5-2022

Experimental Validation of Gene Expression of MYBL1, MYBL2, UBXN8, and ADRM1 Genes in Triple Negative Breast Cancer Cell Lines

Esther Jensinne Nsende

EXPERIMENTAL VALIDATION OF GENE EXPRESSION OF MYBL1, MYBL2, UBXN8, AND ADRM1 GENES IN TRIPLE NEGATIVE BREAST CANCER CELL LINES

THESIS

Presented in Partial Fulfillment of the Requirements for

the Degree Master of Science in the Graduate School

of Texas Southern University

By

Esther Jensinne Nsende, B.S.

Texas Southern University

2022

Approved by

Dr. Audrey Player
Chairperson, Thesis Committee

Dr. Gregory Maddox
Dean, The Graduate School

Approved by

Dr. Audrey Player	03/21/2022
Chairperson, Thesis Committee	Date
Dr. Erica Cassimere	03/21/2022
Committee Member	Date
Dr. Mario Hollomon	03/21/2022
Committee Member	Date
Dr. Tuan D Phan	03/21/2022
Committee Member	Date

EXPERIMENTAL VALIDATION OF GENE EXPRESSION OF MYBL1, MYBL2, UBXN8, AND ADRM1 GENES IN TRIPLE NEGATIVE BREAST CANCER CELL LINES

By

Esther Jensinne Nsende, M.S.

Texas Southern University, 2022

Associate Professor Audrey N. Player, Ph.D., Advisor

A previous study conducted in our laboratory demonstrated V-Myb Avian Myeloblast Viral Oncogene Homolog Like 1 (MYBL1) gene over-expression in triple negative breast cancer (TNBC) compared to normal, some luminal, and a subpopulation of other TNBC. The MYBL1 gene belongs to the Avian myeloblastosis virus (MYB) family and is classified as a proto-oncogene that functions as a strong transcription factor. The MYBL1 gene is related to cancer progression which involves dysregulation of cell cycle signaling, apoptosis and differentiation processes. A primary goal of our laboratory is to further characterize MYBL1 gene expression in TNBC samples. To achieve this goal, we performed a knockdown study to identify genes that co-operate with MYBL1 to affect the phenotype of TNBC. The MDA MB231 TNBC cells were transduced with a short hairpin ribonucleic acid (shRNA) lentiviral knockdown of the MYBL1 gene. When MYBL1 was knocked down, MYBL2 and Adhesion Regulating Molecule 1 (ADRM1) genes were down regulated and UBX Domain Protein 8 (UBXN8) gene was unregulated. Since MYBL2,

UBXN8 and ADRM1 were affected by MYBL1 knockdown, for the current study, we compared the gene expression patterns of MYBL2, UBXN8 and ADRM1 to that of MYBL1 using different methods. Two approaches are utilized to achieve our goal. For approach 1 we utilized polymerase chain reaction and immunohistochemistry to assess RNA and protein expression levels, respectively. For the second approach, we analyzed MYBL1, MYBL2, UBXN8 and ADRM1 transcript levels in TNBC patient samples retrieved from Gene Expression Omnibus. Results from this project should assist in our understanding of MYBL1 in TNBC.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

(ATCC®)	American Type Culture Collection
AMV	Avian Myeloblastosis Virus
Anti-HER2	Antibody therapy targeted treatments
ADRM1	Adhesion Regulating Molecule 1
BL1	Basal-like types 1
BL2	Basal-like types 2
CCNB1	Cyclin B1
cDNA	Complementary DNA
DEG	Differentially Expressed Genes
DBD	DNA binding domain
dT	Oligo dT
DCIS	Ductal Carcinoma in Situ
DMEM	Dulbecco's Modified Eagle Minimum essential media dNTPs Deoxyribonucleic triphosphate
DUSP7	Dual Specificity Phosphatase 7
E2F4	Elongation factor 4
E2F6	Elongation factor 6
ESR1	Estrogen Receptor
EtBr	Ethidium Bromide
G1	GAP 1 phase
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase

GEO	Gene Expression Omnibus
Her2	Human epidermal growth factor receptor
HRP	Horseradish peroxidase
IBC	Inflammatory Breast Cancer
IDC	Invasive Ductal Carcinoma
IHC	Immunohistochemistry
ILC	Invasive Lobular Carcinoma
IM	Immunomodulatory
KD	Knock-down
KIF18B	Kinesin-like protein (kinesin family member 18B) Ki67 Nuclear protein Ki67
LCIS	Lobular Carcinoma in Situ
Luminal A	Hormone-receptor positive (estrogen-receptor positive, progesterone positive HER2 negative).
Luminal B	Hormone-receptor positive (estrogen-receptor positive, progesteronereceptor positive and HER2 positive).
LAR	Luminal androgen receptor
LIN9 LIN9-DREAM component LIN37	MuvB core complex
LIN37-DREAM component LIN52	MuvB core complex
-DREAM component LIN54	MuvB core complex
LIN-54DREAM component	MuvB core complex
Linoo673	Long Intergenic Non-Protein Coding RNA 673
mRNA	Messenger RNA

miRNAs	MicroRNA
MAF1	Repressor of RNA polymerase III transcription
MAF1 M	Mesenchymal group
MSL	Mesenchymal stem-like subtype
MCF7	Luminal breast cancer cell line
MDA-MB231	Triple negative breast cancer cell line
MCF10A	Receptor negative non-tumor cell lines
ul microliter	
ug microgram	
MOPs	3-(N-morpholino) propane sulfonic acid buffer MYB Proto-oncogene, transcription factor
MYBL1 Like 1	V-Myb Avian Myeloblast Viral Oncogene Homolog
MYBL2	Myb-related protein B
NCBI	National Center for Biotechnology Information NFIB Nuclear Factor I B
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PR	Progesterone receptor
PRDM5	PR/SET Domain 5
RANGAP1	Ran GTPase Activating Protein 1
RBBP4	Retinoblastoma Binding Protein 4
RBL1	Retinoblastoma-Like 1 Protein
RBL2	Retinoblastoma-Like 2 Protein
RMA	Robust Multi Array
RNA	Ribonucleic acid

RNAi	RNA interference
RPM	Revolutions per minute
SANT	Swi3, Ada2, N-Cor, and TFIIIB
shRNA	short hairpin RNA
siRNA	small interfering RNA
SLC25A1	Solute Carrier Family 25 Member 1
S phase	Synthesis Phase
STRING™	Search Tool for the Retrieval of Interacting Genes/Proteins
TAQ	Polymerase thermostable DNA polymerase I TBE Tris/Borate/EDTA
TCF19	Transcription factor 19
TNBC	Triple Negative Breast Cancer
TFDP2	Transcription Factor Dp protein 2
TFDP1	Transcription Factor Dp protein 1
TP53	Tumor protein p53
UBXN8	UBX Domain Protein 8
UTSW	University of Texas Southwest Core Facility

VITA

2017 ……	NIH Minority Health and Health Disparities International Research Training Program Department of Biology PI: Dr. Catherine Propper, PhD PI: Dr. Leslie Schulz, PhD Northern Arizona University Flagstaff, Arizona
2017 ……	Summer International Research Assistant Department of Genetics and Biotechnology PI: Dr. Abdelbagi M. Ismail, PhD International Rice Research Institute Los Baños, Laguna, Philippines
2018 …….……	B.S. Biology/Minor: Ethnic Studies Northern Arizona University Flagstaff, Arizona
2018 …….	Poster: *Growth, Physiological, and Molecular Traits Associated with Salinity Tolerance in Rice* Arizona-Nevada Academy of Science University of Las Vegas Las Vegas, Nevada
2021 …….	Publication: *MYBL1 Knockdown in a Triple Negative Breast Cancer Line: Evidence of Down-Regulation of MYBL2, TCF19, and KIF18B Expression* Austin Journal of Cancer and Clinical Research PI: Dr. Audrey Player, PhD
Major Field……..	Biology

ACKNOWLEDGEMENTS

To the strongest person I know, thank you for your endless sacrifice and unconditional support. Thank you for always getting back up and continuing. Mom you are the perfect example of "never give up and follow your dreams". If I can be half the person, you are, I have succeeded in this lifetime. Everything you do for us, does not go unnoticed. If I get to live this life again, Mom I will always pick you! To my siblings, "ONE BAND, ONE SOUND!". This is to show you all that a dream deferred is not a dream denied. To my nieces and nephews, "Aunty did it, so can you!". To my friends, thank you for your constant calls and messages on days when things got tough.

As a young girl, I always imagined myself in places that I was told was "unattainable". I never let what society had to say stop me. My undergraduate experience really shook up that little girl. Being one of the very few brown faces in all my STEM classes was difficult. When all the researchers did not look like me or relate to my experiences, that little girl was shaken. Then, I came across the NIH Minority Health International Research Training Program advertisement. I applied and I got accepted into the program that changed the trajectory of my life. Dr. Propper, Dr. Schulz, Dr. Ismail, and Kathleen Freel made me realize that I must create the spaces I am searching for. I must ask the questions I am passionate about. If I do not ask, then they will not be solved. Thank you for your time and knowledge.

For the constant support and sacrifices you made for us to complete these projects, thank you so very much. When I was not confident in my own abilities, you reassured me and guided me along the way. Dr. Player, thank you for letting me be a part of your vision.

I cannot wait to see what you discover next. Thank you for giving this little girl a reason to keep shooting for the stars.

To my distinguished committee members, Dr. Cassimere, Dr. Hollomon, and Dr. Phan thank you so very much for your time, support, and insight. Dr. Hollomon thank you for growing our cells this past summer. Thank you so much, it did not go unnoticed! I am very grateful! To my lab members, Sierra Cunningham, Christopher Dixon, and Tyra Ivory thank you all for your constant support and kind words. Thank you, Texas Southern University, for giving me the chance to continue to live out my dreams.

CHAPTER 1

INTRODUCTION

Background Information About Breast Cancer and Summary of Breast Cancer Types

The development of adult-onset cancer is an extensive process that involves amendments to the cell cycle, cell growth, cell death, and cellular differentiation that specifically define the disease as cancer (26). The type of cancer is generally defined by the tissue of origin and the progenitor cell type. For breast cancer, the cancers originate in the breast and for *some* breast cancers the progenitor cells are luminal, myoepithelial, and basal-like cells that line the lumen. The assumption is that different progenitor cells along with the contribution of their microenvironments lead to breast cancer heterogeneity.

Breast cancers and all cancers are extremely heterogeneous. Breast cancer can be defined based on their pathological diagnoses and based on molecular characteristics (https://www.breastcancer.org/symptoms/types). A detailed list and short description of breast cancers based on pathology and molecular signatures are listed below. To a large degree, the different molecular diagnoses were defined by DNA microarray analyses. DNA microarrays are a laboratory tool used to simultaneously identify gene expression levels of thousands of genes. Many of the targeted therapies that are being used to treat breast and other cancers are identified by microarray analyses.

The different types of breast cancers are described below, but as a general description, benign tissues are localized to their original site of origin (i.e., in-situ), they are non-cancerous and non-life threatening. The invasive cancers spread (i.e., metastasize) from their primary site to other organs, making them malignant and potentially life threatening. Breast cancers often metastasize to bone, lung, liver, and brain. Recurring cancers are cancers that return at the same or different locations in the body. Just like most other cancers, breast cancers are identified and diagnosed based on their pathology and molecular subtypes (50).

Pathologically Diagnosed Types of Breast Cancers

Even though there are many types of breast cancer in this category, most of them are rare. A detailed list of some o the pathologically diagnosed types of breast cancers include:

- **Cribriform Carcinoma of the Breast -** Cribriform cancers are rare invasive cancers. The cancers are low grade (i.e., look normal), but appear to have 'holes or display cribriform-like configurations.
- **Ductal Carcinoma in Situ (DCIS) -** DCIS is a non-invasive carcinoma which originates in the milk ducts of the breast. DCIS does not metastasize; however, DCIS patients have a significant risk of developing invasive cancers later. It occurs at ~15%.
- **Inflammatory Breast Cancer (IBC) -** IBC is difficult to detect because of it appears as "sheets" rather than "lumps". IBC is an invasive cancer that is very aggressive as well as rare.

- **Invasive Ductal Carcinoma (IDC) -** IDC is the most common kind of breast cancer. It is detected in around 60-70% of breast cancers. IDC is an invasive cancer that occurs in the milk ducts.
- **Invasive Lobular Carcinoma (ILC) -** ILC cancers are invasive cancers that originate in the lobules of the breast. ILC cancers occur in approximately 15% of patients, making it the second most common type of cancer.
- **Lobular Carcinoma in Situ (LCIS) -** LCIS is an abnormal cell growth in the milk glands (lobules) that can signal a higher risk of future invasive cancer. LCIS is also known as a lobular neoplasia (i.e., benign).
- **Male Breast Cancer -** Breast cancer in men is very rare, occurring in less than 1% of men. Although they are rare, they are invasive.
- **Medullary Carcinoma of the Breast -** Medullary carcinomas occur in around 3% of patients. Medullary carcinomas appear as soft, flesh-like masses. The flesh like masses resemble the brain's medulla. Medullary
- carcinomas are rare and invasive ductal carcinomas. These cancers are often 'cluster near TNBC' indicating genetic similarity.
- **Mucinous Carcinoma of the Breast -** Mucinous Cancer is a rare cancer that originates in the milk ducts. It has abnormal cells "floating in a pool of mucin". The mucin then becomes a part of the tumor.
- **Paget's Disease of the Nipple -** Paget's disease is a breast cancer involving cancer cells around the nipple and breast ducts. Paget's disease can be mistaken for dermatitis because skin around the nipple is flaky, and the nipples are inverted and hardened with an appearance resembling eczema.

- **Papillary Carcinoma of the Breast -** Papillary cancers have small finger-like projections or papules and well-defined borders. They are also a rare invasive ductal breast cancer.
- **Phyllodes Tumors of the Breast -** Phyllodes cancers start in the stromal tissue in the breast. The tumor cells grow in a leaf-like arrangement. The cancer happens in less than 1% of breast cancer patients.
- **Tubular Carcinoma of the Breast -** Tubular breast cancer is a subtype of invasive ductal breast cancer. It accounts for less than 2% of all breast cancers. Like other types of invasive ductal cancer, tubular breast cancer originates in the milk ducts of the breast. It then spreads to the tissues around the duct becoming cancerous. Underneath a microscope, tubular carcinomas look like tubes.

Molecular Breast Cancer Types

The prevalence and definitions of molecular breast cancers are presented in Figure 1. A more elaborate description is given below (31). (https://www.breastcancer.org/symptoms/types). Some breast cancers are referred to as luminal or triple negative / basal-like subtypes. These designations are based on the location of the progenitor cells for luminal and triple negative / basal-like cancers Figure 2 (64).

Molecular Subtypes of Breast Cancer are Luminal A/B, Triple Negative/Basal-like, HER2-enriched, and Normal-like

Yersal et al. (71) conducted a study to characterize biological subtypes of breast cancers. Yersal et al (71) classified luminal A cancers as estrogen receptor positive,

progesterone receptor positive, and HER2 negative. Luminal B are classified as estrogen receptor positive, progesterone receptor positive, and HER2 positive/negative (71). The prognoses of Luminal B cancers versus Luminal A cancers are marginally adverse (71).

Triple negative/basal-like cancers are estrogen receptor, progesterone receptor, and HER2 negative (71). While TNBC and basal-like are similar due to gene expression levels, data indicate there is ~25% difference between them (71). A more thorough description of TNBC is explained below. HER2neu-enriched breast cancers are estrogen receptor negative, progesterone receptor negative, and HER2 positive. Normal-like cancers are like Luminal A cancers because they are estrogen receptor positive, progesterone receptor positive, and HER2 negative. Although they appear low grade with low nuclear protein Ki67 (Ki67) levels (71).

Triple Negative Breast Cancer

Transcriptome data analyses suggest the progenitor cells for TNBC are localized to the basement region of the milk ducts in breast, while luminal progenitor cells or localized nearest to the lumen regions (Figure 2). TNBC are defined as a single subtype. However, using DNA microarray and clustering analyses, Lehmann et al. (38) showed that TNBC are a complex, heterogenous subtype that can be further divided into 6 sub-categories. The 6 sub-categories were defined by specific gene expression profiles, and later, mutational studies. Lehmann et al. (38) defined the subtypes as (a) Luminal androgen receptor (LAR) also known as molecular apocrine cancers, (b) basal-like 1 (BL1), (c) basal-like 2(BL2), (d) an immunomodulatory group (IM), (e) a mesenchymal group (M), and (f) a mesenchymal stem-like sub-group (MSL). The LAR subcategory includes estrogen receptor samples that overexpress androgen receptors and hormonal regulatory genes. The

basal-like 1 (BL1) group is enriched in genes associated with the ribonucleic acid (RNA) polymerase pathway, cell cycle and cell division signaling pathways. TNBCs that are basal-like 2 (BL2) originate in the myoepithelium and include genes involved with growth factor signaling processes, gluconeogenesis, and glycolysis. Immunomodulatory (IM) cancers are like medullary breast cancers. They are enriched with genes that include immune signaling pathways, natural killer cell pathways, cytokine signaling pathways, and antigen identification and processing pathways. The M and MSL sub-category include genes aligned with cell motility, proliferation, mesenchymal-like differentiation, and extracellular matrix proteins. Lehmann et al (38) studies were instrumental in defining TNBC. Results from their data led to identification of potential TNBC biomarkers and improvements in predicting patients' response to therapies.

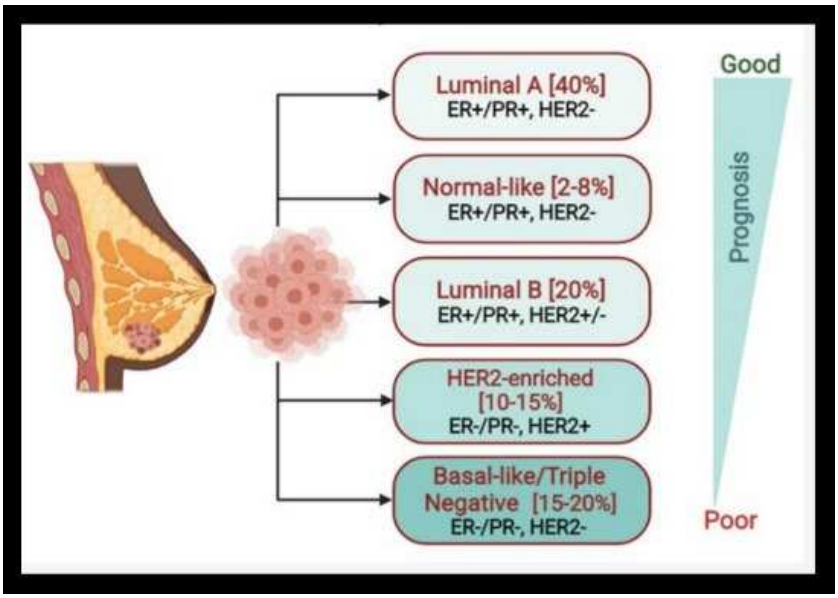


Figure 1: Molecular Signatures and Prevalence of Breast Cancer Types (31)

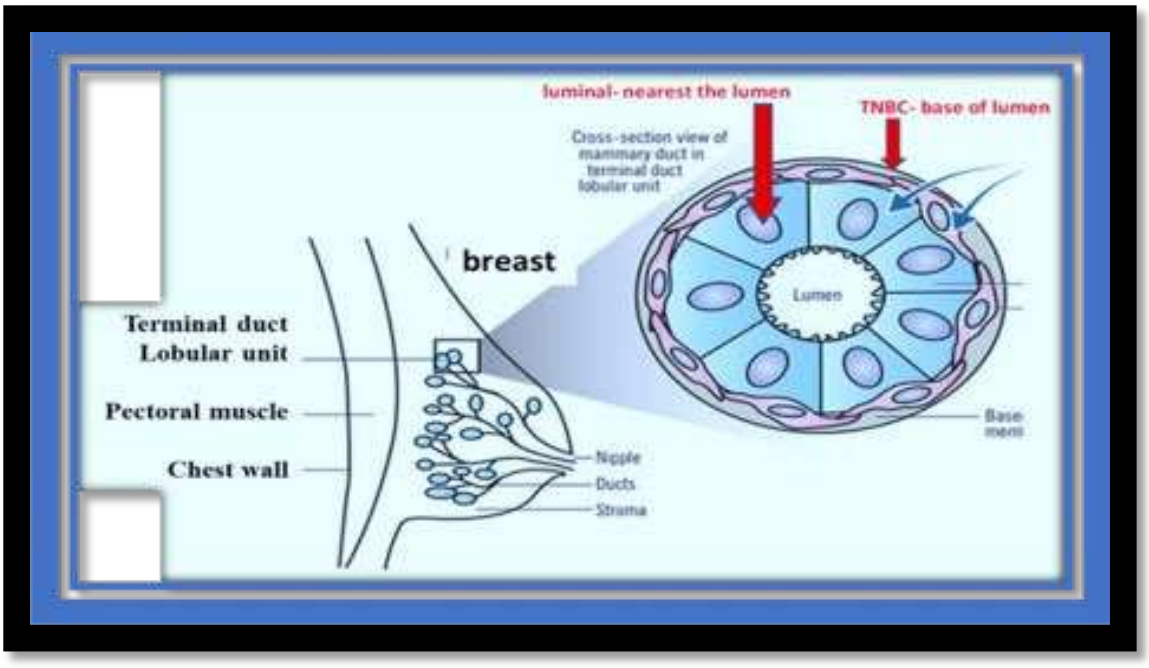


Figure 2: Anatomical Position of Luminal Breast Cancer vs TNBC Progenitor Cells Near Breast Lumen

Breast Cancer Statistics and Other Information Related to TNBC

According to the American Cancer Society, about 1.9 million new cancer cases were diagnosed in 2021 with more than 600,000 estimated deaths (1). In 2021, approximately 281,550 newly diagnosed cases of invasive breast cancer will emerge in women, and about 43,600 women will die from breast cancer in 2021 as reported by the American Cancer Society (3). The American Cancer Society states that 1 in 8 women (13%) will be diagnosed with invasive breast cancer in their lifetime, as well as 1 in 39 women (3%) will die from breast cancer (27). Patients that die are either diagnosed with later stage cancer, or they die due to a lack of therapeutic options (13,51,53). TNBC are basal-like cancers that account for 10-20% of all breast tumor heterogeneities (28).

Approximately 75% of TNBC are basal-like subtypes, which are defined by gene expression profiling (55). Compared to other cancers, the time from initial diagnosis to

death and the time-to metastasis is shorter in patients with TNBC (6,51). In comparison to an 80% survival rate in other subtypes, 20% of the TNBC patient population survive after a 5-year period (54). Because TNBC being negative for all three common receptors, conventional hormone therapies like tamoxifen and anti-HER2 antibody therapies like trastuzumab cannot be used to treat TNBC patients (16). Tamoxifen and trastuzumab are used for patients with positive receptor status, making it ineffective for nearly all TNBC patients (16). There are few therapeutic options for treatment of TNBC patients. Several targeted gene therapies are currently being consider; however, standard care includes chemotherapy and radiation therapy (8). For this reason, it is vital to further characterize TNBC and identify genes that are associated with the tumorigenic process. Once these novel genes have been identified, they can be studied for their possible clinical utility and ultimately used as targeted therapies to affect patient survival.

CHAPTER 2

LITERARY REVIEW

BACKGROUND

Discovery of the MYB Family Genes

The focus of this study is to further characterize MYBL1 in TNBC. MYBL1 is a member of the MYB family of human genes, which also consists of c-MYB and MYBL2. c MYB was the first member of the family to be identified and characterized. The c-MYB gene was identified based on its similarity to the Avian Myeloblastosis Virus (AMV) v-MYB gene sequence (21). AMV is a highly oncogenic chicken leukemia virus which transforms immature hematopoietic cells. Compared to c-MYB, AMV's N-terminal and C-terminal regions are truncated, and the gene contains 16 intragenic point mutations. Using c-Myb as a probe, MYBL1 and MYBL2 were discovered in a complimentary DNA (cDNA) library. All the genes are strong transcription factors that regulate transcript and small RNA expression (39). Sequence analyses reveal both similar and divergent regions in all three MYB family genes. Similarly, the MYB family genes can transcriptionally regulate some of the same and different genes. Each of the MYB family genes is unique in some aspect. Unique to MYBL1, the gene regulates the meiotic process in testis (59) and is over expressed in normal testes during spermatogenesis.

MYB Family Sequences Comparison to Highlight Shared Gene and Unique Gene Expressions

Protein sequences of the MYB family genes share both similar and different regions within their sequence. The genes are similar at the N-terminal DNA binding domain (DBD), their internal transactivation domain and the C-terminal regulatory domain (9). The DBD demonstrates the highest degree of homology, while the transactivation and C-terminal regulatory domain share less homology between the MYB family genes. The NCBI Web CD Search Tool can be used to locate the positions of these regions (44). Collectively the domains are responsible for defining their recognition sites, protein: protein interactions with other transcription factors, coactivators, and sites susceptible to epigenetic modifications. The DBD are an estimated 90% similar between the MYB family genes suggesting that MYB family genes can recognize some of the same genes. Using a reported gene strategy, Rushton et al demonstrated the extent to which MYB family can regulate transcriptional expression of some of the same genes and uniquely different genes. The DBD region is a part of the activation and transcriptional repressor complexes (9). Because the DBD is so homologous between the MYB family genes, investigators suggest that differences in the C-terminal lead to differential regulation of the genes and ultimately their functions. Studies of c-MYB show the N-termini is associated with intra and inter negative regulator functions and truncation of the C-terminal region result in tumors (21).

In addition to the domains noted above, MYB family proteins contain a SANT domain defined by binding to Swi3, Ada2, N-Cor, and TFIIIB genes. SANT domain allows for chromatin remodeling as well as transcriptional regulation. Presence of the SANT regions suggest epigenetic regulation of the MYB family genes.

MYBL1 Expression in Copious Tissues and its Relationship with the Cell Cycle Signaling and Tumorigenesis

Analyses of the MYB family of genes and the genes that they regulate are good resources to examine in search of biomarkers. The genes are good candidates because they are putative oncogenes (73) and are related to processes directly tied to tumor pathogenesis. Several years ago, c-MYB gene was considered a possible therapeutic target for luminal breast because the gene was (a) overexpressed in these tissues (b) vital to growth and cell cycle signaling, and (c) the gene gave rise to tumors in mice (41). Liu et al designed small molecules and applied c-MYB RNA interference (RNAi) strategies to study the use of c-MYB as a target for therapy (42). Liu's experiments were performed several years ago, but today fewer studies are being published touting c-MYB as a target (41). c MYB does not appear to be pursued with the same vigor as earlier.

Our laboratory and others are considering MYBL1 and MYBL2 for their potential clinical utility. Compared to c-MYB, less is known about MYBL2 and even less about MYBL1 associated processes.

The first experiments characterizing the MYBL1 gene showed the genes' involvement in cell cycle signaling. Ziebold et al. (75) conducted one of the first experiments that demonstrated a connection between MYBL1 and cell cycle signaling. Their experiments revealed MYBL1s' involvement in GAP 1 phase (G1) to Synthesis (S) phase progression. Data show that MYBL1 is indirectly affiliated with phosphorylation of cyclin dependent kinases. Marharmati et al. (45) show that MYBL1 works with c-MYB to mediate progression to S phase in smooth muscle cells.

The DREAM or LINC complex is a large assortment of genes that interact in various combinations to regulate cell cycle signaling processes (19,20,56). MYBL1 and MYBL2 are a key part of this complex and subsequent signaling processes. According to

genecards.org, the DREAM complex includes the "elongation factor 4 (E2F4), elongation factor 5 (E2F5), LIN9 complex DREAM MuvB core complex component (LIN9), Lin-37 DREAM MuvB Core Complex Component (LIN37), Lin-52 DREAM MuvB Core Complex Component(LIN52), Lin-54 DREAM MuvB Core Complex Component (LIN54), MYBL1, MYBL2, Retinoblastoma Like 1 Protein (RBL1), Retinoblastoma-Like 2 Protein (RBL2), Retinoblastoma Binding Protein 4 (RBBP4), Transcription Factor Dp1 (TFDP1) and Transcription Factor Dp2 (TFDP2) proteins (19,20,56)". E2F4, LIN37, TFDP2, MY2L1 and MYBL1 genes (of course) were downregulated in our MYBL1 knockdown study (June 2021).

Relationship of MYBL1 to Tumorigenesis

MYBL1 is a putative oncogene which is associated with regulation of cell proliferation, differentiation, and apoptosis, events all of which are hallmarks of cancer. Based on these observations, more and more investigators are beginning to consider MYBL1 as a potential gene to study for its role in cancer. In most of the MYBL1 cancer studies, the gene is over-expressed and the mechanisms leading to dysregulation in many of the cancers is not known. In a few cases the type of mutations has been identified. Several studies in leukemia reveal that changes in expression of MYBL1 gene is triggered by amplifications, rearrangements, and translocation events in tumors (33, 63). The translocations often involve fusions with the Nuclear Factor I B (NFIB) gene (33).

MYBL1 gene is overexpressed in Burkitt's lymphoma (5,23), Triple Negative Breast Adenoid Cystic Carcinoma (63), Cutaneous Adenoid Cystic Carcinoma (37), and TNBC (63). Arsura et al. (5) show survival of lymphomas is maintained by MYBL1 and c-MYB dysregulation at the GI phase of the cell cycle. Their data show the correlation between MYBL1, cell signaling, apoptosis and cancer. Liu et al. (41) examined 181 breast

cancer patient samples using the microarray platform followed by multivariant analyses. The authors identified MYBL1 and 9 additional genes that correlated with poor prognosis in receptor positive patient samples.

Of the 9 genes identified by Liu et al., 3 of the genes were differentially expressed in our current receptor negative MYBL1 knockdown (41,52). The genes include MAF1, Dual Specificity Phosphatase (DUSP7), and Solute Carrier Family 25 Member 1(SLC25A1).

Gorbatenko et al. (25) examined MYBL1 expression levels in basal-like breast samples, normal-like breast samples, and luminal breast samples. Gorbatenko et al. (25) observed high levels of MYBL1 in basal tumors (Figure 4) (arrow). Note that MYBL1 and MYBL2 expression levels in basal-like cancers are similar. Based on gene expression profiles, basal-like cancers and TNBCs are ~75% similar, so we can infer a similar pattern of expression in TNBC. Player et al. observed a similar pattern of overexpression in TNBC patients (53) (Figure 5). For all known genes, Protein Atlas.org gathers RNA, protein staining patterns, pathology and other data related to the gene. More recently the site has included RNA seq profiles generated from cancers. The RNA seq profile for MYBL1 expression in a range of cancer are given in Figure 6. Note MYBL1 levels are highest in breast cancer validating (at least) high levels of the gene in breast cancer. The cancers were *not* examined based on subtype.

MYBL1, MYBL2, UBXN8 and ADRM1 Genes Screened as Part of This Study

It is well established that MYB family genes are co-expressed in a lot of tissues and cancers (24). Studies show that c-MYB can regulate MYBL1 and MYBL2, and MYBL1, c-MYB and MYBL1 are often co-expressed in the same cell lines and cancer patient samples (12, 24, 60).

We examined several GEO microarray datasets where c-MYB was targeted for knockdown or indirectly knocked down in breast cancer cell lines. These data showed that when c-MYB is down-regulated, MYBL1 and MYBL2 are also down-regulated. These data validate the ability of the genes to co-regulate each other. *GEO did not contain MYBL1 knockdown datasets*. The goal of our laboratory is to further characterize MYBL1 in TNBC with the first approach directed at identifying genes that might co-operate with MYBL1 in affecting the cancer genotype in these cells. Towards this goal, the knockdown of MYBL1 is expected to reveal genes that are either directly or indirectly affected by MYBL1 knockdown. This would be consistent with previous data that reveal c-MYB, MYBL1 and MYBL2 can regulate each other. For our knockdown study, we found that when MYBL1 was knocked down, MYBL2 and Adhesion regulating molecule 1 (ADRM1) were downregulated, and UBX Domain Protein 8 (UBXN8) was upregulated. Although the MYB family genes have been shown to be co-expressed, this is the first experiment that showed targeted knockdown of MYBL1 led to down-regulation of MYBL2. *c-MYB is not expressed in MDA MB231 cell line, so c-MYB served as a negative control for assessment of our procedure*. The MYBL1, MYBL2, UBXN8 and ADRM1 expression levels will be assessed as part of this current study.

UBXN8 Gene

In addition to MYBL1 and MYBL2 the UBXN8 gene is being examined to assess its gene expression profile. The UBXN8 gene appeared to be downregulated when MYBL1 was knocked down by shRNA MYBL1 lentiviral transduction, as a result the gene was screened for its pattern of expression as part of the current study. UBXN8 gene is associated with endoplasmic reticulum directed degradation of misfolded proteins. The gene is

localized to both the endoplasmic reticulum and nucleus compartments of the cell. Not much is known about UBXN8, but in certain cancers it functions as a tumor suppressor gene. Over-expression promotes cell cycle arrest and inhibits proliferation and colony forming ability in acute myelogenous leukemia (69). The gene is downregulated by hypermethylation of its promoter region, which in turn leads to leukemogenesis (69). In a separate study, Liu et al. showed that hypermethylation of UBXN8 could be used for diagnosis and prognosis of gastric cancer. One could speculate that UBXN8 functions as a tumor suppressor gene in TNBC because the gene is downregulated while MYBL1 is upregulated. Then when MYBL1 is knocked down, the gene is upregulated activating its suppressor function. We will address this speculation later.

ADRM1 Gene

Adhesion Regulating Molecule 1 (ADRM1) gene is downregulated when MYBL1 was knocked down by shRNA MYBL1 lentiviral transduction, as a result the gene was screened for its pattern of expression as part of the current study. ADRM1 is involved in ATP dependent degradation of ubiquitinated proteins. Like UBXN8, ADRM1 plays a role in removing misfolded, damaged proteins. ADRM1 however functions as a proteasomal ubiquitin receptor (http://www.proteinatlas.org/ENSG00000130706-ADRM1). The gene is up regulated in gastric, colon, prostate, ovarian and breast cancer. Wu et al (67) found that over expression of ADRM1 correlates with poor prognoses in receptor positive breast cancers. Their data show that ADRM1 is a more reliable prognostic marker in estrogen receptor positive cancers compared to receptor negative cancers.

Nonetheless, ADRM1 will be screened for its possible relationship with MYBL1 in TNBC in the current study.

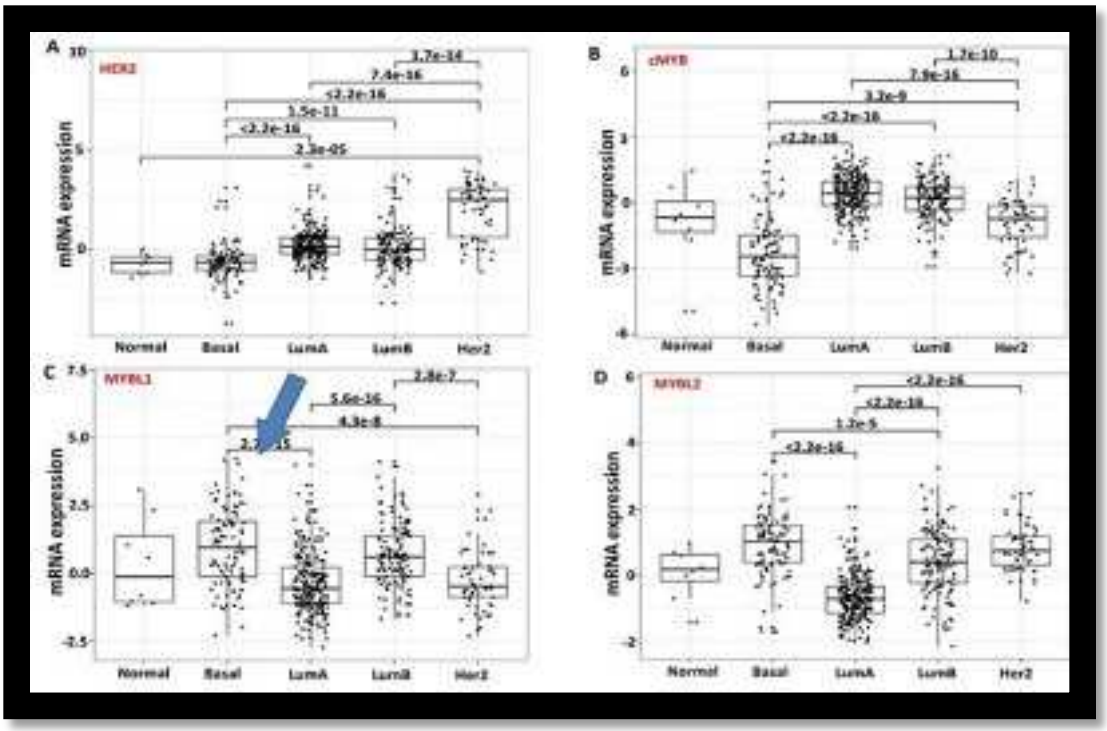


Figure 3: HER2, c-MYB, MYBL1, MYBL2 in Normal, Basal (TNBC) LumA, LumB, HER2 Patient Samples
Arrow points to MYBL1levels in basal (TNBC).

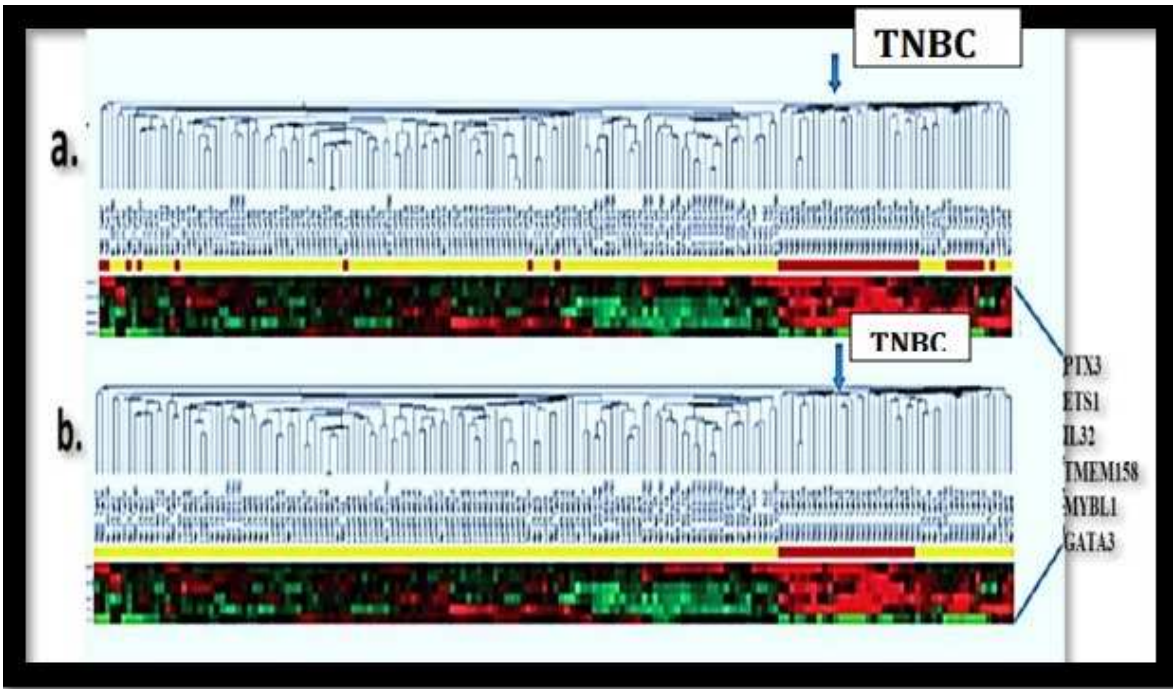


Figure 4: Analyses of MAIRE Patient Samples for Differential Gene Expression of MYBL1 and Other Candidate Genes

TNBC (red bars across the top) are compared to all other sample types (including normal, luminal A/B, HER2neu; yellow). Data shows most (not all TNBC cluster together based on analyses of our 6 gene list. These data led to identification of MYBL1. A and B are the same except the TNBC that cluster with 'all other samples' were removed. TNBC clustering with all others are identified by the individual 'red patient samples that appear mixed with the yellow'. 168 patients are across the top, genes on the slide. Red bars over the red regions designate samples over-expressed for a particular gene. Gene indicates under-expression for a particular gene. Results generated using microarray.

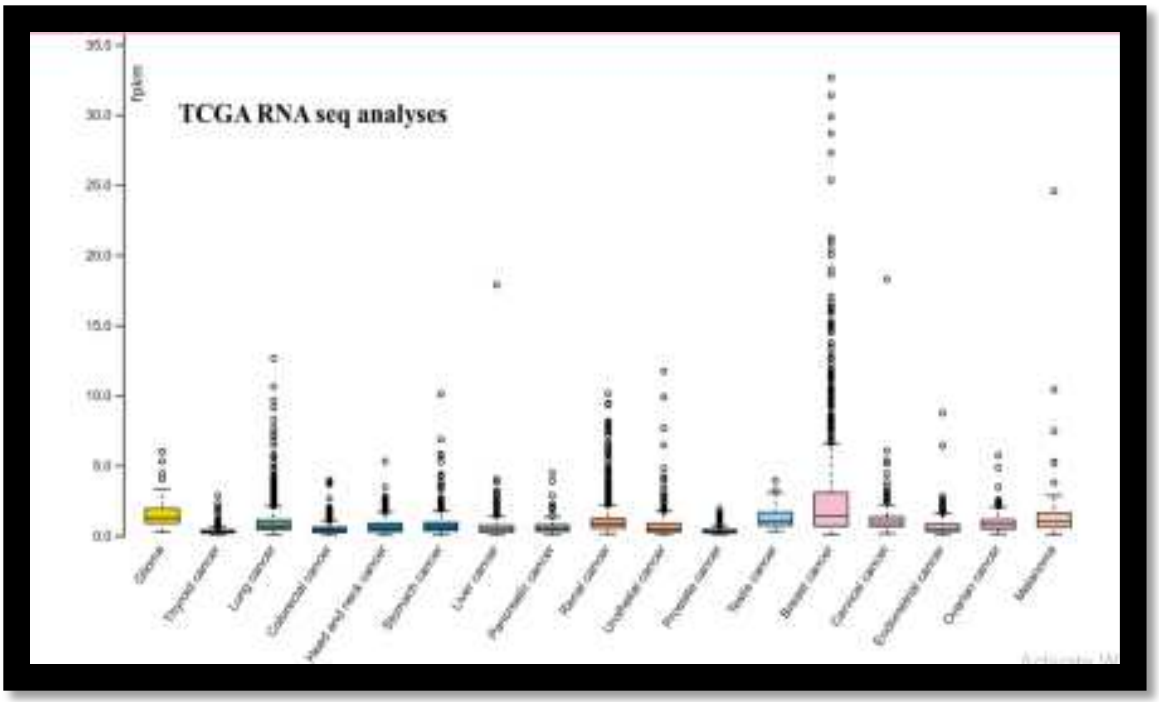


Figure 5: MYBL1 Transcript Levels Determined in Various Types of Cancer

RNA seq data retrieved from ProteinAtlas.org. The cancers were not processed by molecular subtype.

CHAPTER 3

DESIGN OF THE STUDY

MATERIALS AND METHODS

Cell Lines

The cell lines used in this study are MCF7 (luminal receptor positive cells), MDA MB231 (TNBC receptor negative cells), and MCF10A (non-tumor receptor negative cells). The cell lines were purchased from the American Type Culture Collection (ATCC®) atcc.org (Manassas Virginia**)** and used within 1 year of purchase. The cells were fed twice a week using Dulbecco's Modified Eagle Minimum essential media (DMEM) with 1% penicillin and 10% serum and maintained in a 37^{O}C incubator with an additional 5% CO2. The cells were harvested when they reached 80-90% confluency with a 0.25% trypsin solution (Millipore, Sigma, St. Louis Missouri) and either used for experiments or sub-cultured at a lower density.

The patient datasets were obtained from Gene Expression Omnibus (GEO) and included Maire et al (43) (GSE65216), Thorner et al (61) (GSE21371), Muthukaruppan et al. (48) (GSE37820), Richardson et al. (58) (GDS2250), and Farmer et al (18) (GDS1329). The datasets were generated using microarray, so the data contain transcript level analyses.

Ribonucleic Acid (RNA) Isolation

The cell lines in this study were grown in T75 dishes to roughly 90% confluency and harvested by adding 1mlliliter (ml) Trizol (Thermo Fisher Scientific (Waltham Massachusetts) to the culture dish. The sample was collected, and 200 microliters (ul) chloroform was added. The solutions were mixed and centrifuged at 12,000 revolutions per minute (RPM) to separate the aqueous RNA top layer. The aqueous layer was placed in a clean tube with 500ul of 95% ethanol and places at -20°C for 30minutes. RNA was collected by centrifuging the samples for 30 minutes at 10,000 RPM. The RNA pellet was dried, resuspended in 20ul of clean water and heated at approximately 60°C for 1 minute. The purity of the RNA was determined by spectrophotometric analysis. One ul of the RNA solution was added to 9ul of water and the 260/280 absorbance ratio was determined. An A260/280 ratio of 1.8-2.0 represents purified RNA. An aliquot of the RNA mixture was also retrieved for gel electrophoresis. One ul of RNA was added to the RNA sample buffer and heated at roughly 60°C for 1 minute. The 10ul sample was added to a 1% agarose gel. The agarose gel was prepared using of 1 gram of agarose, a 1x solution of 3-(N-morpholino) propane sulfonic acid (MOPs) buffer, 2µl ethidium bromide and 7% formaldehyde. The RNA sample was considered suitable for downstream use if prominent 28S/18S ribosomal RNA bands were observed at equal density.

Generating Complementary DNA (cDNA)

The cDNA was produced using an iScript cDNA kit (Bio-Rad, Hercules CA, USA). A 20 ul mixture containing approximately 1 ug of total RNA, 5x iScript buffer, reverse transcriptase, random hexamers, the Oligo dT mixture, deoxyribonucleic triphosphate acids (dNTPs), and water was combined. The 20 ul mixture was placed at 45°C for an hour. The mixture was then placed at 85°C for ~3 minutes to inactivate the

reverse transcriptase. The mixture was cooled and 80 ul of water was added. The cDNA was placed at –20°C until it was used for polymerase chain reactions (PCR) or -80degrees for long term storage.

Process for Generating Primers

To generate primers for each of the target genes in this study, the primer3 (35) program (http://bioinfo.ut.ee/primer3-0.4.0/) was utilized. The nucleotide sequences for our target genes were retrieved from Affymetrix NetAffx (15) (http://www.affymetrix.com/estore/analysis/index.affx). *So that all our genes could be examined via polymerase chain reaction (PCR) using the same PCR cycle conditions, the default Primer3 conditions were applied. The only parameter that was changed in the program was the amplicon size which ranged from 200-300 nucleotides for each gene*. To determine the specificity of the primer-sets designed by Primer3, the primer-sets were examined using The University of California Santa Cruz Genome database (64) (https://genome.ucsc.edu/) in silico PCR analyses program.

Primer-sets were submitted online to IDTDNA.com (Coralville, Iowa), synthesized, quality controlled at the IDTDNA facility, then shipped to Texas Southern University within 48hours of purchase. Table 3 lists the primer sequences for the genes used in this study.

Polymerase Chain Reaction (PCR)

The PCR procedure was used to process samples for assessment of their differential gene expression levels. PCR was also used to validate the microarray results. Even though the microarray is an invaluable tool, the results ***must*** be experimentally validated. PCR was performed using cDNA. The PCR reactions each contained 2ul (~0.5uM) of forward and reverse primers, 2ul of cDNA (which was generated above),

10ul of 2x thermostable DNA polymerase I TAQ polymerase master mix (optimized for TAQ enzyme, dNTPs and TAQ buffer; Life Technologies, Carlsbad California), and water up to 20ul. The primers generated for each gene were specific for that gene. The samples were placed in PCR quality tubes and positioned in the Bio-Rad Thermal Cycler (Hercules California). The cycler conditions were (a) 5 minutes at 95°C degrees (b) 30-32 cycles for 30 seconds at 95°C, followed by 30 seconds at 58°C degrees, then 30 seconds at 78°C.

Gel Electrophoresis

The gel electrophoresis was performed to analyze the PCR products using a 2% agarose gel. The agarose gel contained 2 grams of agarose added to 100 ml of 1X Tris/Borate/EDTA (TBE) buffer. The mixture was heated, cooled, and an additional 1ul of Ethidium Bromide (EtBr) was added prior to pouring the gel mixture in the chamber. So that direct comparisons between transcript levels could be determined, precisely 10 ul of the PCR product and 2ul of sample buffer were loaded onto the gel for electrophoresis.

Densitometer

LI-COR Imaging System (Lincoln Nebraska) it was used to visualize the PCR products on the gel. The densities of the amplicons were assessed using the LI-COR software. All the values were normalized in comparison to the control gene.

Protein Interaction Analyses

The Search Tool for the Retrieval of Interacting Genes/Proteins (STRING™) program (60) was utilized to examine and demonstrate the protein interactions between the candidate genes. STRING™ analysis relies on the interrogation of millions of data points that was produced using experiments, theory, and published studies. STRING™ analyses were also utilized to determine protein interactions between MYLB1, MYBL2, UBXN8, and ADRM1.

Immunohistochemistry Staining

A low-density paraffin embedded breast cancer TNBC microarray was purchased from BioMax.US (BR498) for this study. The tissue array was deparaffinized by placing the slide in xylene for 15 minutes, followed by rehydration in 100% alcohol for 15 minutes, 95% alcohol for 15 minutes, 70% alcohol for 15 minutes, and clean water for 15 minutes. The slides were rinsed in PBS for 5 minutes. Antigen retrieval was performed by adding the tissue slide to a boiling hot 1x citrate solution (H-3300-250, Vector Laboratory, Burlingame CA) for 20 minutes. The slide was rinsed in PBS for 5 minutes, a PAP blocking pen was applied, then the antibody blocking serum (normal goat or horse serum; supplied with PK-6000; ABC HRP peroxidase staining kit, Vector Laboratory) was added and the slides incubated at room temperature for 30 minutes. Excess serum was removed, and the tissues were incubated overnight at 4^{o}C with their corresponding antibody. ADRM1 human anti-mouse was purchased from Santa Cruz Biotechnology (Santa Cruz CA; sc-166754) and used at a 1:1000 dilution. CD31 blood vessel anti human mouse antibody control was purchased from Santa Cruz Biotechnology (sc-43411) was used at a 1:1000 dilution. Four micrograms per sample of MYBL1 antibody (anti human, mouse antibody; Millipore Sigma, HAP008791) was added to the tissue. Antibodies were diluted in normal horse serum.

The next day, slides were rinsed twice, 5minutes each in PBS/Tween, then incubated with the biotin-conjugated universal secondary antibody at room temperature (RT) for 30 minutes. The slides were rinsed as before and incubated with the SABC reagent at RT for 30 minutes. The tissue was washed and ‘developed’ by adding the DAB peroxidase substrate solution (Vector Laboratories: SK-4600). The samples were allowed to develop and stopped once a desired color was obtained. The reaction was stopped by

placing the slide in distilled water. The slides were briefly counterstained using hematoxylin. The slides were dehydrated by incubation in clean water, 70% alcohol, 95% alcohol, 100% alcohol (each for ~3minutes) followed by xylene for 15 minutes. The slides were dried and permanent mounted.

CHAPTER 4
RESULTS AND DISCUSSION

This chapter discusses the materials and procedures performed prior to this study but critical to interpretation of the results. The data generated as part of an earlier study: *shRNA knockdown of MYBL1 in MDA MB231 cells.*

The knockdown was performed prior to this proposal, however because it was a critical part towards generating the candidate genes studied here, the procedure for achieving the knockdown were described here in Figure 6. The aim of the knockdown study was twofold. The first aim was to further characterize TNBC. Being that MYBL1 was a gene of interest, the second aim of the knockdown study was to determine genes that were directly/indirectly associated with MYBL1 in TNBC. As summary of the procedure, the MYBL1 shRNA lentiviral particles and the scramble control particles were purchased from Origene (Cat # TL303089V; Rockville Maryland). We purchased four MYBL1 target specific particles (packaged from the pGFP-C-shLenti vector; labeled LVA, LVB, LVC, LVD) along with the control from Origene. We transduced each of the lentiviral preparations into MDA MB231 (at a MOI of 10:1) and screened for their ability to decrease MYBL1 levels in the TNBC cell line. MDA MB231 cells were incubated with the targeted or scrambled viral particles for approximately 72 hours with polybrene (sc-134220; Santa Cruz Biotechnology, Dallas Texas) in complete cell culture media.

.

Following removal of the lentiviral particles, fresh media was added to the cells along with 1ug/ml puromycin (CAS 53792; Santa Cruz Biotechnology, Dallas Texas). The LVA lentiviral corresponding to the CTGATCCTGTAGCATGGAGTGACGTTAC sequence, demonstrated the highest downregulation of the MYBL1 mRNA. As a result, the LVA particles were used for future experiments.

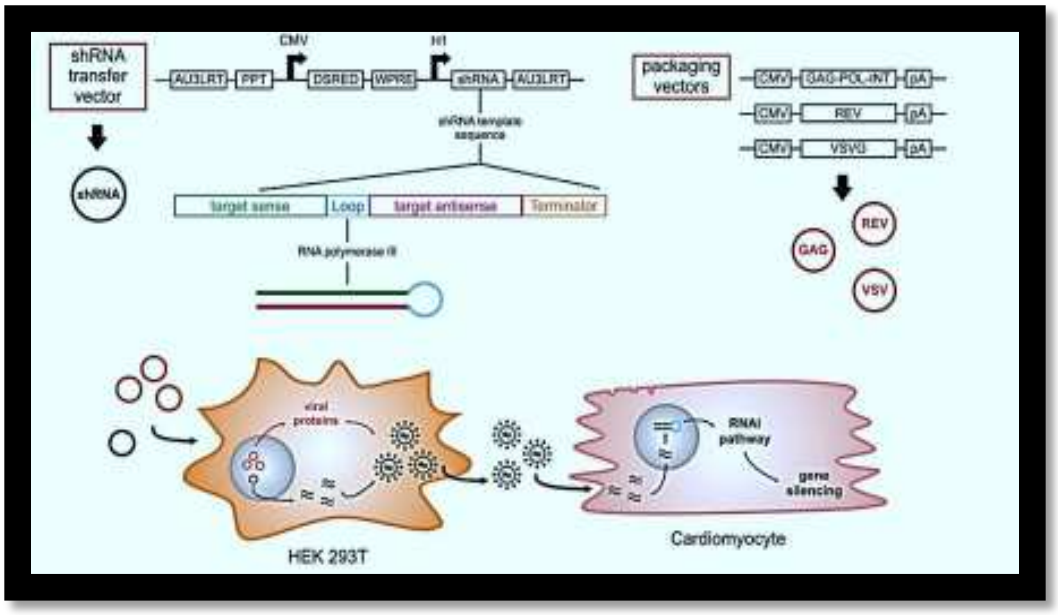


Figure 6: Synopsis of the shRNA Procedure
Transgene of interest and the necessary cis elements for RNA production and packaging are transfected into HEK cells. Before being sent to customers for transduction, the packaged particles are generated and harvested (10). In this figure, HEK293T cells are used to generate viral particles that are transduced into cardiomyocyte cells.

RNA and protein analyses of the LVA shRNA MYBL1 particles demonstrating knockdown of MYBL1 RNA (a) and protein in cell line preparations are shown in Figure 8. RNA was processed as described in the Methods section above. Protein expression levels were determined using Western Blot analyses.

Western Blotting

The western blot procedure was performed as referenced in our publication (June 2021). As documented, the antibodies used for the assay are described below: *Antibodies*: Both antibodies described below were generated as mouse monoclonals. Actin was used at a $1{:}10^4$ dilution (NB600-501SS; Novus Biologicals LLC, Littleton CO), and MYBL1 was used at a 1:500 dilution (sc-514682; Santa Cruz Biotechnology, Santa Cruz CA). Secondary HRP conjugated Anti mouse antibody (HAF007; R and D Systems, Minneapolis, MN) was used at a dilution of 1:4000. Western blotting results were visualized with the Clarity Western ECL substrate (Bio-Rad, Hercules, CA, USA) on a LICOR digital imaging system (LI-COR Biotechnology, Lincoln, NE).

Microarray and Data Analyses

RNA was purified from a scrambled and MYBL1 knockdown preparations of MDA MB231. The RNA was shipped overnight to the University of Texas Southwest Core Facility (UTSW; Dallas Texas, USA). The UTSW core facility prepared the RNA and hybridized it to the Affymetrix Clarion microarray gene-chip which includes approximately 186,000 probe sets (i.e., transcripts, splice variants, siRNA and snoRNA) (see figure 7). The facility provided our laboratory with the hybridization results. The data analyses were conducted (at TSU) by Dr Player utilizing the Affymetrix TAC 4.0 software (Thermo Fisher Scientific, Waltham Massachusetts). The data were normalized utilizing the Robust Multi Array (RMA) program. Utilizing the Limma Bioconductor analysis, the differentially expressed gene levels were produced. If the probe-sets displayed at minimum a 4-fold difference in gene expression between a gene in the MYBL1 LVA

microarray dataset compared to a gene in the scramble microarray dataset, the gene was chosen for further analysis. A 2-fold difference is the industry standard, but the higher the cut-off, the more reliable the resulting data. Gene Ontology analysis (46) was utilized to interrogate the differentially expressed candidate genes. Microsoft Excel was used to create transcript plots and analysis. Molbio-tools (http://www.molbiotools.com/listcompare.html) was used to compare several differentially expressed gene-lists.

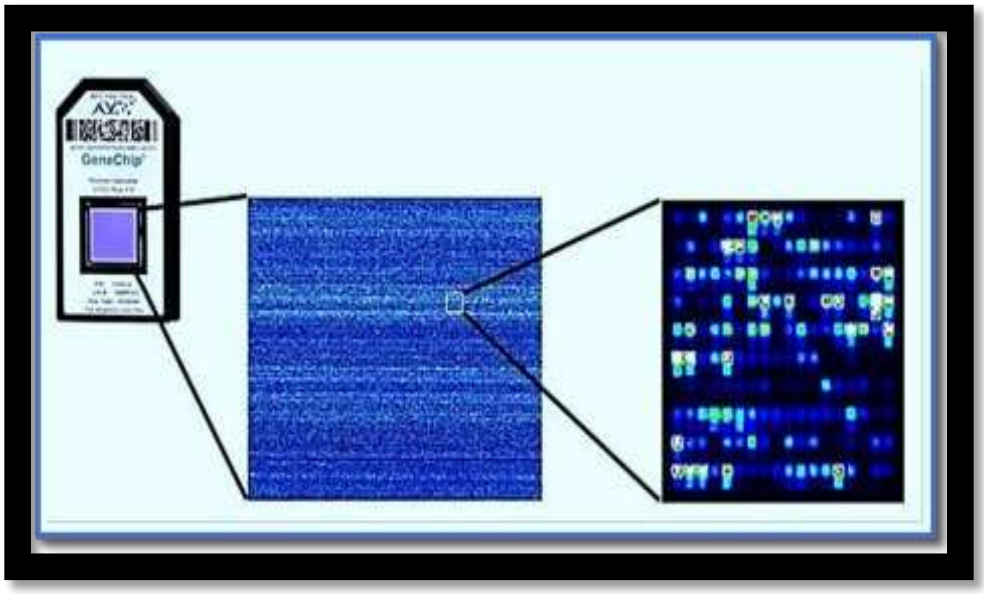


Figure 7: The Affymetrix Microarray Genechip

Immobilized sense strands complementary to the target sequence are on the genechip. Gene-chips containing either 56,000 transcripts or 186,000 transcripts, splice variants and small RNAs displayed as blue dots in the middle figure. Total RNA is labeled and hybridized to the genechip. High copy number corresponds to intense probe-set signals; displayed as lighter spots in the far-right magnified pane. Control sequences at known concentrations are also supplied on the gene chip and used to determine copy number (15).

Based on differential gene expression in microarray and comparison to existing cell line and patient sample datasets, 19 genes were selected for our final list of candidate genes. MYBL1 (which was knocked down) is a strong transcription. As a result, we asked 'of the genes on our list, which ones are direct transcriptional targets of MYBL1'? Online analyses using the ChipX (15) high density transcription factor binding libraries can address this question. ChipX interrogates a larger number of transcription factor (TF) library databases each containing information from thousands of TF experimental assays. Each TF experimental assay assesses the binding of > 1600 well defined transcription factors and accessory proteins. The data presented in Table 2 – was not experimentally validated (by binding assays) but the suggestion was '*our 19 candidate genes show concordant enrichment of MYBL1 binding'.* MYBL1 ranked 11th out of 1632 transcription factors placing our gene list at the top (0.07%) of the transcription assessment list for enrichment of MYBL1 binding. MYBL1 did not show evidence of direct binding to (the promoter of) either UBXN8 or ADRM1, but there was enrichment of NPAS1 transcription factor binding ADRM1. NPAS1 was not identified in our MYBL1 knockdown. Hundreds of genes were identified as differentially expressed. It could very well be that MYBL1 regulates some gene on our list that directly (or even indirectly) regulates NPAS1. Or, it could be 'noise', which is why all experiments must be validated using lots of different datasets and experimental validation platforms.

Table 1: shRNA MYBL1 Target Sequences

Supplied by Origene.com. Each was transduced into MYBL1 and assessed for knockdown of MYBL1 transcript. LVA was most effective at decreasing MYBL1 transcript.

MYBL1 Lentiviral	SEQUENCE
TL303089VA	TCTGATCCTGTAGCATGGAGTGACGTTAC
TL303089VB	CTTGTAATGGTGGCAACAGTGAAGCTGTT
TL303089VC	TAGCACTCCACCAGCCATCCTCAGAAAGA
TL303089VD	CAGGCACTCAACTGTTGACTGAAGACATT

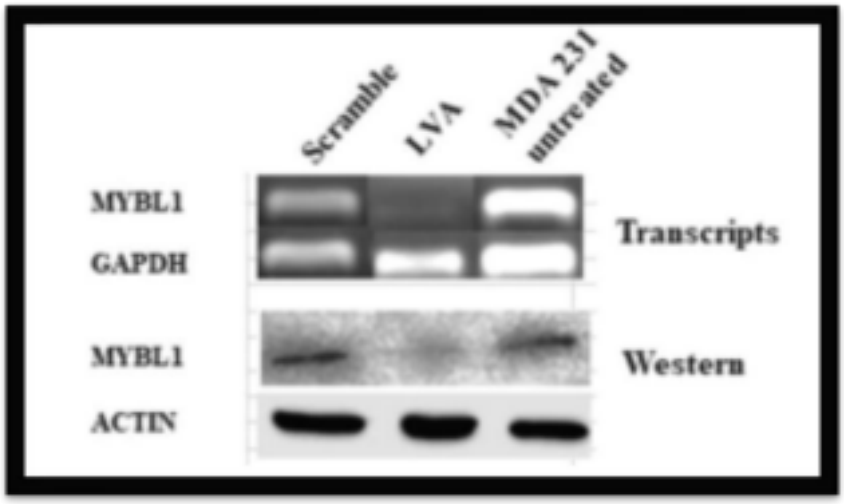


Figure 8: Analyses of LVA Sequence Ability to Knockdown Transcript and Protein Levels

PCR and Western blot were used to compare MYBL1 levels in MDA MB231 cells transduced with scramble compared to LVA particles compared to untreated MDA MB231 cells.

Table 2: Transcription Factor Enrichment Analyses Using ChipX

ChipX program was used to interrogation our candidate genes for transcription factor binding enrichment. Our gene list is enriched for several genes, but also enrichment for binding

Rank	Transcription factor	Score	Library	Overlapping genes
1	FOXM1	6.14E-04	ARCHS4 Coexpression	TCF19,KIF18B,MYBL2,HYOU1
2	ZNF107	6.22E-04	GTEx Coexpression	TCF19,MYBL2,MOB3A,TAPBP
3	ZNF107	7.12E-04	Enrichr Queries	TCF19,KIF18B,MYBL2,MYBL1
4	TP53	0.001229	ARCHS4 Coexpression	TCF19,GABARAPL1,ERCC2,MYBL2
5	ZNF331	0.001245	GTEx Coexpression	RAB3B,GABARAPL1,MAMLD1
6	ZNF367	0.001425	Enrichr Queries	TCF19,KIF18B,MYBL2,MYBL1
7	NPAS1	0.001843	ARCHS4 Coexpression	ADRM1,ERCC2,MYBL2
8	ZHX3	0.001867	GTEx Coexpression	GNPDA2,GNA12,TAPBP
9	CENPT	0.002137	Enrichr Queries	TCF19,KIF18B,MYBL2,TGM2
10	E2F7	0.002457	ARCHS4 Coexpression	TCF19,KIF18B,MYBL2
11	MYBL1	0.002849	Enrichr Queries	TCF19,KIF18B,MYBL2,TGM2

The research goal of our laboratory is to ultimately study genes that are key to the pathogenesis of TNBC. At present, that gene is MYBL1, and toward that goal, we are identifying and characterizing genes that associate with MYBL1 in TNBC cells. The research described in this study totally revolves around analyses of MYBL1 and validating UBXN8 and ADRM1 to determine if the genes might be in some way directly or indirectly associated with MYBL1 in TNBC. The way to accomplish this is to compare the gene expression levels of the two genes to MYBL1. These validations will determine whether UBXN8 and ADRM1 are worth further studies. For UBXN8, currently, the answer is no.

Data presented in this section examined the gene expression patterns of UBXN8 and ADRM1 and compared those patterns to that observed in the knockdown study. For example, data show that ADRM1 follows a pattern of expression concordant with MYBL1; and UBXN8 follows a pattern of expression inverse to MYBL1. These data were

generated using the DNA microarray, so their profiles were further examined and experimentally examined using PCR and IHC protein analyses. In addition, the gene expression levels were compared to patient sample datasets.

According to protocol, these types of analyses should be sufficient to draw conclusions about the utility of our candidate genes. This study is separated into Aim 1 and Aim 2. Aim 1 involves experimental validation of the RNA and protein levels of our candidate genes, using PCR and IHC. For Aim 2 we will compare our candidate genes (MYBL1, MYBL2, ADRM1, and UBXN8) to archived breast cancer patient sample profiles. The patient datasets contain transcript level data. These validations were used to determine whether we would continue studies of UBXN8 and ADRM1 genes.

Choice of Cells Lines Used for the Comparative Analyses in the Current Study

There are at least 92 breast cancer and non-tumorigenic breast cell lines available for studying different aspects of breast cancer (14). Many of these cell lines have been thoroughly defined and characterized based on their origin, morphology, invasive potential, and molecular signatures. Because of the incredible heterogeneity of the breast cancer subtypes, choice of the cell lines to use for a particular study is important and can be a daunting task. For this study, the MDA MB231 was chosen as the TNBC knockdown because of its high levels of MYBL1 and it is well defined. The MCF7 luminal cell types (i.e., MCF7) are used in comparative assays because even though the cell expresses MYBL1 it is receptor positive and to a large degree, the cells express signatures that differ from MDA MB231 and there is a tremendous amount of microarray data available to refer to for validation of our experimental results.

The MCF10A was used for many of the same reasons as the MCF7, except the MCF10A cell line is a non-tumorigenic basal-like triple negative cell line and it serves as a suitable non-tumor source. Search of PubMed, Gene Expression Omnibus, and a myriad of other databases (and journals) you will find experimental comparisons using MCF7 and MDA MB231. We use many of these resources as validation of our bioinformatic and experimental results.

AIM 1 RESULTS SUMMARIZED: PCR AND IHC *[EXPERIMENTAL VALIDATION]*

Question: Does ADRM1 follow a pattern of gene expression like that observed for MYBL1; and does UBXN8 follow a pattern of gene expression inverse to that observed for MYBL1? In aim 1, we compared the gene expression pattern of UBXN8 and ADRM1 to MYBL1 via PCR (for RNA comparison) and IHC (for protein level comparison). The original microarray gene expression profiles extracted from the knockdown study are presented in Figure 9. These results led to our interest in UBXN8 and ADRM1. Data showed that when MYBL1 was knocked down, then UBXN8 was upregulated and ADRM1 was downregulated along with MYBL2. The UBXN8 gene is a tumor suppressor gene (Ts) (69) in several cancers, so we were interested to see the gene upregulated when MYBL1 was knocked down. Did knockdown of MYB1 lead to a more 'non-tumor' genotype, activating the tumor suppressive function of UBXN8? We could speculate that in TNBC, the pathogenesis of TNBC is driven by upregulation of MYBL1 and activation of the Ts function of UBXN8 (i.e., downregulation).

The choice to examine ADRM1 with UBXN8 together because both genes were affected by MYBL1 knockdown, and both are functionally related to dysregulation in protein-folding mechanisms. ADRM1 is currently being studied as a biomarker for luminal breast cancers and although the gene is upregulated in TNBC, it is not being considered for a role in TNBC (66). Nonetheless, it demonstrates a pattern consistent with MYBL1 and that is criteria for studying the gene. Our TNBC cell line (MDA MB231) is negative for c MYB, ESR1, PR and ERBB2 (i.e., Her2), and these genes were negative in the scramble and LVA, so they were included in the figure as negative controls and validation of our protocol.

PCR RNA Expression Analyses

The primer sequences and amplicon sizes for the genes examined in this study are listed in Table 3 and the PCR gel is displayed in Figure 10. Densitometer analyses of the PCR results are presented in Figure 11. If we first compare the PCR results to the knockdown data, gene expression of MYBL1, ADRM1 and MYBL2 in MDA MB231 are consistent with the knockdown results. Each of the genes show a similar direction of expression as observed for MYBL1 in MDA MB231 (which was the cell line used for knockdown). That is not the case for UBXN8; high levels of the gene are detected in UBXN8 in MDA MB231 cells. One would expect low levels of UBXN8 in these untreated TNBC, and higher levels in the MCF10A non-tumor triple negative cell line, which we did not see. Because results of the UBXN8 gene are not consistent, it is not considered a strong candidate for future consideration. ADRM1 and MYBL2 'follow' a similar pattern as MYBL1 (as observed for the knockdown).

IHC Protein Expression Analyses

TNBC tissue array was used to examine MYBL1, ADRM1 and CD31 control levels in invasive carcinoma patient samples. The tissues were a bit old (i.e., lacking optimal protein expression) but some were suitable for analyses. ADRM1 and MYBL1 were detected in a few of the same samples, but there was one patient sample with high ADRM1 and negative for MYBL1; this sample appeared well differentiated (Figure 12). In another patient sample, ADRM1 and MYBL1 protein was detected, but ADRM1 protein expression appeared to be expressed in different cells compared to MYBL1, not different cell types (Figure 13). This sample appeared poorly differentiated. This doesn't mean that ADRM1 is not a good candidate, it's just not the best candidate for being possibly associated with MYBL1 in TNBC. Still, ADRM1 is worth studying further.

AIM 2 RESULTS SUMMARIZED: GEO PATIENT DATASET COMPARISONS ***[BIOINFORMATIC ANALYSIS]***

In aim 2, we compare our candidate genes MYBL1, MYBL2, ADRM1, andUBXN8 against archived breast cancer patient sample profiles. In our laboratory, we consider patient samples the 'gold standard'; nothing is more important than the pattern observed in real women. Our patient sample profiles are from GEO GDS2250, GDS1329, and the MAIRE dataset. Patient samples used for my study include all subtypes. None, however, contain TNBC sub-categories; these samples are difficult to find. The patient samples include normal-like, luminal A/B, and basal-like and TNBC patient profiles.

Analyses of GDS2250

The transcript levels detected for GDS2250 are given in Table 5. A plot of these data is in Figure 14. MYBL1, MYBL2, ADRM1, and UBXN8 transcript levels were examined in archived datasets created from breast cancer patients and breast cancer cell lines. The GEOGDS2250 dataset includes normal-like, luminal-like, and basal-like/ TNBC patient samples (Figure 14). We generated the mean for (a) normal samples and (b) mean values for all cancers combined and presented the values as a ratio. Comparing normal to all samples: MYBL1=253/7224; MYBL2=131/1685; ADRM1=5641/1307; UBXN8=326/353. The normal breast patients demonstrate high levels of ADRM1 even though the trend line shows upward projection. A few highly expressed normal samples screw these results. The ADRM1 is being studied as a biomarker for luminal cancer, but the authors conclude the gene shows little utility for TNBC. There was not total concordance with MYBL1 because some cancers express ADRM1 and not MYBL1 and vice versa, like the pattern observed in IHC.

Also see the RNA Seq data See Figure 5 (generated by ProteinAtlas.org). Nonetheless, the normal values are quite high for this gene in these patient samples. The MYBL1 and MYBL2 genes show a differential pattern of expression. First comparing the patient data values with the knockdown- For example MYBL1, MYBL2 ADRM1 and UBXN8 to levels in TNBC only. MYBL1, MYBL2 and ADRM1 levels were high in TNBC, but the value of ADRM1 is negated based on its expression in normal samples. For UBXN8, the gene expression levels were low in all the samples. If UBXN8 were a Ts gene we would expect high levels in non-tumor, which we did not observe. So, the

UBXN8 is not a good candidate to pursue, and the higher-than-expected values of ADRM1 in normal make it a less than desirable gene candidate. But the gene is worth a second look in other datasets. If UBXN8 functions as a tumor suppressor in TNBC one would expect down-regulation in the cancers and up-regulation upon silencing of MYBL1. This was not the case. UBXN8 will be excluded from further studies.

Analyses of Maire Dataset

For the MAIRE analyses only normal and TNBC were analyzed (Figure 15). The data show that MYBL1, MYBL2 and ADRM1 show a similar pattern of expression like that observed for the GDS2250 dataset. Also, high gene expression levels were observed for ADRM1. There were some samples that express ADRM1 and not MYBL1, like the IHC data; although because of the limited number of IHC data point, it would be difficult to make this comparison.

Cluster Analyses of all Maire Datasets

Even with high levels of ADRM1 and UBXN8, there were differences between the normal and tumor based on cluster analyses (Figure 16). These are likely driven by MYBL1 and MYBL2 and higher values of ADRM1.

STRING Analyses of MYBL1, ADRM1, UBXN8 and MYBL2

The String analyses was performed to examine a possible relationship between UBXN8 and ADRM1 because both genes have been showed to associated with protein misfolding events. There are no published, theorized, or experimental data demonstrating a relationship between these 2 genes (Figure 17).

Table 3: List of Primer Sequences Utilized in the Study

GENE	LEFT PRIMER	RIGHT PRIMER	SIZE OF MPLICON (BP = BASE PAIRS)
MYBL1	AAGTCTGGGCTTATTGGACATAA	TGCAAGTATGGCTGCTACATG	202BP
MYBL2	GAGGGGGTCTGTGAATCTGA	CCATCCTAAGCAGGGTCTGA	265BP
c-MYB	CTTGTTTGGGAGACTCTGCA	TGCAAACACAGGATCCATGC	227BP
ESR1	ACTTGTCCCATGAGCAGGTG	CAAAGCTGCGACAAAACCGA	272BP
ERBB4	ATGCCAATTGTGTGTGGTGT	TTCTTTCCCAAGAGCCAAAA	272BP
GAPDH	TCC CTG AGC TGA ACG GGA AG	GGAGGAGTGGGTGTCGCTGT	217BP
PGR	GTCAGTGGGCAGATGCTGTA	TGTGAGCTCGACACAACTCC	293BP
UBXN8	AGTCGCTGGAGGACATAGGA	AAAATGGCACAGTCCACAGA	266BP
ADRM1	CTGCTTCCCTACTTGCCATC	TCGTCCTTCTTGTCCTTCGT	272BP

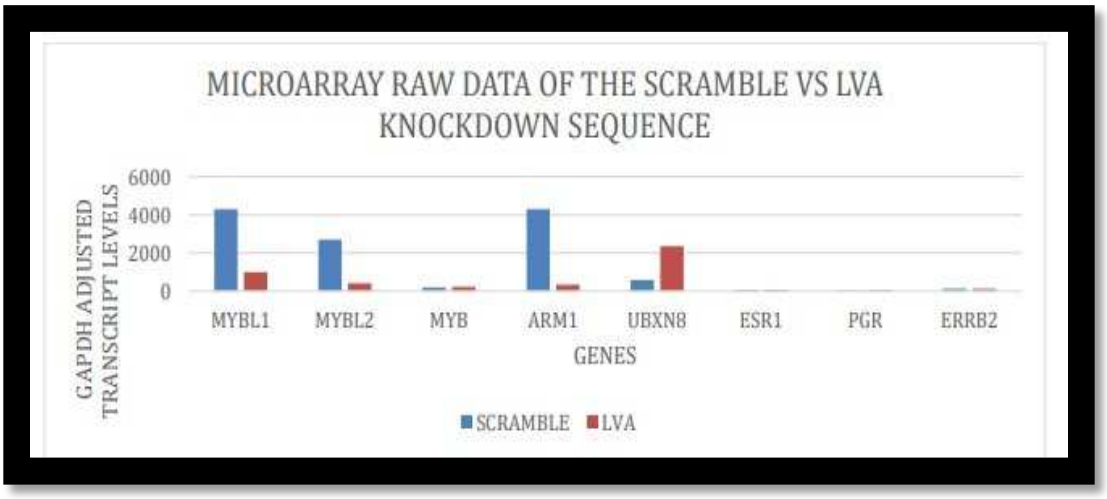


Figure 9: Microarray Data of The Control (Scramble) Vs the Selected Sequence LVA

Table 4: MICROARRAY shRNA knockdown RAW DATA (Gene descriptions from genecard.org)

GENE	AFFY PROBE ID	FOLD CHANGE C/LVA	GENE DESCRIPTION
MYBL1	213906_at	-4X	MYBL1 IS A PROTEIN CODING GENE THAT IS ASSOCIATED WITH ADENOID CYSTIC CARINOMAS AND MITOTIC PROPHASE.
MYBL2	201710_at	-6X	MYBL2 IS A TRANSCRIPTION FACTOR THAT IS INVOLVED IN CELL SURVIVAL, PROLIFERATION, AND DIFFERENTITAION.
ADRM1	201281_at	-5X	A DRM1 ENCODES A MEMBER OF THE ADHESION REGULATING MOLECULE 1 PROTEIN FAMILY. THE ENCODED PROTEIN IS A PART OF THE PROTEASOME WHERE IS BEHAVES LIKE A UBIQUITIN RECEPTOR ADRM1 IS ASSOCAIATED WITH CARCINOGENESIS.
UBXN8	215983_s _at	+4X	UBXN8 IS A PROTEIN IS LOCATED IN THE ENDOPLASMIC RETICULUM (ER) MEMBRANE. ASSOCAITED DEGRADATION OF MISFOLDED PROTEINS.

Figure 10: PCR Visualized Using Densitometer for Numerical Values

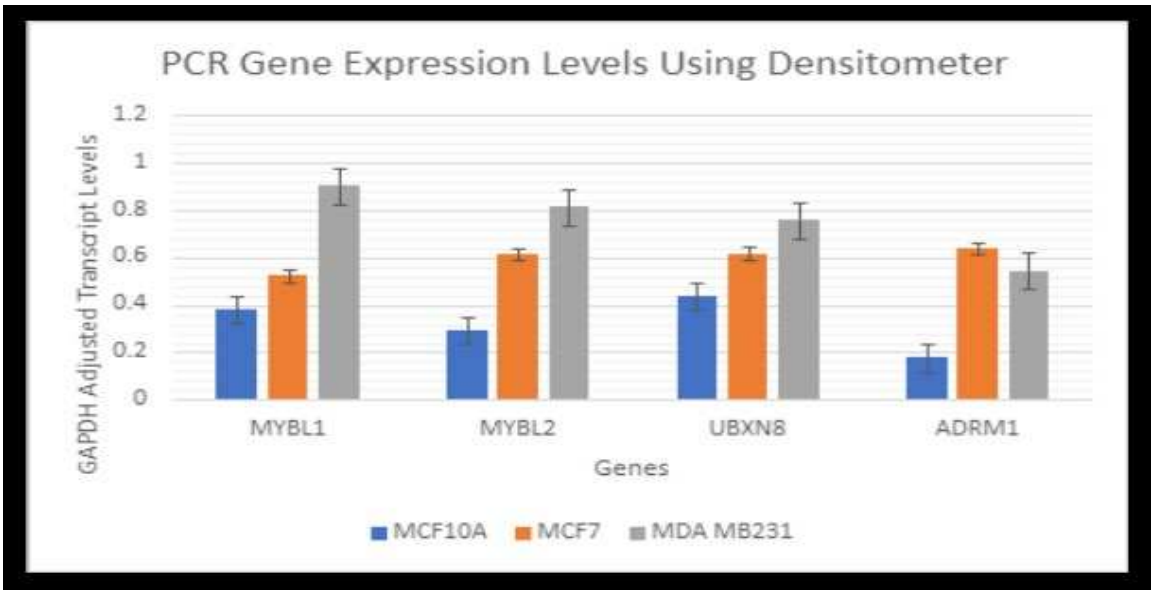
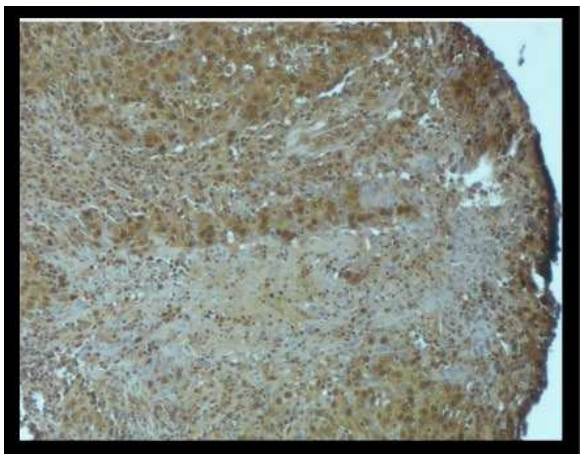
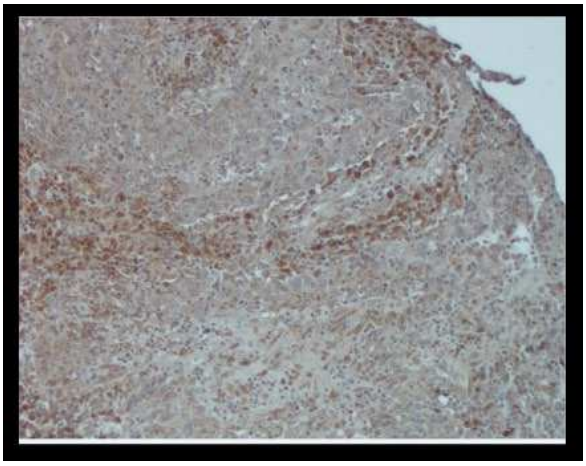


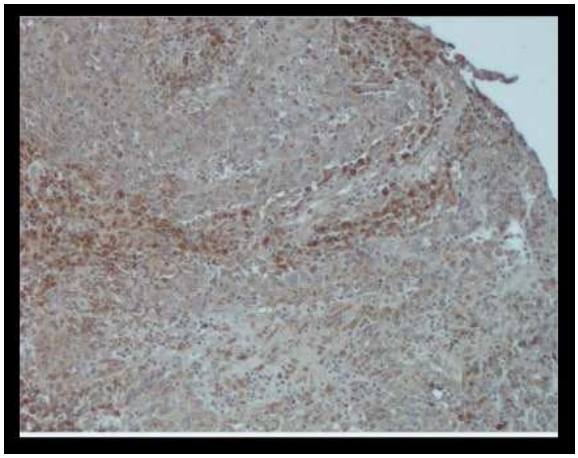
Figure 11: Graph of the Visualized PCR Densitometer for MYBL1, MYBL2, ADRM1, and UBXN8



A. ADRM1



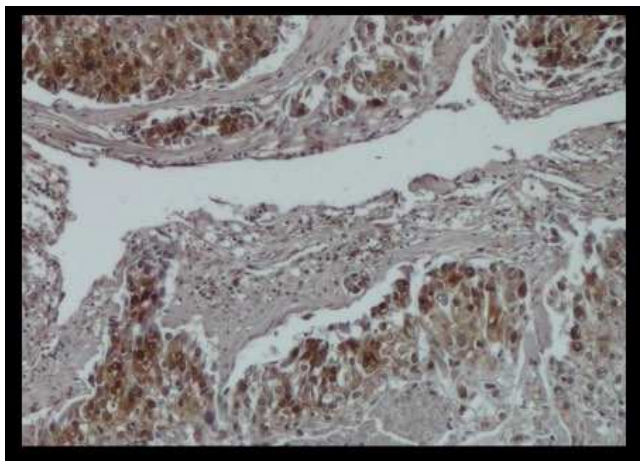
B. MYBL1



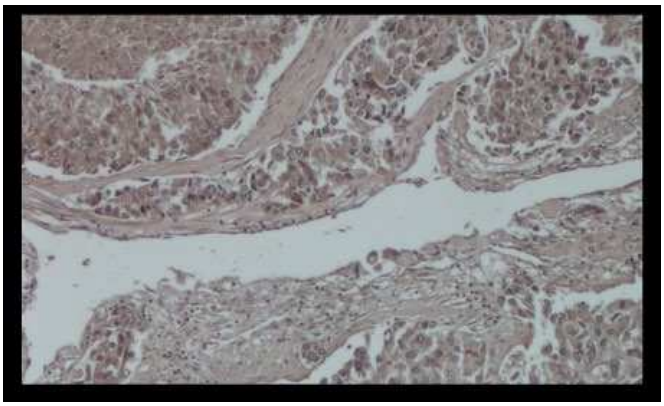
C. CD31 CONTROL

Figure 12: IHC Staining on the TMA for CD31 Blood Vessels as well as Candidate Genes MYBL1 and ADRM1

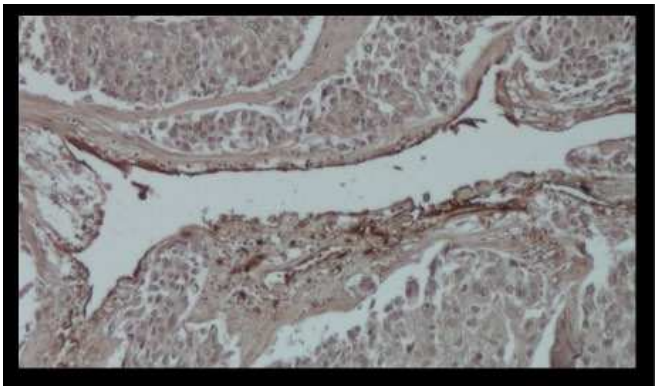
These cells are poorly differentiated.



A.) **ADRM1**



B.) **MYBL1**



C.) **CD31**

Figure 13: IHC Staining of the TMA ADRM1 and MYBL1 Protein were Detected in Different Cells
These cells are well differentiated.

Results from Aim 2:

Table 5: GDS2250 Dataset of MYBL1. MYBL2, ADRM1, and UBXN8 Levels in Patient Samples

GDS2250	BREAST CANCER TYPES	MYBL1	MYBL2	ADRM1	UBXN8
	N1	162.6954	48.93241	6949.882	424.2785
	N2	194.0993	142.2562	7492.367	245.4737
	N3	52.32268	127.6319	4129.824	272.0865
	N4	120.5229	143.4304	4614.921	333.983
	N5	159.9514	87.58327	4167.16	269.1019
	N6	70.33292	267.4814	6747.113	459.0596
	N7	1017.384	103.5019	5390.535	262.1901
	LUM1	1565.03	451.141	9286.899	1184.813
	LUM2	754.7149	2019.388	16756.31	803.3496
	LUM3	295.9558	772.8771	6811.789	1114.945
	LUM4	15561.11	690.0955	10193.14	401.2923
	LUM5	4121.984	918.8933	5938.608	616.3099
	LUM6	25.84998	1215.803	10148.49	116.7389
	LUM7	256.888	473.9112	7904.125	328.048
	LUM8	4549.943	635.2191	15603.96	269.5058
	LUM9	228.5626	228.4255	13785.88	22.89091
	LUM10	2079.681	779.0226	10673.19	861.7024
	LUM11	6670.766	1392.074	4812.827	235.7448
	LUM12	22.70715	1485.208	12760.63	342.3111
	LUM13	312.4736	2484.657	4185.536	324.3425

	LUM14	6741.246	800.791	31533.32	251.9524
	LUM15	100.5666	171.3949	6463.072	280.2466
	LUM16	218.4899	296.2489	7509.844	149.3227
	LUM17	2126.494	702.1099	4113.996	295.6363
	LUM18	4342.302	4547.214	11250.97	195.2186
	LUM19	647.449	1132.293	23701.65	261.1565
	LUM20	91.11204	1180.803	9646.689	166.017
	TNBC1	4938.318	1991.195	12658.45	97.45395
	TNBC2	3339.579	3660.006	21452.15	311.2044
	TNBC3	25210.1	4334.407	22134.66	2062.244
	TNBC4	12666.55	293.4127	12211.46	133.028
	TNBC5	2517.748	916.0766	7718.613	295.9824
	TNBC6	1193.124	2167.132	12478.22	85.60896
	TNBC7	9037.971	1498.965	8123.448	179.3519
	TNBC8	61.00464	974.9342	7741.185	41.29783
	TNBC9	71118.3	3644.338	23077.19	106.2654
	TNBC10	299.1874	556.9637	6927.609	56.83714
	TNBC11	5122.152	925.2463	15215.65	116.1137
	TNBC12	510.9696	3186.951	21940.73	200.415
	TNBC13	58764.9	2988.33	13221.77	171.6693
	TNBC14	108.136	1555.014	5955.855	20.71421
	TNBC15	1889.429	2741.142	9101.277	78.77061
	TNBC16	14284.78	2604.823	13750.77	80.79378
	TNBC17	3494.97	4301.16	19125.54	206.5453
	TNBC18	2043.501	1664.017	27769.68	843.8207

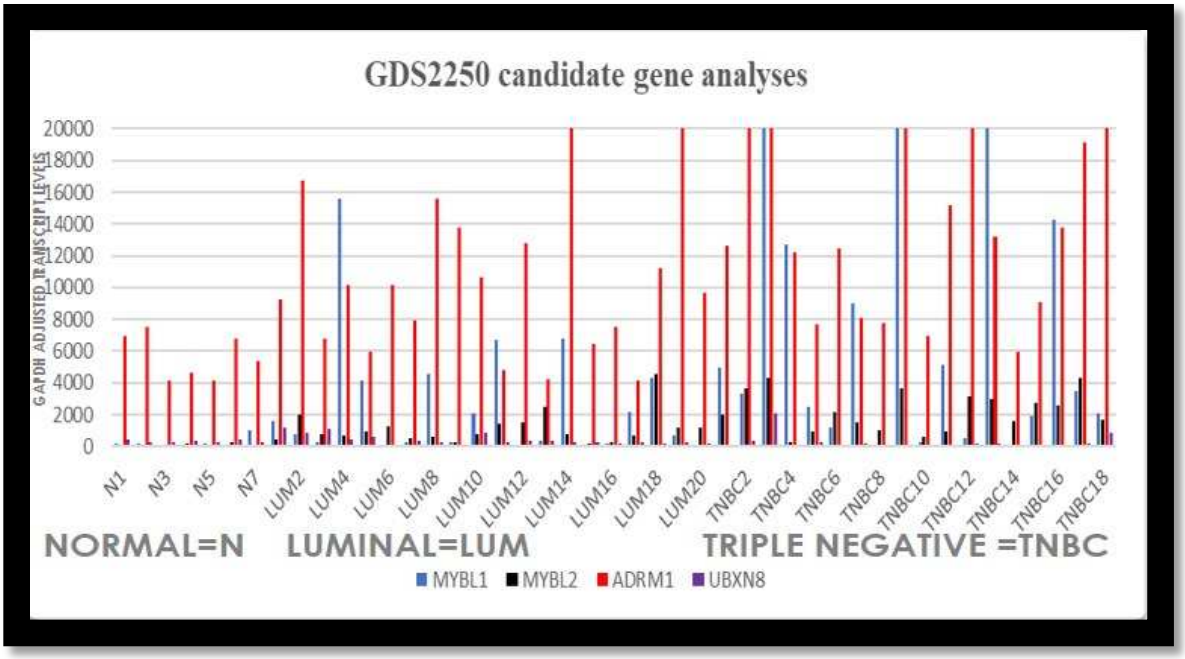


Figure 14: GDS2250 Normal, Luminal, and Basal-TNBC Patients vs Candidate Genes

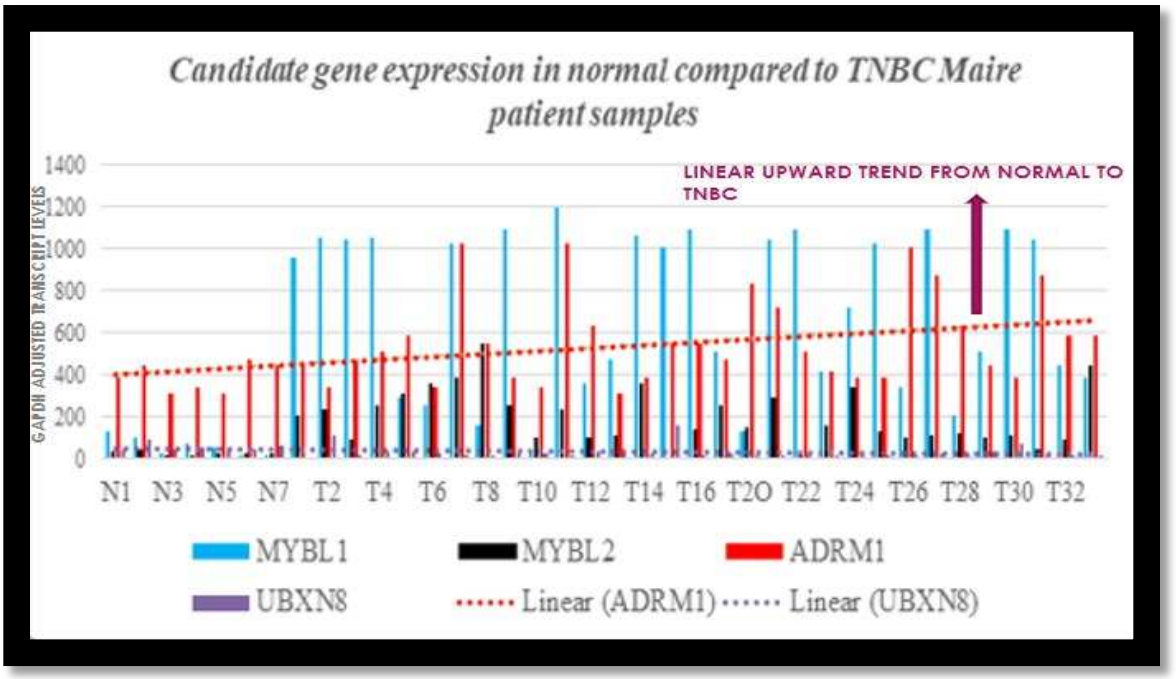


Figure 15: Candidate Gene Expression in Normal vs TNBC in MAIRE Patient Samples Bar Graph

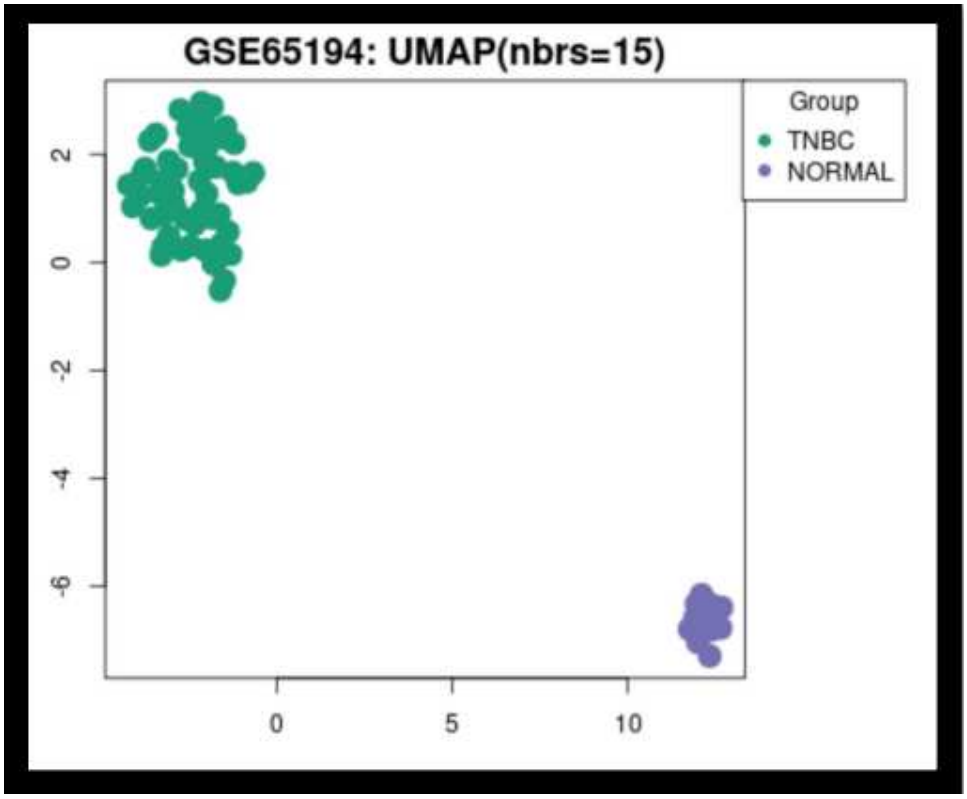


Figure 16: Candidate Gene Expression in Normal vs TNBC in MAIRE Patient Samples

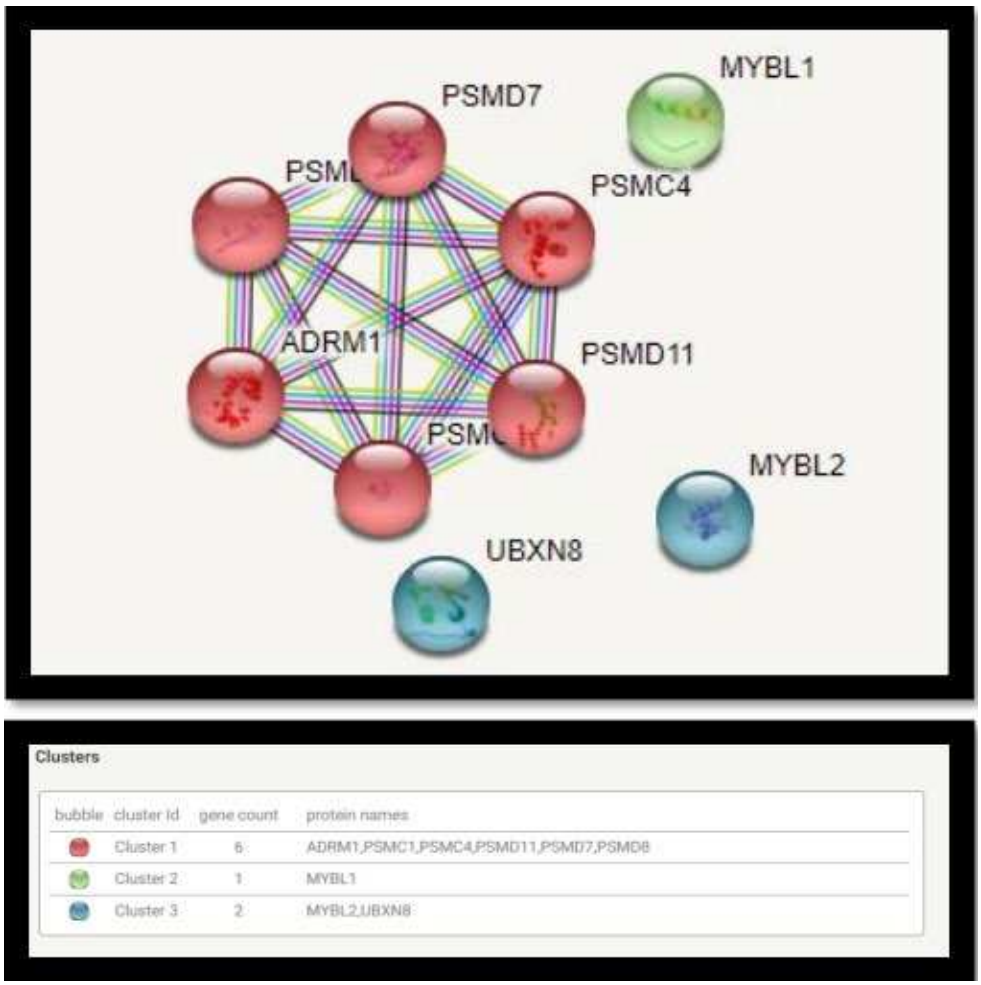


Figure 17: String Analysis of Candidate Genes

CHAPTER 5

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

The goal of this project was to compare the gene expression profiles of ADRM1 and UBXN8 genes to that observed for MYBL1. The focus of our laboratory is further characterizing MYBL1 in TNBC and identify genes that might be either directly or indirectly associating with the gene to affect the genotype of TNBC. Although we have not demonstrated MYBL1's effect on the cancer genotype we suspect MYBL1 does play a role based on its function. Our next experiments will attempt to answer this question.

So, what about ADRM1 and UBXN8? ADRM1 is worth a second look, but UBXN8 is not. The IHC results will be repeated using more suitable arrays and with more patient samples. ADRM1 is proving to be a biomarker for luminal cancers, but we will perform experiments to address the possibility that ADRM1 is indirectly affected by MYBL1 in TNBC because defining genes affected by MYBL1 is one of our main goals. ChipX suggests ADRM1 is directly regulated by NPAS1. The knockdown datasets will be re-analyzed looking for a 'link' between the differentially expressed genes on the list and NPAS1.

Two experimental procedures are critical to the success of the experiments presented in this document. The success of the experiments outlined here depend on the shRNA lentiviral MYBL1 knockdown procedure and the interpretation of the results generated using the DNA microarray platform. We are in the process of repeating the

shRNA lentiviral knockdowns and as we have repeatedly emphasized, even though the microarray is an invaluable resource, the data generated from the platform must be validated. This phase of the experiments will continue, MYBL1 and validation of its expression and search for genes affected by its expression are key goals of our laboratory.

The lab sees MYBL1 as a candidate gene for biomarker studies eventually leading to possible targeted gene therapies for TNBC patients. This current study is one of many to be conducted in our laboratory in relation to MYBL1 and its role in TNBC. The laboratory is also looking into MYBL1's role in tumor progression. This future study will be done using a mice model. The mice model study will further validate MYBL1's position in tumor progression. If MYBL1 plays a part in the tumor development, the knockdown of the gene will then halt the tumor from forming in the mice model. Of the nineteen genes selected from the preliminary knockdown study, TCF19 and KIF18B were the most promising candidates. Further analysis of the genes is underway in our laboratory because they seem to be connected to key signaling processes in TNBC. Studies will continue with these projects to achieve the overall goal of characterizing MYBL1 in TNBC.

.

REFERENCES

1. "Breast Cancer Facts & Figures." American Cancer Society, https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/breast-cancer-facts-and-figures/breast-cancer-facts-and-figures-2019-2020.pdf
2. "Breast Cancer." MD Anderson Cancer Center, www.mdanderson.org/cancer-types/breast-cancer.html/?invsrc=national.
3. "How Common Is Breast Cancer?" American Cancer Society, https://www.cancer.org/cancer/breast-cancer/about/how-common-is-breast-cancer.html
4. Agarwal, V., et al., Predicting effective microRNA target sites in mammalian mRNAs. Elife, 2015. 4.
5. Arsura, M., et al., A-myb rescues murine B-cell lymphomas from IgM-receptor-mediated apoptosis through c-myc transcriptional regulation. Blood, 2000. 96(3): p. 1013-20.
6. Bao L, Guo T, Wang J, Zhang K, Bao M. Prognostic genes of triple-negative breast cancer identified by weighted gene co-expression network analysis. Oncol Lett. 2020 Jan;19(1):127-138. doi: 10.3892/ol.2019.11079. Epub 2019 Nov 11. PMID: 31897123; PMCID: PMC6923995.

7. Barrett, T., et al., NCBI GEO: archive for functional genomics data sets--update. Nucleic Acids Res, 2013. 41(Database issue): p. D991-5.

8. Bianchini, G., et al., Triple-negative breast cancer: challenges and opportunities of a heterogeneous disease. Nat Rev Clin Oncol, 2016. 13(11): p. 674-690.

9. Boyer, L.A., et al., Essential role for the SANT domain in the functioning of multiple chromatin remodeling enzymes. Mol Cell, 2002. 10(4): p. 935-42.

10. Brea, M.S., et al., Epidermal Growth Factor Receptor Silencing Blunts the Slow Force Response to Myocardial Stretch. J Am Heart Assoc, 2016. 5(10).

11. Carey, L.A., et al., Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. Jama, 2006. 295(21): p. 2492-502.

12. Cicirò, Y. and A. Sala, MYB oncoproteins: emerging players and potential therapeutic targets in human cancer. Oncogenesis, 2021. 10(2): p. 19.

13. Costa RLB, Gradishar WJ. Triple-Negative Breast Cancer: Current Practice and Future Directions. J Oncol Pract. 2017;13(5):301-303

14. Dai, X., et al., Breast Cancer Cell Line Classification and Its Relevance with Breast Tumor Subtyping. J Cancer, 2017. 8(16): p. 3131-3141.

15. Dalma-Weiszhausz, D.D., et al., The affymetrix GeneChip platform: an overview. Methods Enzymol, 2006. 410: p. 3-28.

16. Dawood S. Triple-negative breast cancer: epidemiology and management options. Drugs. 2010 Dec 3;70(17):2247-58. doi: 10.2165/11538150-000000000-00000. PMID: 21080741.

17. Engeland, K., Cell cycle arrest through indirect transcriptional repression by p53: I have a DREAM. Cell Death Differ, 2018. 25(1): p. 114-132.

18. Farmer P, Bonnefoi H, Becette V, et al Oncogene. 2005 Jul 7;24(29):4660-71.

19. Fischer, M. and G.A. Müller, Cell cycle transcription control: DREAM/MuvB and RB-E2F complexes. Crit Rev Biochem Mol Biol, 2017. 52(6): p. 638-662.

20. Fischer, M., et al., Integration of TP53, DREAM, MMB-FOXM1 and RB-E2F target gene analyses identifies cell cycle gene regulatory networks. Nucleic Acids Res, 2016. 44(13): p. 6070-86.

21. Franchini, G., et al., Structural organization and expression of human DNA sequences related to the transforming gene of avian myeloblastosis virus. Proc Natl Acad Sci U S A, 1983. 80(24): p. 7385-9.

22. Golay J, Loffarelli L, Luppi M, Castellano M, Introna M. The human A-myb protein is a strong activator of transcription. Oncogene. 1994 Sep;9(9):2469-79. PMID: 8058310.

23. Golay, J., et al., Expression of A-myb, but not c-myb and B-myb, is restricted to Burkitt's lymphoma, sIg+ B-acute lymphoblastic leukemia, and a subset of chronic lymphocytic leukemias. Blood, 1996. 87(5): p. 1900-11.

24. Gonda, T.J., P. Leo, and R.G. Ramsay, Estrogen and MYB in breast cancer: potential for new therapies. Expert Opin Biol Ther, 2008. 8(6): p. 713-7.

25. Gorbatenko, A., et al., HER2 and p95HER2 differentially regulate miRNA expression in MCF-7 breast cancer cells and downregulate MYB proteins through miR-221/222 and miR-503. Sci Rep, 2019. 9(1): p. 3352.

26. Hanahan, D. and R.A. Weinberg, The hallmarks of cancer. Cell, 2000. 100(1): p. 57-70.

27. Howlader N, Noone AM, Krapcho M, et al., eds. SEER Cancer Statistics Review, 1975-2016. Bethesda, MD: National Cancer Institute; 2019. Available from seer.cancer.gov/csr/1975_2016/, based on November 2018 SEER data submission, posted to the SEER web site April 2019.

28. Hurvitz S, Mead M. Triple-negative breast cancer: advancements in characterization and treatment approach. Curr Opin Obstet Gynecol. 2016 Feb;28(1):59-69. doi: 10.1097/GCO.0000000000000239. PMID: 26694831.

29. Iness, A.N., et al., The cell cycle regulatory DREAM complex is disrupted by high expression of oncogenic B-Myb. Oncogene, 2019. 38(7): p. 1080-1092.

30. Ji, Z., et al., KIF18B as a regulator in microtubule movement accelerates tumor progression and triggers poor outcome in lung adenocarcinoma. Tissue Cell, 2019. 61: p. 44-50.

31. Kent, L.N. and G. Leone, The broken cycle: E2F dysfunction in cancer. Nat Rev Cancer, 2019. 19(6): p. 326-338.

32. Kim, J., et al., MYBL1 rearrangements and MYB amplification in breast adenoid cystic carcinomas lacking the MYB-NFIB fusion gene. J Pathol, 2018. 244(2): p. 143-150.

33. Ko, E.R., et al., A conserved acidic patch in the Myb domain is required for activation of an endogenous target gene and for chromatin binding. Mol Cancer, 2008. 7: p. 77.

34. Koressaar, T. and M. Remm, Enhancements and modifications of primer design program Primer3. Bioinformatics, 2007. 23(10): p. 1289-91.

35. Kosok, M., Alli-Shaik, A., Bay, B. H., & Gunaratne, J. (2020). Comprehensive

Proteomic Characterization Reveals Subclass-Specific Molecular Aberrations within Triple-negative Breast Cancer. iScience, 23(2), 100868. https://doi.org/10.1016/j.isci.2020.100868

36. Kyrpychova, L., et al., Small Subset of Adenoid Cystic Carcinoma of the Skin Is Associated with Alterations of the MYBL1 Gene Similar to Their Extracutaneous Counterparts. Am J Dermatopathol, 2018. 40(10): p. 721-726.
37. Lehmann, B.D., et al., Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. J Clin Invest, 2011. 121(7): p. 2750-67.
38. Li, X.Z., et al., An ancient transcription factor initiates the burst of piRNA production during early meiosis in mouse testes. Mol Cell, 2013. 50(1): p. 67-81.
39. Litovchick, L., et al., Evolutionarily conserved multisubunit RBL2/p130 and E2F4 protein complex represses human cell cycle-dependent genes in quiescence. Mol Cell, 2007. 26(4): p. 539-51.
40. Liu, L.Y., et al., A supervised network analysis on gene expression profiles of breast tumors predicts a 41-gene prognostic signature of the transcription factor MYB across molecular subtypes. Comput Math Methods Med, 2014. 2014: p. 813067.
41. Liu, X., et al., Reassessing the Potential of Myb-targeted Anti-cancer Therapy. J Cancer, 2018. 9(7): p. 1259-1266.
42. Liu, Z.L., et al., Expressions and prognostic values of the E2F transcription factors in human breast carcinoma. Cancer Manag Res, 2018. 10: p. 3521-3532.

43. Maire, V., et al., TTK/hMPS1 is an attractive therapeutic target for triple-negative breast cancer. PLoS One, 2013. 8(5): p. e63712.

44. Marchler-Bauer, A., et al., CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. Nucleic Acids Res, 2017. 45(D1): p. D200-d203.

45. Marhamati, D.J., et al., A-myb is expressed in bovine vascular smooth muscle cells during the late G1-to-S phase transition and cooperates with c-myc to mediate progression to S phase. Mol Cell Biol, 1997. 17(5): p. 2448-57.

46. Mi, H., et al., PANTHER version 14: more genomes, a new PANTHER GO-slim and improvements in enrichment analysis tools. Nucleic Acids Res, 2019. 47(D1): p. D419-d426.

47. Muthukaruppan, A., et al., Multimodal Assessment of Estrogen Receptor mRNA Profiles to Quantify Estrogen Pathway Activity in Breast Tumors. Clin Breast Cancer, 2017. 17(2): p. 139-153.

48. Nedeljković, M., & Damjanović, A. (2019). Mechanisms of Chemotherapy Resistance in Triple-Negative Breast Cancer-How We Can Rise to the Challenge. Cells, 8(9), 957. https://doi.org/10.3390/cells8090957

49. Pandya, S. and R.G. Moore, Breast development and anatomy. Clin Obstet Gynecol, 2011. 54(1): p. 91-5.3.

50. Parise CA, Caggiano V. Risk of mortality of node negative, ER/PR/HER2 breast cancer subtypes in T1, T2, and T3 tumors. Breast Cancer Res Treat. 2017 Oct;165(3):743-750. doi: 10.1007/s10549-017-4383-5. Epub 2017 Jul 8. PMID: 28689363.

51. Player A, Abraham N, Abdulrahman N, Nsende E, Cunningham S and Rogers S. MYBL1 Knockdown in a Triple Negative Breast Cancer Line: Evidence of Down-Regulation of MYBL2, TCF19 and KIF18b Expression. Austin J Cancer Clin Res. 2021; 8(1): 1090.

52. Player, A., et al., Identification of candidate genes associated with triple negative breast cancer. Genes Cancer, 2017. 8(7-8): p. 659-672.

53. Player, A., Oguamanam, T., Okanmelu, J. et al. Preliminary characterization of IL32 in basal-like/triple negative compared to other types of breast cell lines and tissues. BMC Res Notes 7, 501 (2014). https://doi.org/10.1186/1756-0500-7-501.

54. Plevritis SK, Munoz D, Kurian AW, et al. Association of Screening and Treatment with Breast Cancer Mortality by Molecular Subtype in US Women, 2000-2012. JAMA. 2018;319(2):154-164.

55. Prat A, Adamo B, Cheang MC, Anders CK, Carey LA, Perou CM. Molecular characterization of basal-like and non-basal-like triple negative breast cancer. Oncologist. 2013;18(2):123-133.

56. Rashid, N.N., H.A. Rothan, and M.S. Yusoff, The association of mammalian DREAM complex and HPV16 E7 proteins. Am J Cancer Res, 2015. 5(12): p. 3525-33.

57. Richardson AL, Wang ZC, De Nicolo A, et al Cancer Cell. 2006 Feb;9(2):121-32.

58. Rushton, J.J. and S.A. Ness, The conserved DNA binding domain mediates similar regulatory interactions for A-Myb, B-Myb, and c-Myb transcription factors. Blood Cells Mol Dis, 2001. 27(2): p. 459-63.

59. Rushton, J.J., et al., Distinct changes in gene expression induced by A-Myb, B-Myb and Ness, S.A., Myb protein specificity: evidence of a context-specific transcription factor code. Blood Cells Mol Dis, 2003. 31(2): p. 192-200.

60. Szklarczyk, D., et al., The STRINGTM database in 2017: quality-controlled protein-protein association networks, made broadly accessible. Nucleic Acids Res, 2017. 45(D1): p. D362-d368.

61. Thorner, A.R., et al., Potential tumor suppressor role for the c-Myb oncogene in luminal breast cancer. PLoS One, 2010. 5(10): p. e13073.

62. Togashi, Y., et al., MYB and MYBL1 in adenoid cystic carcinoma: diversity in the mode of genomic rearrangement and transcripts. Mod Pathol, 2018. 31(6): p. 934-946.

63. Wang, T. and T.S. Furey, Analysis of complex disease association and linkage studies using the University of California Santa Cruz Genome Browser. Circ Cardiovasc Genet, 2009. 2(2): p. 199-204.

64. Weigelt, B., F.C. Geyer, and J.S. Reis-Filho, Histological types of breast cancer: how special are they? Mol Oncol, 2010. 4(3): p. 192-208.

65. Wolter, P., et al., Central spindle proteins and mitotic kinesins are direct transcriptional targets of MuvB, B-MYB and FOXM1 in breast cancer cell lines and are potential targets for therapy. Oncotarget, 2017. 8(7): p. 11160-11172.

66. Wu, Y., et al., KIF18B promotes tumor progression through activating the Wnt/β-catenin pathway in cervical cancer. Onco Targets Ther, 2018. 11: p. 1707-1720.

67. Xu, L.H., et al., MYB promotes the growth and metastasis of salivary adenoid cystic carcinoma. Int J Oncol, 2019. 54(5): p. 1579-1590.

68. Zeng, C.X., et al., TCF19 enhances cell proliferation in hepatocellular carcinoma by activating the ATK/FOXO1 signaling pathway. Neoplasma, 2019. 66(1): p. 46-53.

69. Zhou, Y. and S.A. Ness, Myb proteins: angels and demons in normal and transformed cells. Front Biosci (Landmark Ed), 2011. 16: p. 1109-31.

70. Zhou, Z.H., et al., TCF19 contributes to cell proliferation of non-small cell lung cancer by inhibiting FOXO1. Cell Biol Int, 2019.

71. Ziebold, U. and K.H. Klempnauer, Linking Myb to the cell cycle: cyclin-dependent phosphorylation and regulation of A-Myb activity. Oncogene, 1997. 15(9): p. 1011-9.