12-2023

Analyses Of Mybl1 In Triple Negative Breast Cells And Validation Of Genes Colocalized To Chromosome 8q Loci

La Tajia D. Thirston
Texas Southern University

ANALYSES OF MYBL1 IN TRIPLE NEGATIVE BREAST CELLS AND VALIDATION OF GENES COLOCALIZED TO CHROMOSOME 8Q LOCI

THESIS

Presented in Partial Fulfillment of the Requirements for

the Degree Master of Science in the Graduate School

of Texas Southern University

By

La Tajia Diamonéé Thirston, B.S.

Texas Southern University

2023

Approved by

Dr. Audrey Player
Chairperson, Thesis Committee

Dr. Mahesh Vanjani
Dean, The Graduate School

Approved by

Dr. Audrey Player	10/26/2023
Chairperson, Thesis Committee	Date
Dr. Hector Miranda	10/26/2023
Committee Member	Date
Dr. Ayodotun Sodipe	10/26/2023
Committee Member	Date
Dr. Roderick Holmes	10/26/2023
Committee Member	Date

ANALYSES OF MYBL1 IN TRIPLE NEGATIVE BREAST CELLS AND VALIDATION OF GENES COLOCALIZED TO CHROMOSOME 8Q LOCI

By

La Tajia D. Thirston, M.S.

Texas Southern University, 2023

Professor Audrey Player, Ph.D., Advisor

The research goal of our laboratory is to further characterize MYBL1 in triple negative breast cancers. In order to identify genes either directly or indirectly associated with MYBL1 in these cancers, we knocked down MYBL1 gene (in a triple negative breast cancer cell line) and performed microarray analyses of the cell line transcriptome. The MYBL1 gene is located at chromosome location 8q13.1. Bioinformatics analyses of genes either directly or indirectly effected by MYBL1 knockdown show a substantial number of genes located on chromosome 8q near and distance from the MYBL1 gene locus. The current project involves PCR transcript and protein validation of MYBL1, and other genes located on at the chromosome 8q loci. The current study should lead to further characterization of MYBL1 in triple negative breast cancers so that we can better understand the cancer. Fifteen percent of breast cancer are characterized as the triple negative breast cancer subtype. The cancers are recognized as negative for estrogen, progesterone and ErRB2 receptors, genes that function as transcription and growth factor

functions, respectively. Previous data show the MYBL1 transcription factor is over-expressed in some triple negative breast cancers-suggesting dysregulation of the gene in the cancers. MYBL1 gene is a strong transcriptional regulator associated with regulation of cell cycle proliferation which is a key event in cancer progression. To determine genes either directly or indirectly associating with MYBL1 in these cancers we knocked the gene down, performed microarray analyses and identified a list of genes effected by MYBL1 knockdown (some up-regulated and others down regulated). MYBL1 gene is localized to chromosome region 8q.13.1 and bioinformatic analyses show a substantial number of genes either at the same loci or colocalized to the 8q loci. For this study, based on the knockdown, microarray and subsequent bioinformatic analyses, 14 genes were initially identified as possible candidates for validation using bioinformatic based approaches, and polymerase chain and western blotting experimental analyses. Data presented here are a summary of our results which include analysis of the 14 genes and selection of a final subset of 4 genes as final candidates for this study.

TABLE OF CONTENTS

LIST OF TABLES

Tables Page

LIST OF FIGURES

Figures Page

LIST OF ABBREVIATIONS

ATCC	American Type Culture Collection
AMV	Avian Myeoblastosis Virus
ATR	Attenuated Total Reflectance
ATM	DNA repair
BL1	Basal-like types
BL2	Basal-like types 2
cDNA	Complementary DNA
DBD	DNA binding domain
dT	Oligo dT
DCIS	Ductal Carcinoma in Situ
DREAM	Dimerization partner (DP), retinoblastoma (RB)-like, E2F and MuvB)
G1	GAP 1 phase
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GGH	Gamma Glutamyl-hydrolase
HER2	Human epidermal growth factor receptor
IBC	Inflammatory Breast Cancer
IDC	Invasive Duct al Carcinoma
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 B	Hormone-receptor positive (estrogen-receptor positive, progesterone-receptor positive and HER2 positive)

LAR	Luminal androgen receptor
LIN9	LIN9-DREAM MuvB core complex component
LIN37	LIN37-DREAM MuvB core complex component
LIN52	LIN52-DREAM MuvB core complex component
LIN54	LIN54-DREAM MuvB core complex component
mRNA	Messenger RNA
MAF1	Repressor of RNA polymerase III transcription MAF1
M	Mesenchymal group
MSL	Mesenchymal stem-like subtype
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	V-Myb Avian Myeloblast Viral Oncogene Homolog 1
MYBL2	Myb-related protein B
MYC	Proto-oncogene, transcription factor, cell cycle, apoptosis, differentiation
NCBI	National Center for Biotechnology Information
NFIB	Nuclear Factor I B
NE	Nebraska
PCR	Polymerase Chain Reaction
PR	Progesterone receptor
RBL1	Retinoblastoma-Like 1 Protein
RBL2	Retinoblastoma-Like 2 Protein
RBBP4	Retinoblastoma Binding 4 Protein
RMA	Robust Multi Array
RNA	Ribonucleic acid
RPM	Revolutions per minute

SANT	DNA-chromatin recognition sequence
SC	Santa Cruz
shRNA	short hairpin RNA
S phase	Synthesis Phase
STRINGTM	Search Tool for Retrieval of Interacting Genes/Proteins
TAQ	Polymerase thermostable DNA polymerase I
TBE	Tris/Borate/EDTA
TCF19	Transcription factor 19
TFIIIB	Transcription Factor protein
TNBC	Triple Negative Breast Cancer
TFDP2	Transcription Factor Dp protein 2
TFDP1	Transcription Factor Dp protein 1
TP53	Tumor protein p53

VITA

2020 …….……	B.S. Biology; Minor: Chemistry Prairie View A&M University Prairie View, Texas
2022-2023 …….....	Graduate Teaching Assistant Department of Biology Texas Southern University Houston, Texas
Major Field…….	Biology

ACKNOWLEDGEMENTS

Throughout this journey, since I was a young child this is something that has been heavy on my heart to ensure I completed all of my goals. Along the way, there have been so many people near and far that have been supportive in this journey. First and foremost, none of this would be possible without my Lord and Savior, without him I am nothing. Next, I'll like to give thanks to my loving family, partner, and friends,

My sincere appreciation goes to the most amazing and phenomenal advisor, Dr. Audrey Player and my wonderful committee members: Dr. Ayodotun Sodipe, Dr. Roderick Holmes, and Dr. Hector Miranda

To my supportive lab partners, Fawaz Ibikunle, Deshai Philio, and other lab mates thank you. To all who has played a part in getting me where I am today from grade school on up and lastly Texas Southern's Graduate School and the entire Biology Department…. Thank you.

CHAPTER 1

INTRODUCTION

Background Information About Breast Cancer and Summary of Breast Cancer Types:

According to The American Cancer Society, breast cancer is the most common cancer in women, besides skin cancers (5). About 281,550 new cases of invasive breast cancer will be diagnosed in women and about 43,600 women will die each year from the cancer. Cancers in general are characterized as a disease in which cells grow uncontrollably, ignore signals related to programmed cell death and differentiate usually to a mesenchymal-like genotype and phenotype (5). Like most cancers, breast cancer can be classified based on pathological diagnoses; but some breast cancers can be classified based on the molecular receptor status of the cancer cells. The molecular receptor status of these tumors can drastically affect the type of cancer, prognosis, invasive potential and subsequent treatment (5).

Cancers related to the breast can affect both men and women, but most commonly the disease affects women. Some breast cancers originate from luminal, myoepithelial, and basal-like progenitor cells that line the lumen (Figure 1). Data show the progenitor cells, and the microenvironment cells drive cancer progression and contribute to the heterogeneity of breast cancers. Breast cancers, like other cancers, can be characterized as invasive or noninvasive. Noninvasive cancers do not travel beyond the site of origin.

Invasive cancers travel from their initial site of origin to other organs where they are considered life-threatening. Cancers that leave and return in the same or different location in the body are reoccurring cancers. Benign tissues, which are non-cancerous and unharmful, are localized to their original site of origin.

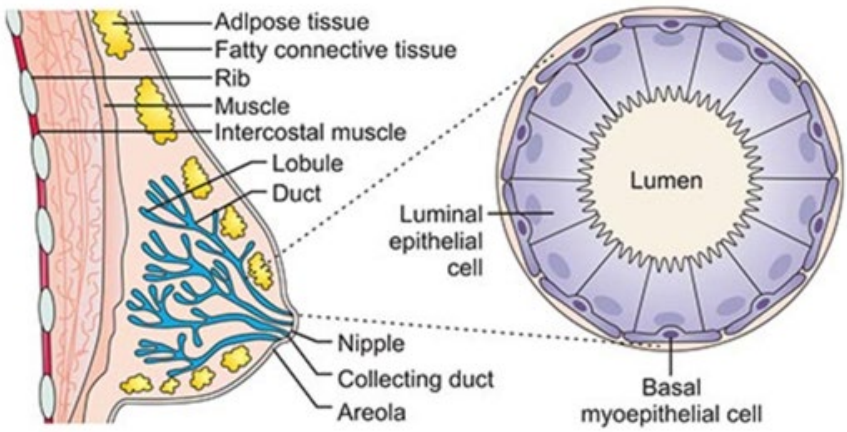

Figure 1: Location of Breast Cancer Progenitor Cell Types. Location of Luminal Cells and Basal Cell Progenitor Cell Types

Before the advent of the DNA microarray studies, breast and all cancers were almost exclusively characterized based on pathological diagnoses. All cancers, including breast cancers are extremely heterogeneous. Microarray analyses have been instrumental in better characterizing the molecular signatures associated with cancer and other diseases. Currently, breast (and other cancers) are characterized based on both pathological and molecular characteristics (60). DNA microarray analyses followed by polymerase chain reaction are the most common methods used to define the molecular signatures that define breast cancers. As proof of how beneficial microarray are, nearly all of the current targeted therapies used to treat cancers are identified by microarray analyses. A list of the breast cancers based on pathology and molecular signatures are in the section below.

PATHOLOGICALLY DIAGNOSED BREAST CANCER TYPES:

Male Breast Cancer: Male breast cancers are rare, occurring in less than 1% of men. They are defined based on the same pathological criteria as women cancers. They can be deadly and invasive.

Cribriform Carcinoma of the Breast: Cribriform cancers are rare invasive cancers, which occur at a rate of 0.3-3% of breast cancers. The cancers are low grade (i.e., look normal), but appear to have 'holes or display cribriform-like configurations.

Lobular Carcinoma in Situ (LCIS): LCIS is also known as a benign lobular neoplasia. LCIS presents an abnormal cell growth in the milk glands (lobules), occurring at a rate of less than 2% of cases. Similar to Ductal Carcinoma in Situ, it can signal a higher risk of future invasive cancer.

Inflammatory Breast Cancer (IBC): IBC are rare invasive and often deadly breast cancers that occur at a rate of 1-5% of breast cancer. They are difficult to detect because unlike most other breast cancers that present as lumps, IBC cancers present as "sheets" of differentiated cells; this also makes the cancers difficult to treat.

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 occurring in ~2% of breast cancers. The cancers originate in the milk ducts. The cancers present as 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 of the breast involves cancer cells associated with the nipple and breast ducts. Paget's disease occurs in ~ 1-4% of breast cancers and is often mistaken for dermatitis because skin around the nipple is flaky and the nipples are inverted and hardened 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 occurring in ~1% of patients.

Phyllodes Tumors of the Breast: Phyllodes cancers originate in stromal breast cells. The cancer occurs in less than 1% of breast cancer patients and present as leaf-like arrangements.

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, then spreads to surrounding ducts. Microscopically the cells are present as tube-like structures.

Ductal Carcinoma in Situ (DCIS): DCIS is a non-invasive carcinoma which originates in the milk ducts of the breast. DCIS cancers occur at ~15% and they do not metastasize. Patients with DCIS cancers have a significant risk of developing invasive cancers later.

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.

Invasive Ductal Carcinoma (IDC): IDC is the most common kind of breast cancer, occurring at approximately 60-70% of breast cancers. The IDC cancers are invasive cancers that occur in the milk ducts of patients.

BREAST CANCER TYPES IDENTIFIED BASED ON GENETIC SIGNATURES:

The molecular characterization of breast (and other) cancers was primarily determined using microarray analyses and validated using polymerase chain reaction (54). The definition and incidence of these cancers are outlined in Figure 2, and more detailed characterizations are listed in https://www.ncbi.nlm.nih.gov/books/NBK583808/. Many breast cancers are referred to as luminal or triple negative / basal-like subtypes based on their progenitor cell types and genomics. Luminal subtype originates from luminal cells of the breast and triple negative/basal-like breast cancers originate from the cells lining the basement portion of the lumen (Figure 1).

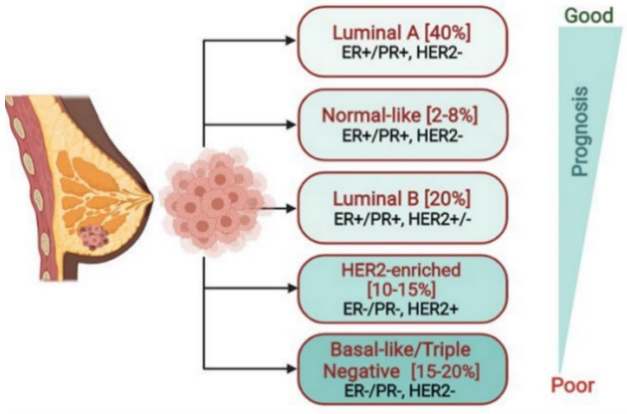

Figure 2: Breast Cancer Types Based on Molecular Signatures

Molecular Subtypes of Breast Cancer are Luminal A/B, triple negative/basal-like, HER2-enriched, and Normal-like:

Based on molecular gene expression signatures breast cancers are identified as luminal A/B, triple negative / basal-like, HER2-enriched and Normal-like. *Luminal A* cancers are classified as estrogen receptor positive, progesterone receptor positive, and HER2 negative Yersal et al (84). *Luminal B* cancers are estrogen receptor and progesterone receptor positive, and HER2 positive or negative (84). The prognoses of Luminal B cancers versus Luminal A cancers are marginal (84). *Triple negative/basal-like* (TNBC) cancers are negative for all three markers, estrogen and progesterone receptor, and HER2 negative (84). The TNBC and basal-like are similar based on gene expression levels, but data show the cancers are at least 25% different based on gene expression markers. Triple negative breast is the subtype examined in the current study and are characterized in greater detail in a section below. *HER2neu-enriched* breast cancers are estrogen receptor negative, progesterone receptor negative, and HER2 positive. *Normal-like* cancers are low grade cancers, with low nuclear Ki67 levels, and they share a gene expression pattern similar to Luminal A cancers in that they are estrogen receptor positive, progesterone receptor positive, and HER2 negative.

Triple Negative Breast Cancer

Microarray transcriptome analyses suggest the TNBC progenitor cells are localized to the basement region of the breast milk ducts. TNBCs account for 15-20% of the overall breast cancers (11). Women of African descent have a higher risk of occurrence and 40% higher mortality rate than Caucasian women with the same diagnosis (Siegel and Brooks 2022). Patients that die are either diagnosed with later stage or they have a cancer that has

limited therapy options available (6,18,20). Without racial disparity, cells associated with the TNBC grow rapidly, are aggressive, have a high recurrence rate and there are limited treatment options available unlike receptor positive breast cancer which are treated with hormone or targeted gene therapies (2,16,20). Since TNBC lacks the three common receptors estrogen receptors (ER), progesterone receptors (PGR), and human epidermal growth factor receptors 2 (HERS) patients have limited treatment options. Chemotherapy and radiation therapies are the treatments relied upon by patients with TNBC. Our goal is to further characterize TNBC.

Often TNBC are referred to as a single subtype, however using transcriptome microarray analyses Lehmann et al (47) show TNBC can be further divided into at least 6 subcategories based on gene expression profiles and mutational studies. Hence TNBC are a very heterogeneous subtype. Lehmann et al (44) studies are considered a seminal study towards characterizing TNBC. Results of these studies have led to identification of TNBC biomarkers and markers useful for predicting a patient response to therapies. Lehmann et al (47) 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 is defined by expression of estrogen and androgen receptors and hormonal regulatory genes. The basal-like 1 (BL1) group is enriched in genes associated with cell cycle and cell division signaling pathways and ribonucleic acid (RNA) polymerase pathway. The basal-like 2 (BL2) originates in the myoepithelium and includes genes associated with glycolysis and gluconeogenesis, and growth factor signaling

processes. The Immunomodulatory (IM) cancers are similar to medullary breast cancers and are enriched with genes related to immune signaling pathways, cytokine signaling pathways, and antigen identification, natural killer cell pathways. The M and MSL sub-category include genes related to proliferation, cell motility, mesenchymal-like differentiation, and extracellular matrix proteins.

CHAPTER 2

LITERATURE REVIEW

Background information concerning the MYB family of genes and comparative analyses of their sequence and functions:

MYBL1 gene belongs to a family containing 3 known genes, c-MYB, MYBL1 and MYBL2 genes. The genes are strong transcription factors involved in cell survival, proliferation, and differentiation. The MYB family of genes were identified based on their sequence similarity to the v-MYB oncogene, a transforming gene of the avian retrovirus which causes lymphomas and myelomas in birds (45). cMYB was the first of the family to be identified in humans, followed later by MYBL1 and MYBL2, each based on cDNA sequence similarities.

The MYB family of genes shares approximately 90% similarity at their 5'prime DNA binding domain (DBD) regions. The DBD is a highly conserved helix-turn-helix (HTH) domain DNA binding sequence at the N-terminus of the gene. The DBD includes the SANT domain (Figure 3). The SANT domain is named because it binds SW13, ADA2, N-CoR and TFIIIB genes which binds proteins with MYB-like G-C rich motifs. The SANT domain is a DNA-chromatin recognition sequence involved in transcriptional regulation. Similarity in the DBDs of the MYB genes suggest the genes interact with and regulate some of the same genes. In addition to the DBD, the genes also contain a transactivation and C-terminal region. The transactivation and C-terminal regions are less well conserved

(than the DBD) in MYB family proteins and are involved in negative regulatory functions. The domains are involved in protein interactions associated with other transcription factors, co-activators, and epigenetic mechanisms.

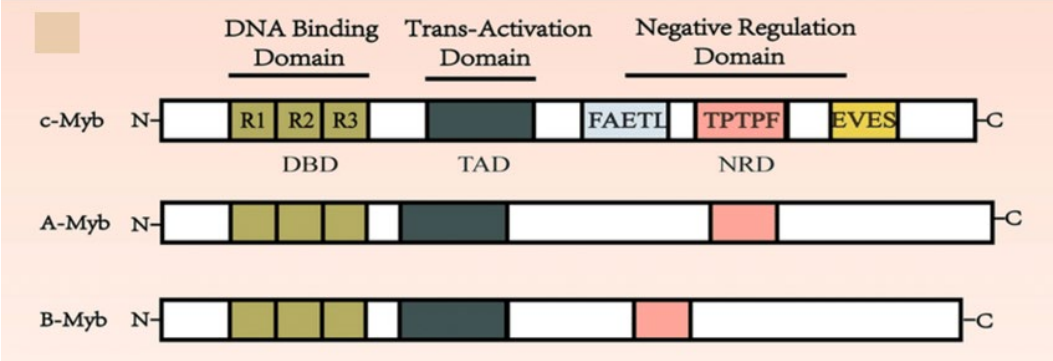

Figure 3: Comparison between cMYB, A-MYB (MYBL1) and B-MYBL2 Protein Sequence Regions- From ". Intricate Crosstalk between MYB and Noncoding RNAs in Cancer'

The MYB family of genes are strong transcriptional regulators. Lei et al (Rushton, Ness, davis) used recombinant adenoviruses followed by microarray analyses to examine comparative gene expression analyses of cMYB, MYBL1 versus MYBL2 genes in an effort to determine the expression of the genes in different cell types, and genes transcriptionally regulated by the MYBs. Their data showed (a) collectively, the MYB genes showed a different pattern of expression depending upon the type of cell (b) when the same cell type was examined, each of the MYB genes regulated a distinct set of genes and (c) the MYB genes regulated some of the same genes. Results from the Ness et al study supports the ability of the MYB family of genes to regulate the same genes, and different genes depending upon the cell type. The transactivation and C-terminal regions likely contribute to differential expression patterns observed in the microarray studies performed by Ness et al.

Of the three MYB genes, there has been a substantial number of studies performed in an attempt to better characterize the cMYB gene with comparatively fewer studies of MYBL1 and MYBL2 genes. cMYB gene is co-expressed and is regulated by estrogen receptor (64). cMYB is of particular interest because estrogen is linked to an increased risk of breast cancer in postmenopausal women. There is also an interest in cMYB because studies show that truncation of the cMYB C-terminal domain (which differ compared to MYBL1 and MYBL2) results in 'activation of cMYB' and subsequent tumorigenesis (64). Because of these and other supporting data the cMYB gene was considered a promising therapeutic target for estrogen receptor positive breast cancer patients. A substantial amount of effort was spent designing oligonucleotide and small molecule compounds to target, bind and inhibit expression of the cMYB gene product. Considerable experimental, modest preclinical and pretrial data continue to support the possible use for cMYB as a therapeutic target.

Review of processes related to gene expression of MYBL1- Involvement in cell cycle:

MYBL1 is a protooncogene involved in cell cycle regulation and proliferation, differentiation and apoptosis, mechanisms all of which are hallmarks of cancer progression. One of the earliest studies demonstrating over expression of MYBL1 gene was in studies of male Xenopus spermatogonia cells of the testis. MYBL1 was virtually absent in certain stages of embryogenesis but over-expressed in proliferating, early-stage progenitors of spermatozoa (Sleeman 1993). Based on its pattern of expression and involvement in meiosis, the investigators suggest MYBL1 is a master regulator of meiotic genes of the spermatocytes (Takahashi). When subscribing a particular function to a gene, we should remember that the same gene will likely have a somewhat different function in different

cells, attributed to cell specific expression and microenvironment; Ruston et al demonstrated this when comparing MYB family genes in different cell lines.

Ziebold and Klempnauer (2001) suggest MYBL1 acts as cell cycle sensor in human cells, with optimal gene expression levels during late G1 and early S-phases of the cell cycle. The cell cycle is a complex signaling process involving a substantial array of genes dynamically regulated. One such regulatory process involves the DREAM complex of genes, which include MYBL1and MYBL2. The DREAM complex involves interactions at various stages of the cell cycle between E2F4, E2F5, LIN9, LIN37, LIN52, LIN54, MYBL1, MYBL2, RBL1, RBL2, RBBP4, TFDP1 and TFDP2 proteins (86). Progression from S phase is dependent on the formation of subcomplexes between LIN9, LIN37, LIN52, LIN54 and MYBL2. Of the genes listed above, E2F4, LIN37 and TFDP2, MYBL2 and MYBL1 (of course), were identified as differential expressions in our gene knockdown studies supporting interactions of MYBL1 AND MYBL2 with cell cycle regulatory gene and functions.

Relationship of MYBL1 to tumorigenesis:

MYBL1 is a protein coding gene that plays a key role in regulation of proliferation, differentiation, and apoptosis programed cell death mechanisms. Dysregulation of MYBL1 has been found in low grade glioma (5), suggested as an immunotherapy biomarker in clear cell renal cell carcinoma (5), transcriptional activation of tumor angiogenesis of ANGPT2 in hepatocellular carcinoma (5), tracheobronchial adenoid cystic carcinoma (5) and genomic rearrangement in breast adenoid cystic carcinoma (5). Except for the breast adenoid cystic carcinomas, the mechanisms leading to the dysregulation in many of the other cancers and the precise cause of MYBL1 dysregulation is unknown. The breast

adenoid cystic carcinoma are rare TNBCs. The tumors have genomic rearrangements in MYBL1 in the form of fusions with ACTN1 and NFIB genes leading to MYBL1 overexpression. But these are rare events. A similar type of MYBL1/NFIB fusion gene product was identified in salivary gland carcinomas (fujii, murase, Beppu). The precise role of the gene in many of the tumorigenic processes has not been determined. Whether the gene is involved in (a) early stage signaling, (b) direct cause of or (c) late-stage progression has yet to be determined. As mentioned, data show that truncation of the 3' region of the cMYB family member is tumorigenic. However, the protein domains responsible for the tumorigenic activity in cMYB are not present in MYBL1 or MYBL2 gene, suggesting a different mode of regulation. Little is known about MYBL1 compared to cMYB gene and even less about MYBL2 gene.

Background describing our identification of MYBL1 expression in TNBC:

We initially became interested in MYBL1 following microarray and bioinformatic meta-analyses of TNBC cell datasets available in Gene Expression Omnibus. The datasets contained high through-put transcriptome cell line gene expression values. The datasets were first partitioned and analyzed using unsupervised analysis methods, then partitioned based on supervised analyses for enrichment of immune-type biomarkers. Six genes were identified, two of which were previously known as associated with TNBC and 4 novel-type genes which included MYBL1. As mentioned above, MYBL1 is considered as an immunotherapy biomarker in clear cell renal cell carcinoma. We identified the gene based on enrichment studies where the gene showed differential gene expression amongst a panel including cytokines. Further studies show MYBL1 overexpression in a subpopulation of TNBC, not all. The fact that MYBL1 is also a transcription factor makes the gene an

interesting candidate to study. Because MYBL1 transcriptionally regulates other genes, the current, primary goal of our laboratory is to determine genes that are directly or indirectly regulated by MYBL1, as a start. Since our initial observation of differential expression of MYBL1 in TNBC cell lines we have performed bioinformatic analyses of thousands of breast cancer patient samples as validation of MYBL1 overexpression in breast cancers. Curated patient datasets used for our studies are available via Cbioportal.org (15).

Experimental approach to study MYBL1 gene in TNBC

The immediate goal of our laboratory is to characterize MYBL1 gene expression in TNBC. There are many questions that we can ask at this point, but we have decided to start with a rather simple, yet informative question related to the transcription factor. MYBL1 is a strong transcriptional regulator and as such it can regulate transcription of a substantial number of genes. The first question that we ask is '*what genes are MYBL1 either directly (i.e., via binding) or indirectly regulating in TNBC"*? A substantial number of genes are affected by the MYBL1 transcription factor, as a result, our hypothesis is '*collectively these genes surely contribute to the genotype and phenotype of the cancers'*.

In order to determine genes either directly or indirectly affected by MYBL1 gene expression in TNBC, we chose to (a) knockdown MYBL1 expression in a TNBC cell line using the lentiviral procedure which compared a lentiviral construct without the MYBL1 silencing target (control) versus a construct with the MYBL1 silencing-target (b) perform transcriptome analyses of the constructs using DNA microarray followed by (c) bioinformatic analyses to identify genes affected by the knockdown process. The control and MYBL1 lentiviral constructs were purchased from Origene.com, and transformations were performed at Texas Southern University. The control construct led to MYBL1

expression in the TNBC and the MYBL1 silencing-targeted-construct demonstrated substantially low MYBL1 gene expression. We were able to compare TNBC with MYBL1 expression (i.e., control) vs TNBC without MYBL1 gene expression. Transcriptomes were examined using a high-throughput Affymetrix DNA microarray platform. Bioinformatic analyses was performed to identify differentially expressed genes. Of course, *MYBL1 and MYBL2 were substantially knocked down, along with* genes *both up and downregulated by the targeted knockdown process.* As expected, we also observed an enrichment of genes associated with proliferation and cell cycle regulation. cMYB along with estrogen, progesterone and ERRB2 receptor genes are not affected because the TNBC cell line used for knockdown is negative for these genes; expression levels of these and other genes served as positive control of our procedural approach. We are currently repeating every aspect of our preliminary knockdown experiment (which includes knockdown, microarray and bioinformatic analyses).

Experimental outline for the current study:

MYBL1 gene is located on the chromosome 8q13.1 locus. We identified approximately 650 genes (out of 186,000 transcripts, expressed sequence tags, miRNA and redundant DNA probes) differential expression in MYBL1 control vs the MYBL1 gene knockdown. A more detailed outline of the methods will be addressed below, but in summary, as part of the 650 genes, we identified 14 differentially expressed genes. Further analyses of the 14 genes show that all are located at the chromosome 8q loci. The current project involves validation of these genes as differentially expressed (with MYBL1) in TNBC. When MYBL1 was knocked down, other genes identified as located on 8q chromosome were also affected by the knocked down; we refer to these as our candidate

genes. *Some genes were down-regulated when MYBL1 was knocked down, while other genes were up-regulated when MYBL1 was knocked down.* For the current project we examined transcript expression levels of our candidate genes in TNBC as validation of the knockdown and subsequent knockdown results. Based on the knockdown study, for genes that were down regulated with MYBL1, we expect the genes to show a similar pattern of dysregulation with MYBL1 via transcript analyses in untreated cell line preparations. For genes upregulated when MYBL1 is knocked down, we expect opposite gene expression levels when these analyses are performed in untreated TNBC cell lines and patient samples.

Screening genes identified as differentially expressed upon knockdown of MYBL1:

The current study began as an analysis of 14 genes that were both differentially expressed (when MYBL1 was knocked down), co-localized to chromosome 8q regions near the MYBL1 gene 8q13.1 locus. Bioinformatic analyses of patient samples show that all 14 genes were mutated in a substantial number of the same patient samples as MYBL1. The current project involves (a) further bioinformatic analyses of the MYBL1 and the other 14 genes and (b) experimental analyses of the transcript and protein levels of the various genes.

Related to the current study-Genes reproducibly affected by targeted lentiviral knockdown of MYBL1

All 14 genes did not validate, as a result, only genes reproducibly validating as differentially expressed are described below. The other genes (included in the set of 14) will be explained in the Results section of this document.

MYBL1

The description of MYBL1 gene is outlined in the sections above. MYBL1 was the target of the knockdown process. The gene is located at position 8q13.1 chromosomal

locus. MYBL1 is a strong transcription activator and protooncogene involved in cell differentiation, proliferation and apoptosis related signaling events.

VCIP1

MYBL1 is localized to chromosome region 8q13.1. Following knockdown of MYBL1 we observed VCPIP1 was downregulated with MYBL1 and located the same region of 8q13.1. Further analyses show VCPIP1 significantly co-expressed with MYBL1 in breast cancers patient samples. Based on these observations, we performed bioinformatic analyses of our MYBL1 knockdown dataset in search of additional genes localized to the chromosome 8q loci. ***The VCPIP1 gene is the subject of another study and will only be referenced in this document.*** VCPIP1 is a key gene involved in the assembly of Golgi and Endoplasmic reticular during p97/p47-mediated membrane fusion events related to mitosis. VCPIP1, initially name VCPIP135 (valosin containing protein [VCP][p97]/p47 complex-interacting protein, p135) binds to the p97/p47/syntaxin5 complex during mitosis and is necessary for p97 (VCP)-mediated reassembly of the Golgi stacks after similarity. VCPIP1 is also involved in ATM or ATR DNA repair events related to targeting DNA-protein cross-links.

MYC

MYC gene is located at chromosomal locus 8q24.21 near the end region of the chromosome. The gene is downregulated with MYBL1 knockdown. MYC (V-Myc Avian Myelocytomatosis Viral Oncogene Homolog) is a proto-oncogene and strong transcriptional activator that codes for a nuclear phosphoprotein. The gene is involved in cell cycle progression and apoptosis signaling events, and frequently identified as mutated and amplified in various cancers. Translocation events have been identified in Burkitt

lymphoma and multiple myeloma human cancers. There is compelling evidence of MYC and MYBL1 co-expression in tumors (41), apoptosis (41) and cell progression to S phase of cell cycle signaling (41). The is frequently amplified in cancers, however translocations are observed in multiple myeloma and Burkitt lymphoma (41).

BOP1

Block of proliferation 1 (BOP1) is located in chromosomal region 8q24.3. The gene is involved in RNA binding activities related to cell cycle proliferation and regulation of p53 signal transduction (34). BOP1 expression is upregulated in prostate tissue and the PC3 cell line (Wu, jing, jing). Knockdown of the gene activates apoptosis and inhibits metastasis of PC3 prostate cells in vitro. BOP1 is also overexpressed in gastric (80), lung (80), pancreatic (80), colon (80) and breast cancer (80). Data show the gene associated with chemoresistance in TNBC by promotion of Beta Catenin acetylation (34), and promotion of cell migration and invasion of tumor progression in colon cells (34) via the JNK pathway (34).

MAF1

MAF1 gene is a homolog to the Yeast MAF1 gene which is a negative regulator of RNA polymerase III (GC). The human MAF1 gene is located on chromosome 8q24.3. The gene is associated with poor prognosis and colon tumor progression (34). Mechanistically, MAF1 is involved in repression of RNA polymerase III mediated transcription, playing a key role in cell fate determination events in early mesoderm induction.

CHAPTER 3

DESIGN OF THE STUDY

Cell Lines and Patient Datasets:

he cell lines utilized in this study were purchased from the American Type Culture Collection (ATCC®) atcc.org and utilized within 6 months of purchase. The cell lines used for the study are MCF10A (derived from benign proliferative breast tissue; triple negative non-tumor) and HS578 and MDA MB231 both tumorigenic TNBC cell lines. The cells were feed twice weekly with Dulbecco's Modified Eagle Minimum essential media (DMEM) supplemented with 1% penicillin and 10% serum and grown in a 37OC incubator with an additional 5% CO2. The cells were grown until 80-90% confluent and then typsinized using a 0.25% trypsin solution for sub-culturing.

The patient datasets were available on the Cbioportal.org online database. Approximately 2500 breast cancer patient samples were interrogated based on gene expression analyses using the data source.

Ribonucleic Acid (RNA) Isolation:

Cell lines in this study were grown to approximately 90% confluency in T75 dishes and RNA was extracted using a 1 mL Trizol solution. Trizol extraction was performed as suggested by the manufacturer. In summary, 200ul of chloroform was added to the Trizol/cell preparation. The suspension was mixed and centrifuged at 10,000

revolutions per minute for separation of the total RNA preparation. The top RNA layer was placed into a clean tube with 500 µL of 95% ethanol and stored at -20ºC for 30 minutes. The RNA preparation was collected by centrifugation, 30 minutes at 10,000 RPM. The RNA pellet was desiccated, resuspended in 20 µL of clean water then heated at approximately 60ºC for 1 minute. A 10 µL aliquot of the RNA preparation was removed for A260/280 analyses and RNA gel electrophoresis. The gel was prepared by adding 1 gram of agarose to a 1x solution of 3-(N-morpholino) propane sulfonic acid (MOPs) buffer, 2 uL of 7% formaldehyde and 2 µL of ethidium bromide. Samples displaying intact 28S/18S profiles and A260/280 ratios of 1.8-2.0 were used for downstream studies.

Generating Complementary DNA (cDNA)

The iScript cDNA kit from Bio-Rad (Bio-Rad, Hercules CA, USA) was used for conversion of mRNA to cDNA. A 20 µL mixture containing 5x iScript buffer, reverse transcriptase, random hexamers, total RNA, the Oligo dT mixture, deoxyribonucleic triphosphate acids (dNTPs) and water was placed at 45^{O}C for an hour. The reverse transcriptase enzyme was deactivated by heating the samples at 85degC for 1 minute. The mixture was chilled and 80 µL of water was added. The cDNA was stored at –20^{O}C until it was utilized for polymerase chain reactions (PCR) or -80 degrees for long term storage.

Generating PCR Gene Primer Sets

The primer 3^{TM} (42) program (http://bioinfo.ut.ee/primer3-0.4.0/) was used to generate primers for each of the 14 target genes. The sequences for only 4 final candidate genes: GAPDH (control), BOP1, MAF1, MYC and MYBL1 are given in this document (see below). Affymetrix NetAffx (20) was used to obtain the nucleotide sequences of our

target genes (http://www.affymetrix.com/estore/analysis/index.affx). Primer sets generated using Primer3 program were further analyzed using the Genome Browser tools available on the University of California Santa Cruz website (https://genome.ucsc.edu/) in silico PCR analyses program. The primer-sets were synthesized by IDTDNA.com (Coralville, Iowa), manufactured, then shipped to Texas Southern University within 48 hours of purchase. Primer sets for all 14 genes analyzed in this study were generated and used for analysis of triple negative samples. However, only sequences for primer sets of genes demonstrating reproducible results are included below. The primer sequences for the 4 genes (plus the control gene) used in this study are given in Table 1.

Table 1: PCR Primer Sequences

Gene	Left Primer	Right Primer	Size of Amplicon (Bp=base pair)
MYBL1	TGGATAAGTCTGGGCTTATTGG	CCATGCAAGTATGGCTGCTA	210BP
MYC	TGGCTGCTTGTGAGTACAGG	TGAACTGGCTTCTTCCCAGG	229BP
BOP1	CACCCCCAGCTTCTATGACC	CTGGATGAAGCGTCCGTAGG	256BP
MAF1	CAGCTTCAACCAATGCCTGG	GTGTGCAGGCCCAGTAGAAA	297BP
GAPDH	TCCCTGAGCTGAACGGGAAG	GGAGGAGTGGGTGTCGCTGT	217BP

Polymerase Chain Reaction (PCR):

PCR was performed to determine and display the differential gene expression levels between the different cell line preparations. The PCR reactions consisted of 2 µL cDNA, 2ul of forward and reverse gene primer sets, 10 µL of 2x DNA polymerase I TAQ

polymerase master mix which contained the TAQ enzyme and buffer and dNTPs. The samples were placed in PCR quality tubes and analyzed using the Bio-Rad Thermal Cycler (Hercules California). PCR cycling consists of (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 PCR products were analyzed using a 2% agarose gel. The gel consisted of 100 mL of 1X Tris/Borate/EDTA (TBE) buffer with 2 grams of agarose and 1ul of ethidium bromide. A 10 μL aliquot of the PCR product and 2 μL of the sample buffer were loaded onto the gel for electrophoresis. The gels were run up to 1-hour at 98volts.

Densitometer of the PCR Results

The LI-COR Imaging System (Lincoln Nebraska) was used to visualize the samples on the agarose gel. To determine the different levels of transcript, present in the different samples, the LI-COR software was also used to assess the densities of the amplicons present on the agarose gels. The densitometer plots were each normalized against control GAPDH levels.

Western Blotting

The western blot procedure was performed as outlined in a previous document (62). Antibodies used in the current study are outlined below: *Antibodies*: MYBL1 anti-mouse antibody was utilized at a dilution 1:500 (SAB14002280; Sigma Aldrich/Millipore, St. Louis, MO). Actin anti-rabbit antibody was used at a $1{:}10^4$ dilution (SAB14002280; Sigma Aldrich / Millipore, St. Louis, MO). MYC, MAF1 and BOP1 anti-mouse antibodies were

purchased from Santa Cruz Biotechnology Santa Cruz CA. MYC anti-mouse antibody (SC-40) was used at 1:500 dilution. BOP1 anti-mouse antibody (Sc-365595) was utilized at a 1:500 dilution. MAF1 anti-mouse antibody (Sc-515614) was utilized at a 1:500 dilution. Secondary HRP conjugated anti-mouse antibody (HAF007) and anti-rabbit (HAF008) antibodies were purchased from R and D Systems (Minneapolis, MN). Western blotting filters were processed and developed using the Clarity Western ECL substrate (Bio-Rad, Hercules, CA, USA) on a LICOR digital imaging system (LI-COR Biotechnology, Lincoln, NE).

Data Analysis Using cBioPortal Open-source Database

The cBioPortal.org website was used to analyze our genes against patient samples. The website is an interactive open-source website that contains thousands of curated cancer patient samples collected from a number of different database services, originally developed at Memorial Sloan Kettering Cancer Center (MSK). Current data is collected from MSK, Dana Farber Cancer Center, Princess Margaret Cancer (Toronto), and Children's Hospital of Philadelphia and more. Breast samples utilized in this study were part of the (a) METABRIC dataset (published in 2012 and 2016; 2509 patient samples) and (b) TCGA Breast Invasive Carcinoma Pan Cancer dataset (1084 patient samples). Graphic display of the OncoPrint analyses allows for mutation and gene expression analyses of our genes in the various patient datasets. Co-expression analyses were done using this approach.

Data Analyses of Protein-protein Interaction Using STRING Open-source Database

The Search Tool for the Retrieval of Interacting Genes/Proteins (STRING^{TM}) was utilized to determine the possible functional relationship between the gene validated as part of this study. STRING^{TM} analysis online database tool utilizes millions of data points validated by experimental, text mining and bioinformatic websites to predict protein-protein interactions of gene-lists.

CHAPTER 4

RESULTS AND DISCUSSION

MATERIALS AND PROCEDURES PERFORMED PRIOR TO THIS STUDY BUT CRITICAL TO INTERPRETATION OF THE RESULTS

MYBL1 gene knockdown and selection of differentially expressed list of candidate genes to study.

The primary goal of our laboratory is to characterize TNBC. Towards this goal, we performed DNA microarray analyses and identified MYBL1 as over-expressed in ~ 25% of TNBC (17). MYBL1 gene is a strong transcription factor and controls expression of a number of genes with a range of functions associated with tumor progression events. In efforts to better understand the possible role of MYBL1 in TNBC our first approach was to knockdown the gene and determine genes that might be co-operating with MYBL1 to affect the genotype of the cancers. Because MYBL1 is a strong regulator, we decided to knockdown MYBL1 in MDA MB231 TNBC cells and determine genes either directly or indirectly affected by the knockdown process. The shRNA lentiviral process was chosen for gene knockdown. For the detailed approach and analyses procedures leading to selection of the candidate genes see the previous publication (17). As summary, several regions of the MYBL1 gene were examined for the ability to knockdown (i.e., down-regulate) the MYBL1 in TNBC. The MYBL1 shRNA lentiviral particles and the scramble

control particles were purchased from Origene (Cat # TL303089V; Rockville, Maryland). A control (lentiviral without MYBL1 target) and the MYBL1 knockdown preparations were analyzed using DNA microarray. The microarray platform allows for analyses of the transcriptome (~186,000 transcripts, splice variants, expressed sequence tags and small RNAs) (Figure 4). Genes differentially represented in the Control vs the MYBL1 targeted gene preparation were initially processed using the Affymetrix TAC 4.0 software (Thermo Fisher Scientific, Waltham Massachusetts), then later using gene-set enrichment and transcription factor analyses tool available on Molbio-tools (http://www.molbiotools.com/listcompare.html) (25). Genes differentially expressed in the control vs MYBL1 knockdown gene-chips were selected based on the 'industry standard protocol', in that (a) genes must demonstrate at least a 2-fold difference in transcript levels and (b) demonstrate p-value of < 0.05. The genes could be easily sorted based on these criteria and datasets generated. A final list ~ 650 genes (starting from 186,000 probe-sets) were considered our final list of candidate-genes-of-interest. Knockdown of the MYBL1 gene in TNBC cell line led to genes both downregulated with MYBL1 expression and upregulated when MYBL1 was knocked down. Some of the genes are likely directly regulated by the MYBL1 transcriptional activator, while other genes are likely affected by regulation of genes that MYBL1 directly regulates. Genes utilized in this study represent a subset of the original 650 differentially expressed candidate gene dataset. As shown in Figure 4 the disabled sense strands complementary to the target sequence are on the genechip. Gene-chips are composed of 186,000 transcripts, splice variants and small RNAs displayed as blue dots in the center figure. Total RNA is tagged and hybridized to the genechip; one genechip per patient sample. High copy number is presented as densely

colored / white spots as demonstrated in the magnified pane. Control sequences spotted at known concentrations were supplied on the gene-chips to allow for quantitative calculations (20).

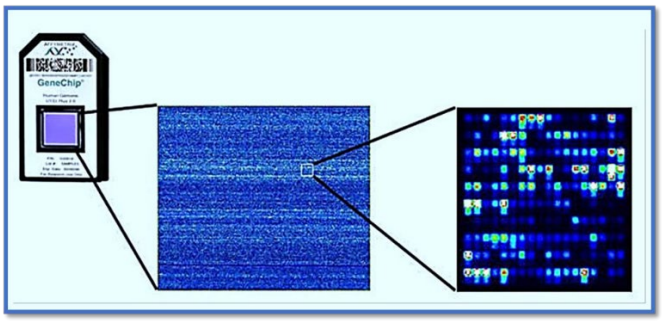

Figure 4: The Affymetrix Microarray Genechip

RESULTS AND DISCUSSIONS OF STUDY

The focus of our laboratory is to further characterize MYBL1 gene in TNBC cells and patient samples. The MYBL1 gene is a proto-oncogene and strong transcriptional activator. The gene is also associated with many of the events' key to tumor progression such as cell cycle regulation, differentiation and apoptosis. Since we identified MYBL1 as over-expressed in ~25% of TNBC it was important for us to determine the role of the gene in the breast cancers. The first approach towards our goal was to identify genes that might be either directly or indirectly cooperating with MYBL1 to affect the genotype and phenotype of TNBC. To achieve this goal, we knocked down MYBL1 in a TNBC cell line, and identified genes significantly affected by decreased expression of MYBL1 gene levels in the knockdown cells compared to control preparations based on T-test analyses (64). Genes were both down-regulated and up-regulated when MYBL1 gene expressed was decreased by knockdown. A total of ~650 genes were identified as differentially expressed

by this process and labeled as candidate-genes-of-interest. The genes examined in the current study represent a subset of the 650 genes.

How were the genes identified as associated with chromosome 8q loci?

We entered all 650 genes into the cBioPortal.org analysis asking 1 question… based on the genes in the 650 list, are any of the genes co-expressed with MYBL1 in patient samples? The experimental outline is shown in Figure 5). The answer was yes. The cBioPortal results (following analyses of the 650 genes) are outlined in Figure 6. There is not complete concordance, but an estimated half of the patients show mutations similar to MYBL1 gene. Some genes showed nearly complete concordance with MYBL1. Notice that for BPNT2 and GGH there is nearly complete concordance with MYBL1; the same patients show MYBL1, BPNT2 and GGH amplification. The red bars for each gene correspond to gene amplification events identified in the mutated genes. Previous data validate frequent amplification of MYC gene in various human cancers. There is no published scientific evidence of amplification of BOP1 or MAF1 gene. Very few mutation deletions (i.e., blue bars) are identified in MYBL1 and the other genes; see yellow arrow in Figure 6.

The bottom panel of Figure 6 shows the pattern when MYBL1 is mutation in a particular set of patients compared to two other genes. In the bottom panel (Figure 6b, MYBL1 gene mutations are compared to MYBL2 and cMYB mutations in the same patient samples. MYBL1, MYBL2 and cMYB belong to the same gene family. Analyses of the same patient population show little relationship between mutations in MYBL1 compared to MYBL2 and cMYB; hence patients with MYBL1 amplification mutations do not show evidence of amplification mutations (or other type mutations) in MYBL2 or cMYB.

Published data by other investigators show that MYBL2 can be mutated in some of the same patients as MYBL1, compared to cMYB. These data appear to support those observations because there are more patients (hence red bars) showing concordance between MYBL1 and MYBL2 compared to cMYB gene. And, based on our observations, MYBL1 gene is over-expressed in TNBC and our and other studies show TNBCs are negative for cMYB. Of importance related to the cBioPortal analysis- the patient samples are not stratified based on 'the type of breast cancer'. **All breast cancer types, not just TNBC are included in the combined datasets and analyses.** All of the genes that demonstrate amplification in the same patients as MYBL1 (in cBioPortal) were identified as located to chromosome 8q loci (Table 2). There were 2509 patients in 2 combined curated datasets in the analyses. Datasets were previously published in Nature 2012 and 2016.

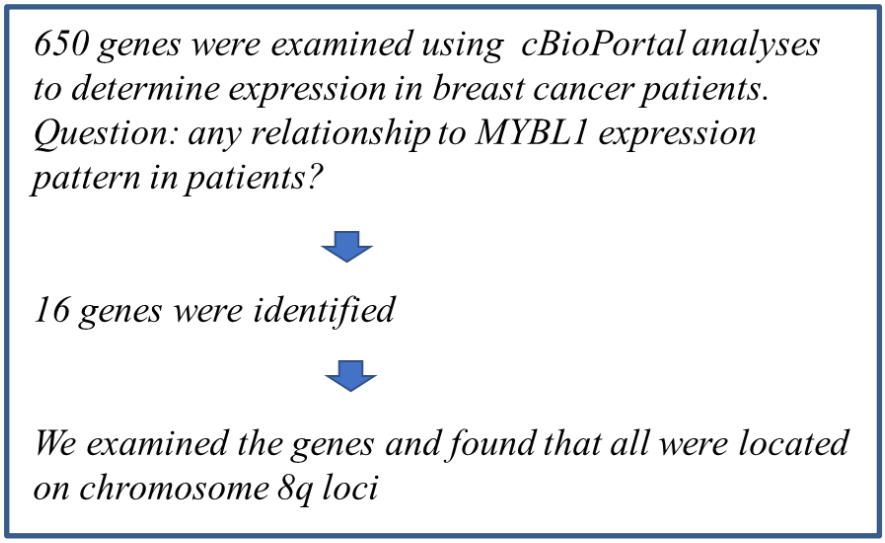

Figure 5: Experimental Outline Summarizing Identification of Genes on Chromosome 8q loci

In Figure 6 the red bars represent patients, and red represents amplifications. The genes are listed on the far-left side of the image. The thick arrow points to the MYBL1 gene; the red bars represent the patients in which MYBL1 is amplified. The blue bar separates patients which display amplifications in MYBL1 and our 14 genes. Note the overlap with MYBL1 patients. Nearly all of the mutations are amplifications (red bars). Few mutations are deletions (blue bars designated by gold arrow) (b) The bottom panel demonstrates the pattern of amplification in genes **not** dysregulated with MYBL1 gene (thick blue arrow). The genes include MYBL2 and cMYB, both belonging to the same gene family as MYBL1.

Figure 6: **cBioPortal Analyses of the (650) Genes Led to Identification of 14 Genes Amplified in the Same Patient Samples as MYBL1**

Table 2: Fourteen Genes Demonstrating Amplifications in the Same Patients as MYBL1

Table 1.	shRNA scramble control (RNA levels)	mybl1 knock down (RNA levels)	FOLD CHANGE	CHROMOSOME
	LOG 2 VALUES	LOG2 VALUES	(up or down regulated after MYBL1 KD)	
MYBL1	12.2	10.2	4X KD	**8q13.1**
RMDN1	10.87	12.69	6.8X up	8q21.3
MRPL13	10.12	12.89	6.8X up	8q24.12
MAF1 GENE	12.12	10.21	3.7X down reg	**8q24.3**
SDC2	12.61	10.9	3X down reg	8q22.1
GGH	10.1	12.5	5.2X up	8q12.3
MYC	12.7	10.6	4.4X down reg	**8q24.21**
BOP1	11.3	8.5	7X down reg	**8q24.3**
PUC60	10	8.3	3X down reg	8q24.3
SNX16	9.66	11.4	3.3X up	8q21.13
IMPAD1/BPNT2	11.44	9.3	4X down reg	8q12.1
AZIN1	15.2	12.9	4.8X down reg	8q22.3
EBAG9	12.4	14.4	3.9X up	8q23.2
MRPS28	10.3	12.1	3.X up	8q21.13
VCP1P1	10.1	9	2.1X down reg	**8q13.1**
UBXN2B (p37)	NOT ON LIST			8q12.1

In Table 2 all of the genes are located to chromosome 8q loci. KD= knocked down (or downregulated) when MYBL1 is knocked down. Up= up regulated when MYBL1 is knocked down; for example, when MYBL1 is knocked down (or silenced), that gene is upregulated. Table 2 shows the results from the knockdown study. Note that all of the genes identified via cBioPortal in patient samples are located to chromosome 8q loci. Many are localized to the end region of 8q closer to MYC than to the MYBL1 locus. This observation will be addressed later. We cannot draw conclusions at this point, but it appears that genes closer to MYBL1 locus show a more similar pattern in patient samples.

The transcript levels in Table 2 were obtained from the DNA microarray of the knockdown experiment. The DNA microarray detects the normalized transcript levels in the control lentiviral gene-chip compared to the gene-chip processed using the shRNA lentiviral containing the MYBL1 targeted (knockdown) preparation. The fold change represents the normalized value of the transcript levels of the control versus the knockdown values. All of the values were generated by the TAC4 software. Notice that some genes are

upregulated when MYBL1 is knocked down (like RNDN1) while most are downregulated when MYBL1 is knocked down. From these data it is impossible to determine if genes affected by MYBL1 knockdown are directly regulated by MYBL1 transcription factor binding. For genes up-regulated when MYBL1 was knocked down, it could be 'release of some sort of regulation' or the effect of MYBL1 regulating some other gene that regulates the up-regulated gene. This could be the cause of either down-regulated or up-regulated genes. Since MYBL1 is a transcription factor and transcription factors directly bind to and regulate transcription, we are analyzing a panel of genes for the presence of MYBL1 transcription factor binding sites to address this question.

As a summary of Table 2, (a) all of the genes from cBioPortal were localized to chromosome 8q loci (b) some genes were upregulated when MYBL1 was knocked down (like RMDN1) and (c) most of the genes are localized to the end of the chromosome, some distance from the MYBL1 locus.

When validating the microarray, it is important to perform transcript analyses using the same sequence utilized to on the gene-chip. We generated primer-sets for all of the genes in Table 1 and performed conventional PCR to compare transcript levels in the non-tumorigenic MCF10a cell line compared to MDA MB231 TNBC (which was used for the knockdown study). The data were inconclusive in that all of the genes did not validate (i.e., the results were not consistently reproducible). Genes that did not validate on all PCR attempts were excluded. PCR results were reliability reproducible for MYBL1 (as expected), MYC, BOP1 and MAF1. Microarray data showed that when MYBL1 was knocked down, MYC, BOP1 and MAF1 were each downregulated. That means that in untreated lentiviral samples, MYBL1, MYC, BOP1 and MAF1 should demonstrate the

same pattern of expression in TNBC preparations, which is what we observed following PCR. MYBL1, MYC, BOP1 and MAF1 demonstrated high levels in TNBC (MDA MB231) compared to non-tumor triple negative (MCF10a) Figure 7. PCR and densitometer analyses of the four genes are shown in Figure 7b. More than 20 experiments were performed as validation. A summary of the function of the four genes are summarized in Table 3. Note that MBYL1 and MYC are both transcriptional activators. Figure 8 demonstrates the chromosomal loci for each of the candidate gene. VCIP1 (which is only referenced in this study) is the only gene that was localized immediately neighboring MYBL1 gene at 8q13.1 locus.

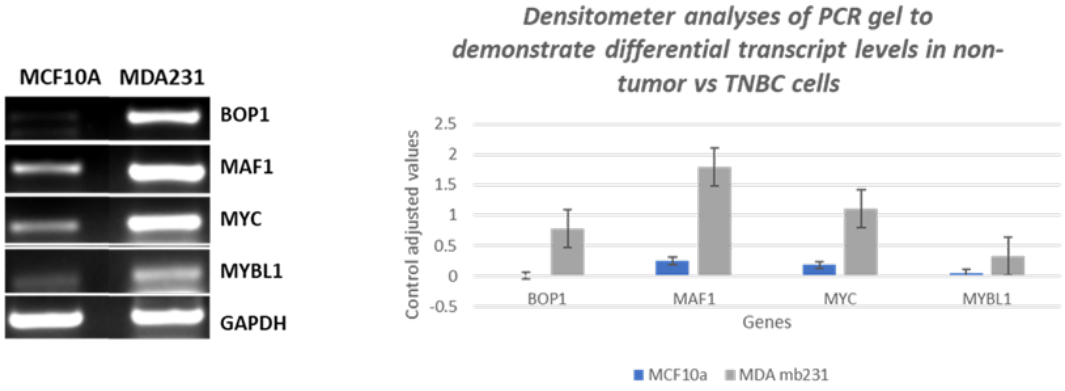

Figure 7: **PCR Validation of MYBL1, MYC, BOP1 and MAF1. MCF10a is a Non-Tumor Triple Negative Cells and MDA MB231 is TNBC**

The agarose gel profiles were analyzed via densitometer analyses and plotted in the (b) panel.

Table 3: Summary of MYBL1, MYC, BOP1 and MAF1 Function

GENE	FUNCTION
Mybl1	Proto-oncogene, transcription factor regulating cell proliferation, differentiation
MYC	Proto-oncogene, transcription factor, cell cycle , apoptosis, differentiation
BOP1	Regulation of cell cycle, signal transduction by p53 class mediator
MAF1	Negative effector of RNA polymerase III

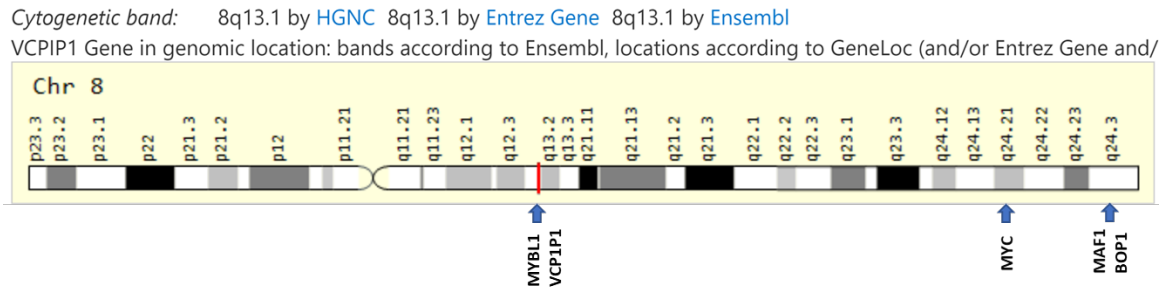

Figure 8: Chromosome loci for MYBL1, MYC, BOP1 and MAF1

Note that BOP1 and MAF1 are closer to MYC locus. Figure obtained from GeneCards.org.

Western analyses for MYBL1, MYC, BOP1 are consistent with the PCR transcript analyses, in that higher levels of protein are detected in TNBC compared to non-tumor triple negative MCF10a sample preparation (Figure 9). The MAF1 protein levels were not detected using western analyses. We suspect this could be due to our choice of antibodies to use for the assay. We will continue to process the samples using different commercial antibody sources. A decision was made to include an additional TNBC cell line in the

protein analyses as validation of the differential pattern of gene expression in TNBC. The HS578 cell line is a TNBC similar in genomic expression to the MDA MB231 TNBC cell line. We recognize that additional TNBC samples should be examined as part of this study to validate gene expression in TNBC compared to other breast cancer subtypes. Additional TNBC cell line samples will be included as part of this study, in future experiments.

Data by other investigators show that BOP1 and MAF1 are regulated by MYC. So, it could be that MYBL1 regulates MYC (7) and MYC regulated BOP1 and MAF1. The transcript data for MYBL1, MYC and BOP1 are validated by western blot analyses of the triple negative cell lines (Figure 9), but not for MAF1. We are not sure why MAF1 protein levels will not validate, but we suspect our selection of the antibody. The corresponding densitometer plot for the western is given in Figure 9b.

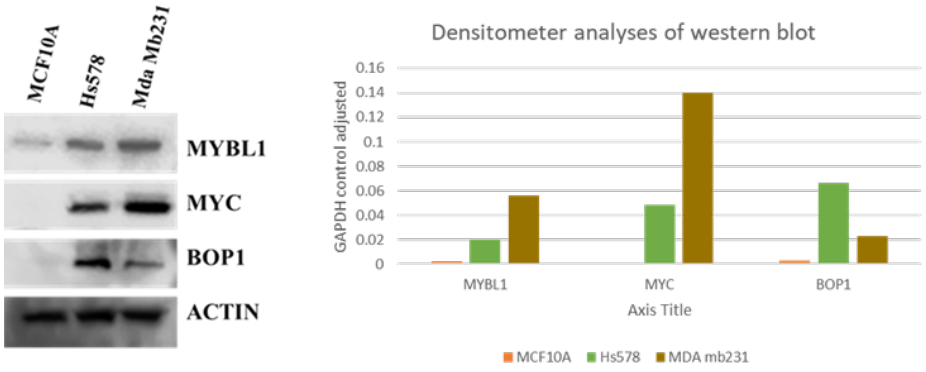

Figure 9: Western Blot Protein Analyses of MYBL1, MYC, BOP1 and MAF1 and Actin Control (a) Western blot (b) Densitometer of Western Blot

As further analysis of the possibility that MYBL1, MYC, BOP1 and MAF1 are involved in documented protein: protein interactions, the genes were processed using the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING^{TM}) program (Figure 10). STRING graphically displays 'linkages' (i.e., connections via lines) when there is documented evidence of some sort of relationship or protein: protein interaction between specific genes. STRING presented here shows evidence of protein interactions between MYBL1, MYC and BOP1.

As an example of the type of data utilized by STRING, investigators show that MYBL1 can regulate MYC in mouse cell (Golay, M Introna, G E Sonenshein), and MYBL1 and MYC overexpression may be risk factors for disease free survival in Adenoid cystic carcinoma (Nobuhiro Hanai , Yasuhisa Hasegawa , Hiroshi Inagaki); and MYBL1 and MYC are co-expressed in pharynx cancer. These types of data are presented in STING as linkages between individual genes such as the linkages (i.e., lines) displayed between MYBL1 and MYC. Related to our knockdown data, because MYBL1 and MYC are both decreased in the knockdown study, we think that MYBL1 directly binds to and downregulates MYC expression in TNBC.

There is also evidence of a relationship (hence a linkage) between BOP1 and MYC genes. Data show that BOP1 is a target for MYC regulation (Ulrich H. Weidle, Dirk Eick), and over-expression of MYC activates BOP1 in colorectal cancer cell lines (Sha Li, Pei Xue, Xun Diao). As for MAF1, bioinformatic analyses show MYC binding site in the MAF1 promoter region, while experimental validation is lacking. There are fewer validating studies in the literature supporting associations between MAF1 and MYC genes.

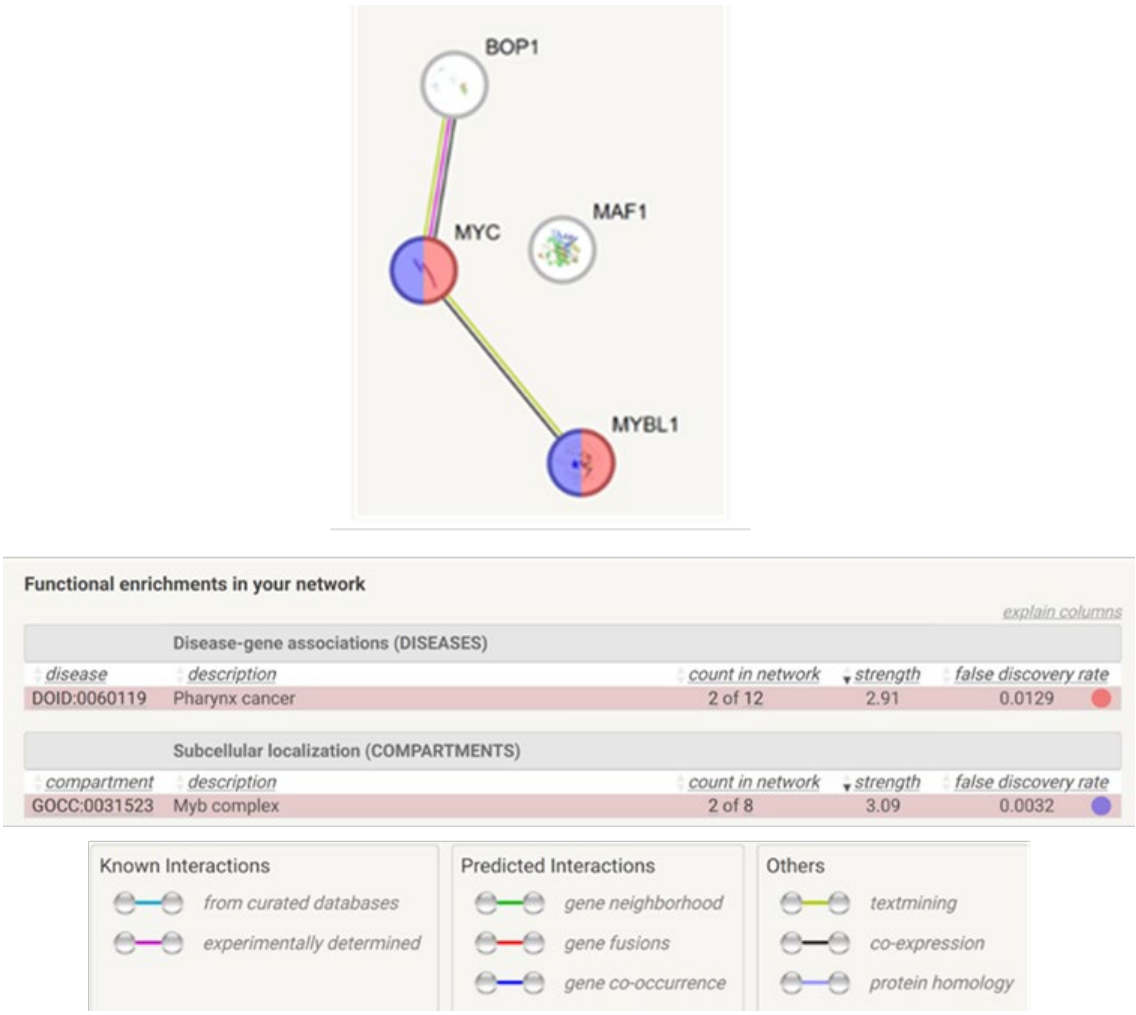

Figure 10: STRING Analyses of Protein: Protein Interaction between MYBL1, MYC, BOP1 and MAF1. Data based on experiments, text-mining and co-expression analyses. See colorimetric legend.

Literature Search to Examine the Significance of Candidate Genes Localized to Chromosome 8q loci.

There is a substantial amount of data generated based on chromosomal copy-number analyses of Basal-like / TNBC samples. Many of the data cited in this section of the thesis are consistent with our data in that (at least from an observation point of view) chromosomal loci on 8q are amplified in breast cancers and in TNBCs. Many of the studies

sited below conclude that (a) large segments of chromosomal 8q are often amplified in breast cancers and (b) regions near the end of the 8q loci are more frequently amplified than regions nearer the centromere. Many of these studies are validated by either Fluorescent In situ hybridization (FISH) or Comparative Genomic hybridization (CGH). These studies support the observations that amplification of chromosome 8q (specifically near 8q24) is common in breast cancers and associated with tumor progression. We would like to contribute to this conversion and suggest that more upstream loci be considered as 'potential amplification hotspots', including the MYBL1 gene and its genomic region at the 8q13.1 position.

As mentioned, a number of studies support data showing that 8q is a hot spot frequently amplified in breast cancers. Milioli et al examined genomic and transcriptional analyses of basal-like/TNBC samples and found significant gains/amplifications associated with chromosome regions 8q13.2-8q24.22 loci. The authors showed that cancers with mutations in 8q were high grade (i.e., poorly differentiated), with frequent MYC over-expression. No mention of MYBL1 in their study, but MYBL1 is located at 8q13.1 and data show the gene is amplified in breast cancers. In another study, Rummukainen et al. examined chromosome 8 in 16 breast cancer cell lines including receptor positive and TNBC subtypes. The authors found high-level amplifications in 8q in all of the breast subtypes. After further analyses, they concluded that genes on 8q played a significant role in breast cancer pathogenesis; but MYC likely played a less significant role.

Also consistent with results of this thesis, Nupponen et al. show that genes in regions 8q21-8q24 are frequently amplified and overexpressed in both breast cancer. Nupponen identified a similar pattern of amplification in prostate cancer. More directly

consistent with our theories, Parris et al examined invasive breast cancers and identified amplifications spanning regions from 8q12.1-q24.22. The authors suggest that genes located between these loci play a pivotal role in the aggressive phenotype in breast cancers.

Collectively data cited in this section of the thesis support the need to better understand and define genes localized to the chromosomal 8q loci in all breast cancers. Our studies focus on MYBL1, and genes directly and indirectly associated with the gene. We suggest that MYBL1 and MYC and BOP1 be considered for their combined contributions to the TNBC genotype. This is a preliminary study, culminating in observations that are in part supported by previous investigators. Even though the genes in Table 2 do not reproducibly validate, we have confidence that several of them (do in fact) function with MYBL1, contributing to the TNBC phenotype. We are currently (a) repeating our knockdowns (adding additional TNBC cell lines to the study) (b) repeating the microarrays and (c) data analyses of the results presented here. The immediate goal of our laboratory (in line with the study presented here) is to characterize MYBL1 in TNBC and ultimately determine the role of the gene in tumorigenesis.

CHAPTER 5

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

In summarizing the following conclusions and recommendations are made:

1. Knockdown of MYBL1 in TNBC results in identification of an enrichment of genes on chromosome 8q; all of the genes appear to be and amplified. These data are consistent with other data supporting observations that chromosome 8q is frequently amplified in TNBC and other breast cancers.
2. Both MYBL1 and MYC are strong transcription activators. This is the first observation that MYBL1 might regulate MYC in TNBC. Other investigators show that MYC can bind to and regulate MYBL1 expression in other cell lines. We will validate MYBL1 binding to MYC promoter in TNBC (via electromobility shift assay; via MYBL1 binding MYC promoter).
3. Our previous bioinformatic analyses show that many (not all) of our 650 candidate genes show an enrichment of MYBL1 transcription factor binding, suggesting MYBL1 direct binding to and regulating a subset of these genes. We will re-examine the 650 gene list to determine which genes do indeed have MYBL1 binding sites.
4. We suggest MYBL1 regulates MYC, and MYC down-regulates, BOP1 and MAF1 in TNBC.

5. We will reexamine cBioPortal patients using only TNBC samples for our most immediate study. This is key to our focus and will be done soon.
6. Negative MAF1 westerns are likely due to the poor performance of our antibody. We will try another antibody for Western (protein validation). We will revisit analysis of all 14 genes identified on the 8q loci.
7. We are currently repeating the knockdowns with more TNBC cell lines, which will later involve repeating the microarrays and data analyses.

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. Breast Cancer - Symptoms, Types & Risk Factors | MD Anderson Cancer Center." MD Anderson Cancer Center, https://www.mdanderson.org/cancer-types/breast-cancer.html/?invsrc=national&cmpid=BRE_ORG_infoaboutbreastcancer_G_SE&gclsrc=aw.ds&&k_clickid=81fbf833-02aa-421c-88ca-747e339ff485&gclid=Cj0KCQjw06OTBhC_ARIsAAU1yOXNA8gxjXffrk5fO4EZr1yTWEQzJrS_5WElTuvOwN249_PnMUsUqGgaAm3UEALw_wcB). Accessed 25 May 2022.
4. Bolcun-Filas, Ewelina et al. "A-MYB (MYBL1) transcription factor is a master regulator of male meiosis." Development (Cambridge, England) vol. 138,15 (2011): 3319-30. doi:10.1242/dev.067645
5. "How Common Is Breast Cancer?" American Cancer Society, https://www.cancer.org/cancer/breast-cancer/about/how-common-is-breast-cancer.html

6. Agarwal, V., et al., Predicting effective microRNA target sites in mammalian mRNAs. Elife, 2015. 4.

7. 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.

8. 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.

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

10. 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.

11. 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.

12. 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).

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

14. Cell Biology of Cancer | SEER Training.” Welcome to SEER Training | SEER Training, https://training.seer.cancer.gov/disease/cancer/biology/. Accessed 25 May 2022.

15. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, Jacobsen A, Byrne CJ, Heuer ML, Larsson E, Antipin Y, Reva B, Goldberg AP, Sander C, Schultz N. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012 May;2(5):401-4. doi:

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

17. Cole SW, Hawkley LC, Arevalo JM, Sung CY, Rose RM, Cacioppo JT. Social regulation of gene expression in human leukocytes. Genome Biol. 2007;8(9):R189 doi: 10.1186/gb-2007-8-9-r189. PMID: 17854483; PMCID: PMC2375027.

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

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

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

21. 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.

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

23. “Essential Role for the SANT Domain in the Functioning of Multiple Chromatin Remodeling Enzymes - PubMed.” PubMed, https://pubmed.ncbi.nlm.nih.gov/12419236/. Accessed 25 May 2022.

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

25. 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.

26. 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.

27. 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.

28. 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.

29. 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.

30. 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.

31. 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.

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

33. 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.

34. 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.

35. 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.

36. "InterPro." EMBL-EBI: EMBL's European Bioinformatics Institute | EMBL's European Bionformatics Institute, https://www.ebi.ac.uk/interpro/entry/InterPro/IPR001005/. Accessed 25 May 2022.

37. 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.

38. Kavarthapu, R., Anbazhagan, R., & Dufau, M. L. (2021). Crosstalk between PRLR and EGFR/HER2 Signaling Pathways in Breast Cancer. Cancers, 13(18), 4685. https://doi.org/10.3390/cancers13184685

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

40. 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.

41. 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.

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

43. 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

44. Krautkramer, Kimberly A et al. “Tcf19 is a novel islet factor necessary for proliferation and survival in the INS-1 β-cell line.” American journal of physiology. Endocrinology and metabolism vol. 305,5 (2013): E600-10. doi:10.1152/ajpendo.00147.2013

45. 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.

46. 46. Lee YM, Kim E, Park M, Moon E, Ahn SM, Kim W, Hwang KB, Kim YK, Choi W, Kim W. Cell cycle-regulated expression and subcellular localization of a kinesin-8-member human KIF18B. Gene. 2010 Oct 15;466(1-2):16-25. doi: 10.1016/j.gene.2010.06.007. Epub 2010 Jun 23. PMID: 20600703.

47. 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.

48. 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.

49. 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.

50. 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.

51. Liu, X., et al., Reassessing the Potential of Myb-targeted Anti-cancer Therapy. J Cancer, 2018. 9(7): p. 1259-1266.

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

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

54. 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.

55. 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.

56. 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.

57. "MYB Oncoproteins: Emerging Players and Potential Therapeutic Targets in Human Cancer | Oncogenesis." Nature, Springer Nature, https://doi.org/10.1038/s41389-021-00309-y. Accessed 25 May 2022.

58. 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

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

60. 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.

61. 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.

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

63. 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.

64. 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.

65. Qiu MJ, Wang QS, Li QT, Zhu LS, Li YN, Yang SL, Xiong ZF. KIF18B is a Prognostic Biomarker and Correlates with Immune Infiltrates in Pan-Cancer. Front Mol Biosci. 2021 May 24;8:559800. doi: 10.3389/fmolb.2021.559800. PMID:34109209; PMCID: PMC8182049

66. 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.

67. 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

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

69. Rolland, T., et al., A proteome-scale map of the human interactome network. Cell, 2014. 159(5): p. 1212-1226

70. 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.

71. 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.

72. "Silencing of KIF18B Restricts Proliferation and Invasion and Enhances the Chemosensitivity of Breast Cancer via Modulating Akt/GSK-3β/β-Catenin Pathway." Biofactor, 31 May 2021, https://iubmb.onlinelibrary.wiley.com/doi/10.1002/biof.1757.

73. 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.

74. 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.

75. 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.

76. 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.

77. 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.

78. Yang GH, Fontaine DA, Lodh S, Blumer JT, Roopra A, Davis DB. TCF19 Impacts a Network of Inflammatory and DNA Damage Response Genes in the Pancreatic

β-Cell. Metabolites. 2021 Aug 4;11(8):513. doi: 10.3390/metabo11080513. PMID: 34436454; PMCID: PMC8400192.

79. 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.

80. Wu W, Zhong J, Chen J, Niu P, Ding Y, Han S, Xu J, Dai L. Prognostic and Therapeutic Significance of Adhesion-regulating Molecule 1 in Estrogen Receptor-positive Breast Cancer. Clin Breast Cancer. 2020 Apr;20(2):131-144.e3. doi: 10.1016/j.clbc.2019.07.009. Epub 2019 Sep 3. PMID: 31669266.

81. 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.

82. 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.

83. Yang E, Guan W, Gong D, Li J, Han C, Zhang J, Wang H, Kang S, Gao X, Li Y, Yu L. Epigenetic silencing of UBXN8 contributes to leukemogenesis in t (8;21) acute myeloid leukemia. Exp Mol Med. 2021 Dec;53(12):1902-1910. doi: 10.1038/s12276-021-00695-8. Epub 2021 Dec 17. PMID: 34921223; PMCID: PMC8741982.

84. Yang, J., Ikezoe, T., Nishioka, C., Nobumoto, A., Udaka, K. and Yokoyama, A. (2013), CD34+/CD38− acute myelogenous leukemia cells aberrantly express Aurora kinase A. Int. J. Cancer, 133: 2706-2719. https://doi.org/10.1002/ijc.28277

85. Yersal, O., & Barutca, S. (2014). Biological subtypes of breast cancer: Prognostic and therapeutic implications. World journal of clinical oncology, 5(3), 412–424. https://doi.org/10.5306/wjco.v5.i3.412

86. 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.

87. 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.

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

89. 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.

90. Hall, W. J., Bean, C. W. & Pollard, M. Transmission of fowl leukosis through chick embryos and young chicks. Am. J. Vet. Res. 2, 272–279 (1941).

91. Boyle, W. J., Lipsick, J. S., Reddy, E. P. & Baluda, M. A. Identification of the leukemogenic protein of avian myeloblastosis virus and of its normal cellular homologue. Proc. Natl Acad. Sci. USA 80, 2834–2838 (1983).

92. Westin, E. H. et al. Differential expression of the amv gene in human hematopoietic cells. Proc. Natl Acad. Sci. USA 79, 2194–2198 (1982).

93. Nunn,M. F., Seeburg, P. H., Moscovici, C. & Duesberg, P. H. Tripartite structure of the avian erythroblastosis virus E26 transforming gene. Nature 306.