5-2024

# Experimental Analysis Of Mybl1, Vcpip1 And Ubxn2b In Triple Negative Breast Cancers

Fawaz Opeyemi Ibikunle
*Texas Southern University*

# EXPERIMENTAL ANALYSIS OF MYBL1, VCPIP1 AND UBXN2B IN TRIPLE NEGATIVE BREAST CANCERS

**THESIS**

Presented in Partial Fulfillment of the Requirements for

the Degree Master of Science in the Graduate School

of Texas Southern University

By

Fawaz Ibikunle B.S.,

Texas Southern University

2024

Approved by

Dr. Audrey Player
Chairperson, Thesis Committee

Dr. Mahesh Vanjani
Dean, The Graduate School

Approved by

| | |
|---|---|
| Dr. Audrey Player | 03/28/2024 |
| Chairperson, Thesis Committee | Date |
| Dr. Shodimu-Emmanuel Olufemi | 03/28/2024 |
| Committee Member | Date |
| Dr. Ayodotun Sodipe | 03/28/2024 |
| Committee Member | Date |
| Dr. Tuan Phan | 03/28/2024 |
| Committee Member | Date |

# EXPERIMENTAL ANALYSIS OF MYBL1, VCPIP1 AND UBXN2B IN TRIPLE NEGATIVE BREAST CANCERS

**By**

Fawaz Ibikunle, M.S.

Texas Southern University, 2024

Professor Audrey Player, Ph.D., Advisor

The focus of our laboratory is to characterize triple negative breast cancers. In a previous study we identified six genes over-expressed in a subset of the cancers. MYBL1 was one of the genes. MYBL1 is a strong transcriptional activator responsible for regulation of genes associated with cell cycle signaling, differentiation and apoptosis, all processes that drive cancers. Because of MYBL1's involvement in processes key to cancer pathogenesis, in an earlier study, we knocked down the MYBL1 gene in efforts to identify genes that were either directly or indirectly associating with the gene to affect the triple negative breast cancer processes. The implication is that MYBL1, being a strong transcriptional activator, can affect the cancer processes. Analyses of the knockdown dataset show that when MYBL1 is knocked down, the VCPIP1 gene is also knocked down. Both MYBL1 and VCPIP1 genes are located at the chromosomal 8q13.1 locus. And bioinformatic analyses of pan cancers and breast cancers show some of the same patients with MYBL1 alterations have VCPIP1 alterations, with almost complete concordance. The precise nature of the alterations in TNBC has not been identified, but the alterations in cell

lines lead to over-expression of the genes. Data mining via STRING™ analyses show high confidence associations between MYBL1, VCPIP1 and another gene UBXN2B, which is in a VCPIP1 mitotic signaling pathway. Although direct interactions have not been identified, all three genes are involved in mitotic signaling mechanisms. As summary, the current project describes bioinformatic and experimental data validating the possible relationship between MYBL1 and VCPIP1, and possibly the UBXN2B gene. Our data convincingly demonstrate an association between MYBL1 and VCPIP1 and we suspect this association is related to cooperation between the genes in mitotic signaling events.

## TABLE OF CONTENTS

## LIST OF TABLES

**Tables** Page

## LIST OF FIGURES

**Figures** Page

## LIST OF ABBREVIATIONS

| | |
|---|---|
| ACTN1 | Actinin alpha 1 |
| ADA2 | Adenosine deaminase 2 |
| ANGPT2 | Angiopoietin 2 |
| ATCC | American Type Culture Collection |
| ATR | Attenuated Total Reflectance |
| ATM | DNA repair |
| BL1 | Basal-like types |
| BL2 | Basal-like types 2 |
| CA | California |
| cDNA | Complementary DNA |
| CO2 | Carbon (ii) oxide |
| 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 |
| GEO | Gene Expression Omnibus |
| HER2 | Human epidermal growth factor receptor |
| HS578t | Human breast cancer cell line |
| IBC | Inflammatory Breast Cancer |

| | |
|---|---|
| IDC | Invasive Duct al Carcinoma |
| ILC | Invasive Lobular Carcinoma |
| IM | Immunomodulatory |
| IL | Interleukin 32 |
| KD | Knock-down |
| 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 |
| MA | Massachusetts |
| M | Mesenchymal group |
| MSL | Mesenchymal stem-like subtype |
| MCF10A | Receptor negative non-tumor cell lines |
| MDA-MB-231 | M.D. Anderson and Metastasis Breast cancer |
| 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 |
| N-COR | Nuclear receptor corepressor |
| NFIB | Nuclear Factor I B |
| NSFL1C | N-ethylmaleimide-sensitive factor (NSF) and valosin-containing protein |
| NE | Nebraska |
| P47 | NSFL1 cofactor |
| P97 | valosin-containing protein |
| PA | Pennsylvania |
| PCR | Polymerase Chain Reaction |
| PR | Progesterone receptor |
| RBL1 | Retinoblastoma-Like 1 Protein |
| RBL2 | Retinoblastoma-Like 2 Protein |
| RBBP4 | Retinoblastoma Binding 4 Protein |
| RNA | Ribonucleic acid |
| RPM | Revolutions per minute |
| SANT | DNA-chromatin recognition sequence |
| SC | Santa Cruz |
| shRNA | short hairpin RNA |

| | |
|---|---|
| S phase | Synthesis Phase |
| STRING™ | Search Tool for Retrieval of Interacting Genes/Proteins |
| SW13 | Epithelial Carcinoma Cell line |
| TAD | Trans activation Domain |
| NRD | Negative regulation Domain |
| 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 |
| UBA1 | Ubiquitin like modifier activating enzyme 1 |
| UBXN2B | UBX Domain Protein 2B |
| UBXN8 | UBX Domain Protein 8 |
| VCPIP1 | Valosin Containing Protein Interacting Protein 1 |

## VITA

| | |
|---|---|
| 2021 …….…… | B.S. Biology; Minor: Chemistry<br>Prairie View A&M University<br>Prairie View, Texas |
| 2022-2024 …….……... | Graduate Teaching Assistant<br>Department of Biology<br>Texas Southern University<br>Houston, Texas |
| Major Field……. | Biology |

## ACKNOWLEDGEMENTS

I would like to express my gratitude to God for continually providing solutions, even when I felt all was lost, allowing me to graduate. I owe immense thanks to my incredible advisor, Dr. Audrey Player, for her continuous support, as I would not be where I am today without her.

I also want to honor my late uncle, Dr. Richard Idowu, who instilled in me to never be comfortable with anywhere I am in life and to strive to be better every single day, no matter how small. I express my heartfelt gratitude to my dad, Engr. Rasaq, and my mum, Mrs. Toibat, as well as my friends Idris, Mubaraq, Dapo, Dr. Steve, and once again, my advisor, Dr. Player. who were always there when you needed them.

To my committee members Dr. Sodipe, Dr. Shodinmu, and Dr. Phan, thank you so much for the time, your guidance, and helping me lay the foundation for the next steps ahead. As well as my lab partner LaTaija Thirston for her continuous encouragement.

# CHAPTER 1

# INTRODUCTION

Second to skin cancer, breast cancer is the most common type of cancer in women in the United States (33). An average of about 240,000 women will be diagnosed with breast cancer each year, and about 42,000 women will die from the disease (5). For some breast cancer types, the incidence in women of African descent are higher compared to Caucasian women (9). Data suggest this is in part due to socioeconomic factors related to lack of access to healthcare and late-stage diagnoses, but other investigators attribute this to possible genetic factors. There are studies that both support and dispute that racial disparities are related to genetic parameters (9)(41). Aside from breast cancers in women, men are also affected by breast cancer. Every year about 2100 men are diagnoses with breast cancer and about 500 other (previously diagnosed) men will die from the disease each year. The presentation of the disease and molecular characteristics are similar in both men and women.

Breast cancers are characterized based on the molecular signatures defining the disease and the pathological diagnoses. Prior to the 1990's breast cancers were almost exclusively characterized by pathology. After the development of research tools that allowed for molecular characterization of cancer and other diseases, breast cancer was defined by both molecular and pathological diagnoses. The section below will outline the pathological diagnostic types of breast cancers (24) with reference to female breast

cancers, but even though the incidence of male breast cancers is rare, the same diagnostics terms (based on histology) can be applied to male breast cancers.

**Pathological Characterization of Breast Cancers**

**Ductal Carcinoma in Situ (DCIS):** DCIS occurs in approximately 15% of patients. The cells in DCIS form in the milk ducts are carcinoma, but they are in situ and are non-invasive. The presence of DCIS is often considered a risk to develop invasive cancers.

**Lobular Carcinoma *in Situ* (LCIS)**: LCIS is a benign lobular neoplasia that occurs in ~2% of patients. LCIS presents an abnormal cell growth in the milk glands (lobules), and also consider a risk factor for development into invasive cancers.

**Cribriform Carcinoma of the Breast**: Cribriform cancers are rare invasive cancers, which occur at a rate of 0.3-3% of breast cancers. Histologically, the cancers appear normal-like, but also show cribriform-like configurations.

**Inflammatory Breast Cancer (IBC):** IBC occurs in 1-5% of breast cancer patients. IBC is a rare, invasive and often deadly form of breast cancer. Most cancers are present as a lump, but IBC cancers grow in sheets, which makes it difficult to remove and treat.

**Invasive Ductal Carcinoma (IDC):** IDC is the most common kind of breast cancer, occurring at approximately 60-70% of breast cancers in the patient's milk ducts. There are treatment options for cancers but like any cancer, if they are diagnosed in late stages, they can be deadly.

**Invasive Lobular Carcinoma (ILC):** ILC cancers occur in approximately 15% of patients. Cancers are invasive and originate in the lobules of the breast.

**Mucinous Carcinoma of the Breast**: Mucinous Cancer is a rare cancer occurring in ~2% of breast cancers forming in the patient's milk ducts. The cancers are characterized by over-expression of the high molecular weight mucin glycoprotein. They are often referred to as "floating in a pool of mucin".

**Medullary Carcinoma of the Breast:** Medullary carcinomas occur in around 3% of breast cancer patients. the cancers are shiny in appearance, with brain-medulla-like features. The cancers are basically triple negative breast cancers.

**Papillary Carcinoma of the Breast:** Papillary cancers are rare invasive cancers that occur in approximately 1% of breast cancer patients. The cancers appear like small finger-like projections or papules and well-defined borders.

**Paget's Disease of the Nipple**: Paget's disease occurs in ~ 1-4% of breast cancers. Paget's disease looks like dermatitis with eczema-like features in regions around the nipple are flaky. Cancers are frequently misdiagnosed as eczema.

**Phyllodes Tumors of the Breast:** Phyllodes cancers occur in less than 1% of breast cancer patients. They appear as leaf-like structures and are thought to originate in stromal breast cells.

**Tubular Carcinoma of the Breast (TBC):** Tubular breast cancer is invasive ductal cancers that affect approximately 2% of breast cancer patients. The cancer appears like tube-like structures.

**Molecular Characterization of Breast Cancers:**

Although we often refer to breast cancer as though it were a single disease, breast cancers are extremely heterogenous as noted by the pathological descriptions summarized above. Based on molecular studies there are at least 5 molecular subtypes of breast cancer

and within the individual subtypes there are subcategory distinctions (40). The subtypes are characterized based on molecular signatures and the ability to cluster based on their expression profiles. The more similar the transcriptomes, the closer samples 'cluster' following analyses (further explained below). The five breast cancer subtypes are classified as Luminal A, Luminal B, HER2-like, Triple negative breast cancers (TNBC) /Basal-like and Normal-like (Figure 1).

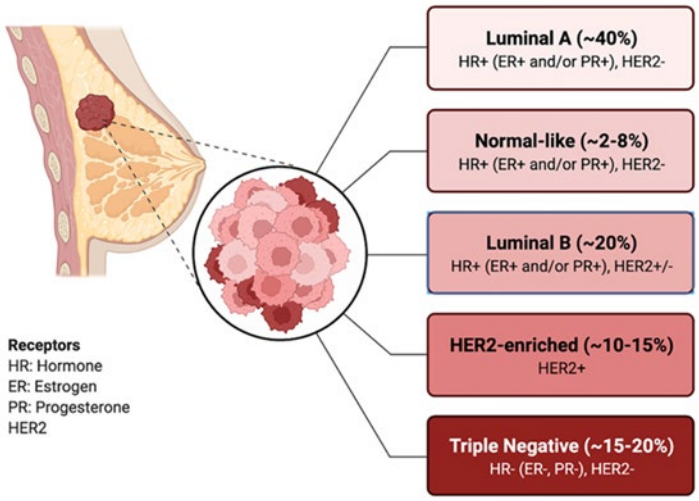


**Figure 1: Molecular Classification of the 5 Breast Cancer Subtypes (2)**

(a) Luminal A occurs in about 40% of women diagnosed and are positive for estrogen and progesterone receptor genes and HER2 negative.
(b) Luminal B occurs in about 20% of women diagnosed and positive for estrogen and progesterone and +/- HER2 expression.
(c) HER2 over-expressed occurs in about 10-15% of women diagnosed and are positive for over-expression of HER2 gene.
(d) Triple negative breast occurs in about 15% of women diagnosed and are negative for estrogen receptor, progesterone receptor and HER2. and
(e) Normal-like occurs in about 2-8% of women and positive for estrogen receptor and progesterone receptor and HER2 negative.

The breast cancers in Figure 1 are ranked top to bottom, from good-to-poor prognoses. Luminal A and Normal-like breast cancers have more favorable prognoses

while TNBCs patients have the poorest disease prognoses. Data show that the Luminal breast cancers likely originate from progenitor cells lining the lumen constituting the inner layer of the ducts, and the TNBC/Basal-like cancers likely originate from progenitor cells on the underside of the luminal cells nearer the basement membrane of the ducts (3) (Figure 2; 2)

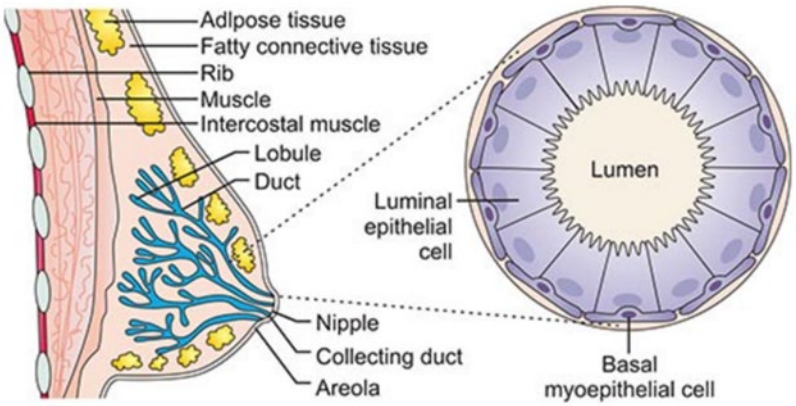


**Figure 2: Diagram of Breast Tissue to Demonstrate the Location of the Cancer Progenitor Cells (14)**

## Gene Expression Microarray Analyses Used for Classification of Breast Cancers

The molecular characterization of the breast cancers described above was accomplished using DNA microarrays. The use of DNA microarrays has been instrumental in characterizing cancers and all types of diseases. As a summary of the technique, for gene transcriptome expression analyses (a) mRNA is reverse transcribed to first strand (cDNA) and then second strand DNA (b) antisense RNA generated from this material is biotin-labeled, fragmented and utilized to hybridize to a high through-put DNA microarray gene-chip. The microarray gene-chip contains sense material corresponding to the transcriptome (i.e., all of the transcripts) sequences generated from a particular cell type/sample. That

way the transcriptomes for particular cancer types can be compared. If after analyses the transcriptome profiles 'cluster together' (i.e., look similar with respect to gene expression patterns), one can assume the sample-types are similar or demonstrate some relationship based on their signaling pathways. See the example profile in Figure 3. The same approach is used to characterize all diseases. Analyses of breast cancers and every other disease can be done using DNA microarrays. The example in Figure 3 is a partial snapshot of what the microarray data looks like following computer analyses. In this example there are about 150 patients examined (across the top), displaying the gene expression pattern of about 32 genes (as columns). Genes with green colors represent over-expression and genes with red colors indicate under-expression of the gene in the samples. The actual DNA microarray gene-chip contains the sense sequences corresponding to 'all of the known gene transcripts' deposited at the National Center for Biotechnology Information (NCBI), not just 32 gene transcripts as used in the example. The final dataset can contain anywhere from 56,000 to 200,000 data points. These data can only be processed, and the data interpreted with the assistance of computer analyses. Note that different sample types will result in distinct patterns of expression (i.e., red/green; decreased/increased gene expression levels). And the genes defining this condition or sample type can easily be retrieved following data cluster-analyses. The computer program will automatically cluster patients and genes with similar patterns.

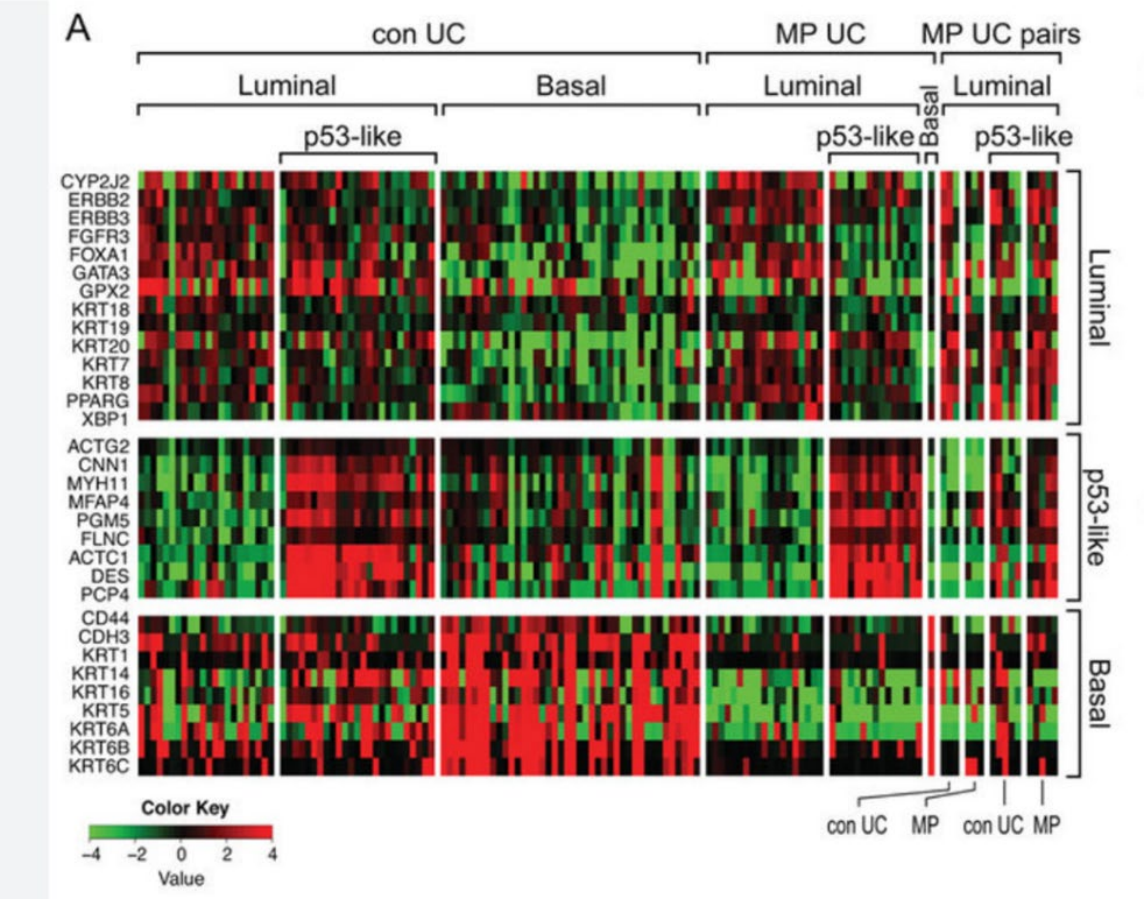


**Figure 3: Example of the Microarray Cluster Analyses Method Utilized to Identify Breast Subtypes and Identify Therapeutic Biomarkers for Patient Care (7).**

***Introduction to the Research Approach for the Current Study***

Our laboratory focuses on studies of TNBC. TNBC are negative for the three genes estrogen receptor (ER), progesterone receptor (PR), and HER2 also called ERBB2 (-v-erb-b2 erythroblastic leukemia viral oncogene homolog 2). TNBC are invasive cancers with poor 5-year survival and are more likely to metastasize than some other breast cancer subtypes. Compared to receptor positive breast cancers, there are not many effective drug therapies for TNBC. There are effective drug therapies for breast cancers that are positive

for the receptors, partly contributing to their improved survival. Because of this, many investigators, including our laboratory, are further characterizing the TNBC with the goal of eventually identifying reliable biomarkers that can be examined for their possible therapeutic potential. For many of our gene screening protocols we utilize non-tumor triple negative cell lines compared to tumor TNBC cell lines.

The microarray gene expression procedure described in the section above can be used for a myriad of applications. It is the main platform utilized in our laboratory for characterizing the transcriptome of TNBC and subsequent gene discovery. In an earlier study we identified MYBL1 transcripts as over-expressed in TNBC samples compared to a non-tumor triple negative breast sample and some luminal breast cancer cell lines, and clinical patient samples using microarrays (1)(25). MYBL1 gene is a strong transcriptional activator which is involved in key processes associated with cancer progression, so a decision was made to further characterize the gene in TNBC with the expectation the gene contributes to the cancer genotype. Our approach was to knockdown MYBL1 gene, and because it regulates other genes, we looked for genes that were either directly or indirectly affected by the knockdown process (25). When MYBL1 was knocked down in a TNBC cell line VCPIP1 gene was also downregulated. Both genes are located at chromosome 8q13.1 locus, are involved in mitotic signaling and show consistent co-expression pattern in many of the same TNBC cell lines and patient samples. The genes demonstrate a similar pattern of mRNA and protein expression in cell lines. When clinical TNBC patient samples were analyzed, MYBL1 and VCPIP1 had alterations (i.e., including different types of mutations including homozygous and heterozygous deletions, missense mutations and RNA and protein over-expression) in the same patient samples with complete concordance.

The current study demonstrates bioinformatic data supporting the relationship between MYBL1 and VCPIP1 in different TNBC samples and presents experimental data validating the co-expression pattern of the genes in TNBC cell line preparations. The VCPIP1 gene is specifically involved in Ubiquitin-dependent Protein Degradation at the Endoplasmic Reticulum signaling pathway which also includes UBXN2B and several other genes. We examine gene expression of these genes as well in TNBC. This data are unclear but the co-ordinate expression between MYBL1 and VCPIP1 appears far more consistent.

# CHAPTER 2

# LITERARY REVIEW

## *Background Information Concerning the MYB Family of Genes*

The human MYB family of genes includes 3 family members, c-MYB, MYBL1 and MYBL2. The genes were discovered based on their similarity and cellular homolog to the v-myb oncogene carried by avian leukemia viruses avian myeloblastosis virus and E26 (9)(38). The c-MYB gene was discovered first following screening of sequence similarities of the bird cDNA library, while MYBL1 and MYBL2 were identified based on their extensive sequence similarity to c-MYB (1)(7). The three genes are expressed in different cell types, where they have been found to have both similar and different regulatory functions. The c-MYB is overexpressed in immature hemopoietic cells and data suggest it is important for maintaining the proliferative state of the progenitor cells (18)(37). In addition, c-MYB and MYBL2 and both expressed in a wide range of cell types, while MYBL1 is found in testis, monocytes and cerebrospinal fluids (34). Ruston et al (28)(29) demonstrated that based on their programmed signatures and microenvironment, the same gene could perform a different function in different cell types. Considered together, these data suggest a major role of the MYB family of genes in normal cells. More recently genes have been studied and shown to be involved in a variety of different cancers.

The human MYB family of genes share approximately 90% similarity at their 5'prime DNA binding domain (DBD) regions (17)(43). 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 4), 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 can interact with and regulate some of the same genes. c-MYB, MYBL1 and MYBL2 also share homology in a transactivation domain and the negative regulatory domain. There is less homology in these regions which could account for the differential expression of the genes in different types of cells, via interaction with different transcription factors, co-activators. Differences related to the negative regulator region likely account for the differential expression in different types of cells as well. Utilizing recombinant adenovirus Rushton, Ness and Davis (23)(28)(29) showed that c-MYB, MYBL1 and MYBL2 could regulate some of the same genes and each could regulate a unique set of genes, validating the similarity and specificity of their gene functions.

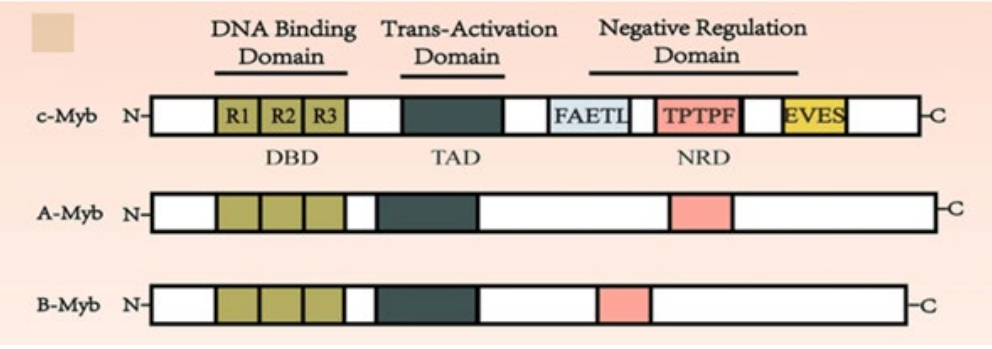


**Figure 4: Diagram Demonstrating the Similarity between MYB Gene Family (15)**

***Review of Processes Related to Gene Expression of MYBL1- Involvement in Cell Cycle***

The current project focuses on characterizing the MYBL1 gene. MYBL1 is a protooncogene with substantial data validating its involvement in cell cycle regulation and proliferation, differentiation, and apoptosis. MYBL1 is not as well characterized as c-MYB and MYBL2 and does not appear expressed as many cell types. Other studies show MYBL1 is associated with early-stages spermatozoa progenitor cells. And other studies show the gene is a master regulator of meiosis in spermatocytes (4). Related to the cell cycle, Ziebold and Klempnauer (45) showed that the highest levels of MYBL1 are observed in late G1 and early S-phases of the cell cycle. Their data suggest the gene serves as a cell cycle sensor regulated by signaling processes in the cell cycle. Their hypothesis was that cyclin dependent kinases might act to phosphorylate MYBL1 via the carboxy-terminal regulatory region of the gene. Other data show MYBL1 gene is a component of the DREAM complex, which functions to regulate cell cycle-dependent genes (30). DREAM stands for the **d**imerization partner **R**B-Like, **E**2F **a**nd **m**ulti-vulval class B (30). The DREAM complex consists of the E2F4, E2F5, LIN9, LIN37, LIN52, LIN54, MYBL1, MYBL2, RBL1, RBL2, RBBP4, TFDP1 and TFDP2 proteins (30). The genes interact at various combinations to regulate the cell cycle. Consistent with document studies, in our studies we frequently find E2F4, LIN37 and TFDP2, MYBL2 and coordinately differentially expressed with MYBL1 in TNBC.

***MYBL1 and Role in Tumor in Various Organs and Cell Types***

Less is known about the MYBL1 gene compared to c-MYB and MYBL2. Until a few years ago c-MYB was considered a promising target for treatment of luminal breast cancers. Small molecules were being designed to target luminal cancers many of which

express high levels of the genes. Modifications to the carboxy-terminal region of the protein are implicated in tumorigenesis. As for MYBL1, the gene is involved in mechanisms shown to be key to the progression of cancers, including involvement in cell cycle processes and differentiation which are hallmarks of tumor progressions. We suspect the gene is involved in early tumor pathogenesis. MYBL1 gene is over-expressed in low grade glioma (32) and genomic alterations in breast adenoid cystic carcinomas which are a rare triple negative breast cancer (20). In the adenoid cystic carcinomas, gene fusions between MYBL1 and ACTN1 or MYBL1 and NFIB genes are suspected of leading to MYBL1 overexpression. And in salivary gland carcinomas gene fusions between MYBL1 and NFIB are thought to contribute to the cancers (11). Recent studies suggest MYBL1 can serve as an immunotherapy biomarker in clear cell renal cell carcinoma (6). These data are consistent with our studies in that MYBL1 was initially identified as differentially expressed in TNBC over-expressing the IL32 pro-inflammatory cytokine. MYBL1 gene is a strong transcriptional activator. A recent study shows MYBL1 binds to the ANGPT2 promoter region and activates tumor angiogenesis in hepatocellular carcinoma (44). Other than dysregulations related to gene fusions, the precise role of the gene in many of the tumorigenic processes has not been determined.

***Background Describing Our Identification of MYBL1 Expression in TNBC***

We identified MYBL1 as one of six genes over-expressed in TNBC cell lines and patient samples that over-express the IL32 cytokine gene (25). Two of the genes identified by our analyses were previously identified as biomarkers for TNBC and 4 genes (including MYBL1) were identified as novel by our laboratory. The analyses were based on use of an unsupervised high-throughput analyses of thousands of patient samples and hundreds of

cell lines including non-tumor triple negative, the five different breast cancer subtypes and various subcategories of TNBCs. All of the samples utilized in that study were previously processed (by other investigators) using the Affymetrix DNA microarray. The samples were obtained primarily from the NCBI Gene Expression Omnibus (GEO). Data show MYBL1 gene consistently over-expressed in a subset of TNBC cell lines and patient samples (Player et al 2024 in press). Because MYBL1 is a strong transcriptional activator and regulates a host of genes, is consistently differentially over-expressed in a subcategory of TNBC cell lines and patient samples, and regulates other key genes important for cancers, we made the decision to further characterize the gene. That is the focus of the current project.

***Experimental Approach to Study MYBL1 Gene in TNBC***

For the past few years our laboratory has focused on attempting to better understand the role that MYBL1 might play in TNBC. The gene is a strong transcriptional activator previously shown to be involved in key cancer signaling processes. So, the expectation is that the gene also plays a role in TNBC pathogenesis. Our goal is to characterize the gene in breast cancers. We understand that there are many questions to address, but because the gene is a strong transcriptional activator and likely contributes to the genotype and signaling events in TNBC, we sought to define genes that associate with MYBL1 in these cancers. Our approach was to utilize the lentiviral shRNA targeted knockdown procedure to silence the MYBL1 gene and define genes affected by the knockdown process. As a summary, we (a) silenced MYBL1 in TNBC cells (b) compared the transcriptome to shRNA control compared to the shRNA MYBL1 knock down and identified differentially expressed genes using the free-online TAC4 software (c) we then utilized CHipXpress to

determine which transcription factors were enriched during the knockdown process. As expected, genes regulated by MYBL1, MYBL2 and MYC transcription factors were enriched in the differentially expressed gene dataset, validating the role of these transcription factors in the knockdown process. We then (d) examined the dataset using the WebGestalt (42) and Pantherdb Gene-set enrichment (21) programs to identify enriched signaling pathway, compartment and other enrichment gene-sets associated with the candidate dataset. As expected, cell cycle signaling and proliferation markers, cancer biomarkers, differentiation and cell death signatures were affected by MYBL1 silencing in the TNBC cells. Cell cycle signaling and proliferation genes were substantially enriched compared to other signaling mechanisms and biomarkers. As validation of our approach and analyses, many of the genes affected by the process were previously defined, but a substantial number of novel candidate genes were also identified. Previous studies suggest (but did not directly show) that MYBL1, MYBL2 and MYC genes are coordinately regulated. Our data clearly demonstrate that when MYBL1 is knocked down, MYBL2 and MYC genes are downregulated, validating previous observations suggesting a relationship between the 3 genes. The current study will focus on MYBL1 gene and two of the novel genes identified.

***Experimental Outline for the Current Study***

Following analyses of the knockdown dataset we find that MYBL1 and VCPIP1 appear to be coordinately expressed in some of the same TNBC cell lines and patient samples. Both are downregulated with MYBL1 knockdown, and both are dysregulated in the exact same patient samples. This data was first observed after analysis of a small number of clinical patient samples. The current study represents analyses of additional

TNBC cell lines, additional patient samples and analysis of different kinds of alterations including deletions, expression, missense mutations et al in TNBC patient samples. Both MYBL1 and VCPIP1 genes are located at the chromosome 8q13.1 locus. This could be coincidental, but it might also represent a deletion 'hot spot'. In addition, bioinformatic analyses show MYBL1 binding site in the VCPIP1 promoter region suggesting MYBL1 can regulate expression of the gene. We also consider the possibility that both genes are involved in the same signaling pathway. The MYBL1 has been documented to be involved in cell cycle signaling, proliferation and drive genes at the mitosis-to-meiosis transition (19). While VCPIP1 gene is involved in ubiquitination processes and a key gene involved in the assembly of Golgi and Endoplasmic reticular during p97/p47-mediated membrane fusion events related to mitosis. Our hypothesis is that at some juncture the MYBL1 and VCPIP1 genes are involved in the same signaling pathways, be it mitotic and/or cell proliferation. A VCPIP1 signaling pathway has been defined (38). As part of the current study we perform expression analyses of several of these genes to determine their dysregulation in TNBC. One of these genes is the UBXN2B gene. In addition, bioinformatic analyses show MYBL1 binding site in the VCPIP1 promoter region suggesting MYBL1 can regulate expression of the gene. Considered together these data support explanations for co-expression of MYBL1 and VCPIP1 genes.

**Summary Description of MYBL1 Gene**

We described the MYBL1 gene in the sections above. MYBL1 was the target of the knockdown process in the MDA-MB-231 TNBC cell line. The gene is located at position 8q13.1 chromosomal locus. And it is a strong transcription activator and

protooncogene involved in cell differentiation, proliferation and apoptosis related signaling events.

**Summary description of VCIP1 gene**

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 named 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 (35). VCPIP1 is also involved in ATM or ATR DNA repair events related to targeting DNA-protein cross-links and deubiquitinating events.

**Summary Description of UBXN2B Gene**

The UBXN2B gene is located at chromosome region 8q12.1 in a region near MYBL1 and VCPIP1 genes. The gene was not affected by the knockdown of MYBL1 gene suggesting neither direct of indirect regulation by MYBL1. UBXN2B is required for Golgi and endoplasmic reticulum development. The gene is also involved in maintenance of the Golgi and endoplasmic reticulum during mitosis. There is evidence that the VCPIP1 and UBXN2B genes are part of a scaffolding complex formed during mitosis (19).

# CHAPTER 3

# DESIGN OF THE STUDY

## Cell Lines and Patient Datasets

Cell lines utilized in this study were purchased from the American Type Culture Collection (ATCC®) atcc.org. The MCF10A cell line was a non-tumor triple negative cell line. The Hs 578T and MDA-MB-231 cell lines were triple negative breast cancer cell lines. For screening purposes, the MCF7 luminal cell line was utilized. The cell lines were maintained according to the recommendation of ATCC.org. The cells were grown in T75 tissue culture dishes and fed twice weekly with Dulbecco's Modified Eagle Minimum essential media (DMEM) supplemented with 1% penicillin and 10% serum and incubated in a $37^{O}C$ incubator with an additional 5% $CO_2$. The cells were grown to ~80% confluent and then typsinized using a 0.25% trypsin solution for sub-culturing or until utilized for experiments.

The patient datasets were available on the cbioPortal.org online database (6). Approximately 13,800 pan cancer patient samples and 2700 breast cancer patient samples were interrogated based on gene expression analyses using cBioPortal analyses resources.

## 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 (Thermo-Fisher Scientific, *Norristown, PA*). As summary

of the protocol, 200ul of chloroform was added to the Trizol/cell preparation. The suspension was mixed followed by centrifuged at 9,000 revolutions per minute for separation of the total RNA. The top RNA layer was placed into a clean tube with 500 µL of 95% ethanol and stored at -20°C for 30 minutes. The RNA preparation was collected by centrifugation for 30 minutes at 9,000 RPM. The RNA pellet was desiccated, resuspended in ~20 µL of clean water. 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 the 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. The components and volumes utilized for cDNA conversion were performed as suggested by BioRad company. As summary, 20 µL mixture containing 5x iScript buffer, reverse transcriptase, random hexamers, total RNA (10ul), 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 85°C for 1 minute. The mixture was chilled and 80 µL of water was added. This mixture was considered cDNA and was stored at –20°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$^{TM}$ (27) program (http://bioinfo.ut.ee/primer3-0.4.0/ ) was used to generate primers to detect the target genes. The sequences for only 4 final candidate genes: GAPDH (control), VCPIP1, UBXN2B and MYBL1 are given in this document (see below). The other genes were processed as above but their experimental results were not reproducible, so the genes were not considered reliable, thus not considered as candidate genes. Affymetrix NetAffx was used to obtain the nucleotide sequences of our target genes (http://www.affymetrix.com/estore/analysis/index.affx) (you must be a registered used to access the Affymetrix / Thermo-Fisher Scientific webpage where the probe-set sequences are deposited). 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/) using the *In-silico PCR* analyses program. The primer-sets were synthesized by IDTDNA.com (Coralville Iowa, USA), manufactured, then shipped to Texas Southern University within 24-48 hours of purchase. Primer sets for the target genes analyzed in this study were generated and used for analysis of breast cell line samples. Primer sequences for the genes demonstrating reproducible results in this study are given in Table 1.

**Table 1: PCR Primer Sequences**

| | LEFT | RIGHT | AMPLICON SIZE |
|---|---|---|---|
| **MYBL1** | TGGATAAGTCTGGGCTTATTGG | CCATGCAAGTATGGCTGCTA | 210BP |
| **VCPIP1** | CAGGCAGCTTGATCCTGATT | CTCCCAGTGCATCTGCTACA | 272BP |
| **UBXN2B** | TTCCTGTCCATGCAGTAGCA | CAACTCATTTGCCTAGAAATCCA | 252BP |

**Polymerase Chain Reaction (PCR)**

For the PCR reaction, approximately 10 nanograms (ng) of cDNA was mixed with 10uL of Amplitaq Gold$^{TM}$ (Life Technologies, Waltham MA, USA), 10uM each of the forward and reverse gene primers, and water up 20uL. PCR conditions were carried out at 95°C 5minutes for one cycle, then 30 cycles at 95°C for 30seconds, 58°C for 90seconds and 70ºC for 90seconds, followed by a final extension of 70ºC for 5minutes. PCR products were separated into a 1.0% agarose gel.

**Gel Electrophoresis**

Ten microliters of PCR product were then mixed with two microliters of 5X sample buffer then analyzed on a 1.0% agarose gel (in Tris Acetate EDTA buffer with Ethidium bromide). The agarose gel was then visualized and photographed using the LiCor Biosciences, Odyssey Bio-Imager (Lincoln, NE, USA).

**Western Blotting**

The western blot procedure was performed as outlined in a previous document (26). 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). VCPIP1 anti-mouse (sc-515281) antibody was purchased from Santa Cruz Biotechnology (Santa Cruz CA, USA) and 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, USA). Western blotting filters were developed using the Clarity Western ECL substrate (Bio-Rad,

Hercules, CA, USA) using the LI-COR digital imaging system (LI-COR Biotechnology, Lincoln, NE).

**Densitometer**

The LICOR densitometer software was used to quantitate the gene expression levels compared to control gene expression levels.

**Data Analysis Using cBioPortal Open-Source Database**

The cBioPortal.org website (6) was used to analyze our genes against pan and breast cancer 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 Onco-Print analyses allows for mutation and gene expression analyses of our genes in the various patient datasets. Co-expression of our genes was determined using this approach.

The Search Tool for the Retrieval of Interacting Genes/Proteins (STRING$^{TM}$) was utilized to determine protein: protein interactions (31) between the candidate genes. STRING$^{TM}$ analysis is a database tool for known and predicted protein-protein interactions. Protein Atlas (22, 36) was utilized to demonstrate the expression of MYBL1, VCPIP1 and UBXN2b in pan-cancers.

# CHAPTER 4
# RESULTS AND DISCUSSION

## *Summary of Experiments Previously Performed but Critical to Understanding This Project*

Data by other investigators demonstrate the MYBL1 gene dysregulated and over-expressed in various types of cancer. Consistent with these data, we find MYBL1 over-expressed in TNBC, and bioinformatic analyses of patient sample datasets demonstrate a variety of different genomic alterations (i.e., over-expression, amplifications, deletions, missense mutation etc.) in TNBC clinical samples. Since our initial observations, our goal has been to further characterize the MYBL1 in TNBC. The current project is an extension of experiments directed towards this goal.

Because MYBL1 gene is a strong transcriptional activator involved in key events associated with regulation of cell proliferation, we knocked down the gene and identified genes either directly or indirectly affected by targeting MYBL1. The VCPIP1 gene was identified. Table 2 summarized the results of the microarray gene-chip analyses. This data were generated in a previous study and are critical to this current study. (a) MDA-MB-231 cells were grown on a dish and exposed to either a lentiviral shRNA control sequence (generated by Origene.org) or lentiviral shRNA MYBL1 targeting sequence for 72 hours at 37°C. (b) the RNA was extracted, purified and processed for hybridization to the Affymetrix Clariom gene-ship microarray (this was performed by University of Texas, Southwest). The datasets were made available to Texas Southern University (TSU) and we analyzed the data and identified genes differentially expressed, in other

words, genes affected by the knockdown process. The table contains log2 values corresponding to transcript levels. In addition to the transcript levels, the values and fold-change or degree of knockdown, are generated by the TAC4 data analyses software. As an explanation of the table, MYBL1 transcript levels in the control shRNA chip are 12.2 (or ~4000 MYBL1 transcripts), and in the knockdown sample MYBL1 are 10.2 (or ~1000 transcripts). Hence, when MYBL1 is target by the lentivirus shRNA-MYBL1 sequence, transcript levels decrease; the MYBL1 gene is 'knocked down' by 4 X times. Thousands of genes are affected by this process. We found a substantial number of genes were localized to the chromosome region 8q loci (25). VCPIP1 located in the same locus as MYBL1, was also affected by the process. VCPIP1 was knocked down from 10.2 (~1000 transcripts) to 9 (~500 transcripts), by a factor of 2X times. Data show the gene UBXN2B, which interacts with VCPIP1 was not on the list, but we chose to examine its expression because both genes are involved in the same Golgi mitotic deubiquitylation signaling process; we will describe this pathway in the next chapter. Related to UBXN2B, processes other than MYBL1 are likely involved in its regulation.

**Table 2: Data Generated from the Original MYBL1 Knockdown in MDA-MB-231 Cells**

The data were collected from the microarray gene-chip. shRNA control transcript levels (a) compared to shRNA MYBL1 targeted transcript levels (b). Fold change in knockdown is also given (c), along with the chromosomal loci of the gene (d). As example, MYBL1 transcript levels decreased from 12.2 log2 to 10.2 log2 transcript levels; antilog values are ~4000 transcripts to 1000 transcripts after knockdown).

| | a. | b. | c. | d. |
|---|---|---|---|---|
| | shRNA scramble control (RNA levels) LOG 2 VALUES | mybl1 knock down (RNA levels ) LOG2 VALUES | FOLD CHANGE (up or down regulated after MYBL1 KD) | CHROMOSOME |
| **MYBL1** | 12.2 | 10.2 | 4X KD | **8q13.1** |
| **VCP1P1** | 10.1 | 9 | 2.1X down reg | **8q13.1** |
| UBXN2B (p37) | NOT ON LIST | | | 8q12.1 |

Our microarray and knockdown data suggest a strong, reliable relationship between MYBL1 and VCPIP1. It is possible the genes are co-regulated. As noted in the previous section, MYBL1 and VCPIP1 are located at the same chromosomal locus, 8q13.1 and UBXN2B is located at the 8q12.1 locus (Figure 5). The arrows point to the locations for MYBL1, VCPIP1 and UBXN2B. The diagram was extracted from GeneCards.org (12)(13). The red line drawn through the chromosome was generated by GeneCards program. The 8q loci has been identified as a region with a high incidence of mutations in TNBC (13). It could be that MYBL1 and VCPIP1 contribute to this mutation frequency.

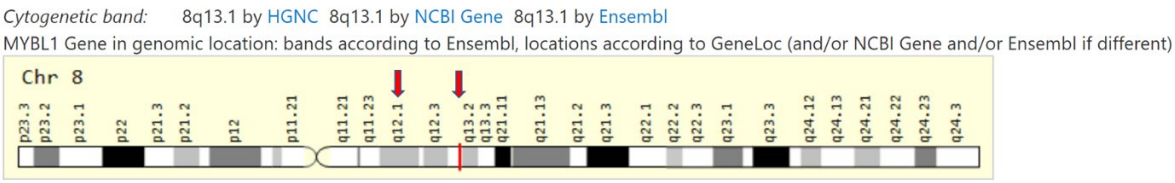


**Figure 5: Chromosomal Region Showing MYBL1 and VCPIP1 Locus and UBXN2B**

MYBL1 and VCPIP1 are both located at 8q13.1 locus, and UBXN2B is located at 8q12.1 locus. The red arrow points to gene locus on chromosome (12)(13).

VCPIP1 gene was first considered for our studies because it was affected by the knockdown process (following analyses of the TAC4 dataset), the gene was located at the same locus as MYBL1 and preliminary analyses of patient sample data show co-expression of MYBL1 and VCPIP1 in some of the same clinical patient samples. Analyses of Reference Sequence high throughput analyses (obtained from ProteinAtlas.org (22)(36) of pan cancers show that when different types of cancers are examined for all the RNA species, MYBL1 and VCPIP1 genes show a similar pattern of expression in breast cancers

and lower RNA Ref Seq levels in other types of cancers. These data validate a pattern of co-expression of MYBL1 and VCPIP1, and to some degree UBXN2B in breast cancers compared to most other cancers. Please note that the diagram was extracted from ProteinAtlas.org website and breast cancers include *all subtypes of breast cancer not only TNBC*. There also appeared to be similarities in lung cancer RNA Ref Seqs.

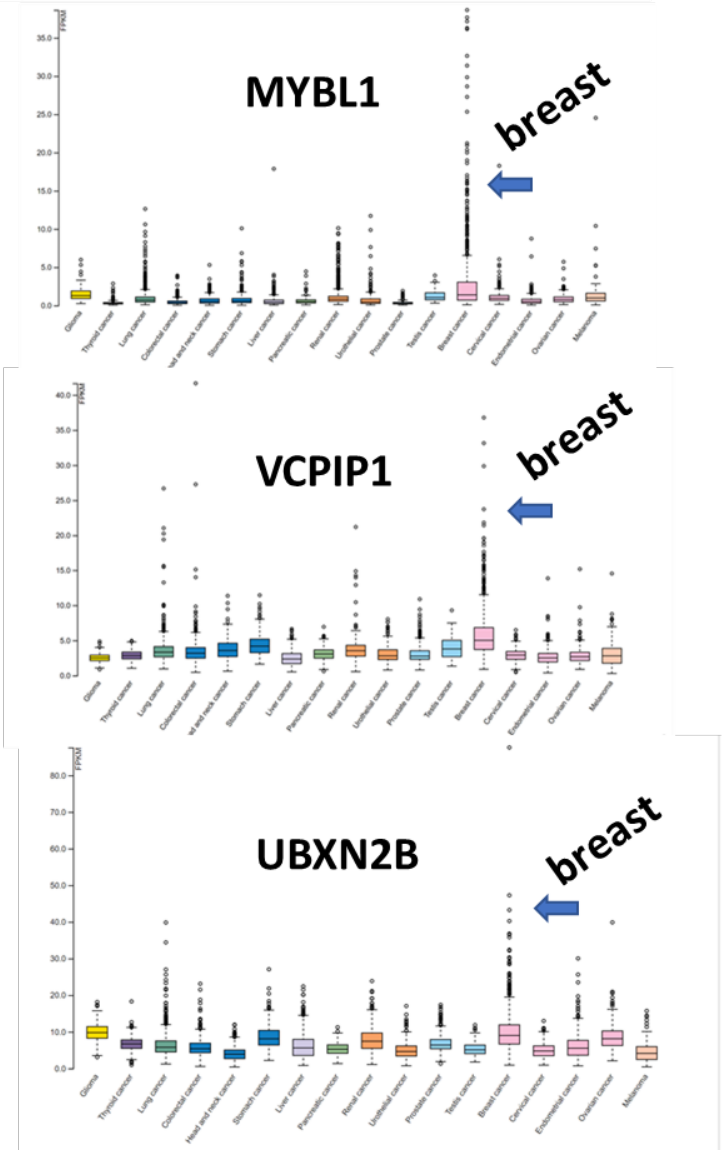


**Figure 6: ProteinAtlas RNA Reference Sequence Analyses of Pan Cancer RNA Levels (22, 36)**

There is substantial experimental data validating interactions between VCPIP1 and UBXN2B. Both genes are involved in mitotic events associated with ubiquitin signaling events involving the Golgi Apparatus and Endoplasmic Reticulum (38) (Figure 7a) during interphase of the cell cycle.

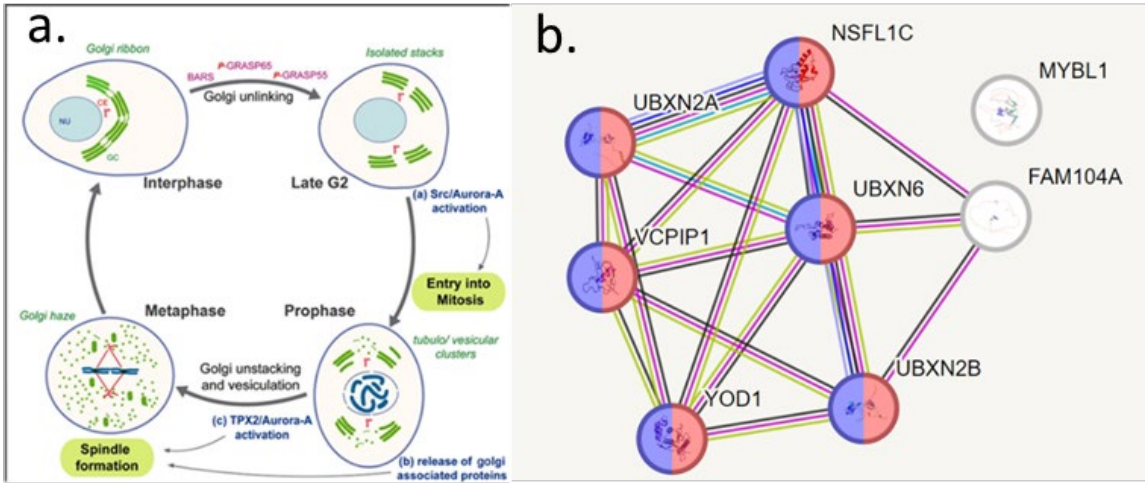


**Figure 7: Pathway Summary and STRING Protein Interaction Figures to Demonstrate the Relationships between Candidate Genes (31)**

(a) Summary of the Golgi/Endoplasmic reticulum involvement in mitosis. (b) STRING protein: protein interaction linkage to show the relationship between the candidate genes. Note – there is no known relationship with MYBL1 (31).

The false discovery rate for these events are significant. For the associations with ubiquitin events associated with the Golgi Apparatus and Endoplasmic Reticulum the false discovery rates is $4 \times 10^{-11}$ (red in STRING diagram) and for containing the Ubiquitin-like domain, the false discovery rate is $1.5 \times 10^{-7}$ (blue in STRING diagram) (Figure 7b). The

STRING program demonstrates protein: protein interactions or associations as determined experimentally, by PubMed citations or based on predicted interactions. Note – there is no known relationship with MYBL1. The more links (i.e., lines between the proteins), the more associations or biological relationships between the different proteins. VCPIP1 and UBXN2B are involved in Golgi and endoplasmic reticulum maintenance during interphase and in reassembly of the organelles at the end of mitosis. Experimental data show the genes along with a cluster of other genes form a protein: protein complex during mitotic signaling. Several of these genes are included in the section below.

We examined clinical patient sample datasets retrieved from the cBioPortal.org (6) online patient resources. The cBioPortal includes hundreds of thousands of cancer patient data points related to nearly every cancer. In Figure 8 we examined MYBL1, VCPIP1 and UBXN2B in approximately 13,000 patients. Each rectangle represents a different patient. The long red area represents patients with positive for MYBL1 gene alterations. Notice that the length of the line for MYBL1 and VCPIP1 are the same length and color, demonstrating both genes are expressed in the same patients and both genes are amplified (as noted by the red color). A similar, but not exact profile is observed for UBXN2B. In summary, MYBL1 and VCPIP1 gene amplifications are detected in the exact same patients, clearly suggesting co-expression of both genes in particular patients. Near concordance is observed for UBXN2B. The other genes, VCP, UBA, NSFL1C and UBXN8 are not dysregulated in the same patient samples as noted by their different pattern. There are substantial deletion mutations in the UBA1 gene in pan cancer patient samples (as determined by the blue region). We examined these genes because they are key to the Golgi Apparatus and Endoplasmic Reticulum mitotic signaling pathway like VCPIP1 and

UBXN2B genes. These genes were not affected by the knockdown process but considered for analyses to examine the possibility that MYBL1 was associated with this pathway. MYBL1 does not appear to be. The data continues to validate a strong relationship between MYBL1 and VCPIP1, and to a lesser degree UBXN2B.

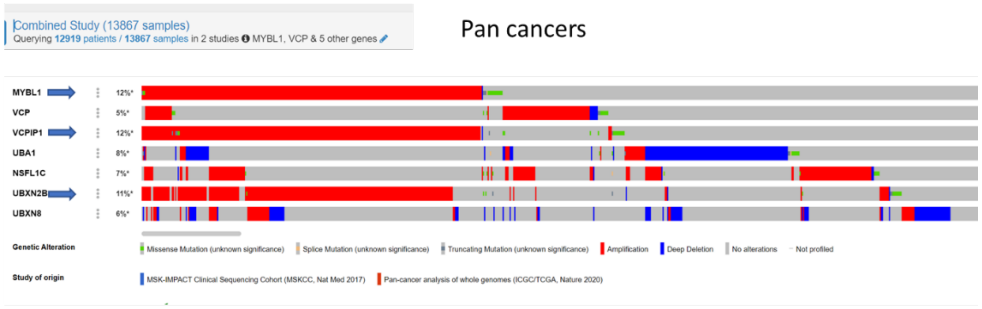


**Figure 8: Analysis of the Gene Alterations in Select Genes in Different (pan) Cancers**

MYBL1, VCPIP1, UBXN2B and VCP, UBA1, UBXN8 and NSFL1C. Each rectangle represents a different patient, approximately 13,000 in these analyses. The red color represents patients with gene amplifications. The blue color represents patients with deletions for a particular gene. Note that MYBL1 and VCPIP1 are dysregulated in the exact same patients, and to a lesser degree UBXN2B (6).

Analyses of breast cancer patients are summarized in Figure 9. Approximately 2700 breast cancer patients were examined. Note also that MYBL1 and VCPIP1 show the same pattern of expression in the same patient samples, as demonstrated by the exact same profile as demonstrated by the red line. UBXN2B was present in nearly, but not the same patients. VCP, UBA1, NSFL1C and UBXN8 are not dysregulated in patients demonstrating MYBL1 and VCPIP1 alterations; these genes show fewer alterations and

UBXN8 shows evidence of deletions in some patients as noted by the blue colored rectangles.

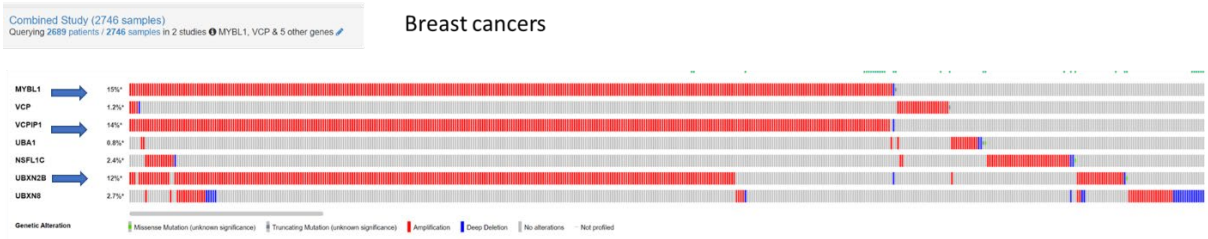


**Figure 9: Analysis of the gene alterations in select genes in breast cancers.**

MYBL1, VCPIP1, UBXN2B and VCP, UBA1, UBXN8 and NSFL1C genes analyzed. Each rectangle represents a different patient, approximately 2700 are analyzed. The red color represents patients with gene amplifications. The blue color represents patients with deletions for a particular gene (UBXN8). Note that MYBL1 and VCPIP1 are dysregulated in the exact same patients, and to a lesser degree UBXN2B (6).

The GeneHancer database contains predicted interactions of enhancer and promoter interactions for particular genes. Because of the close relationship between MYBL1 and VCPIP1 gene we speculated that MYBL1 transcription factor *can* directly bind to and affect analyses of the VCPIP1 gene. So we searched GeneCards and found that MYBL1 recognition sequences and predicted assessory protein bnding sites are found in the VCPIP1 promoter. Based on GenHancer predictions, MYBL1 and possible accessary proteins bind with high confidence to the VCPIP1 promoter sequences (Figure 10).

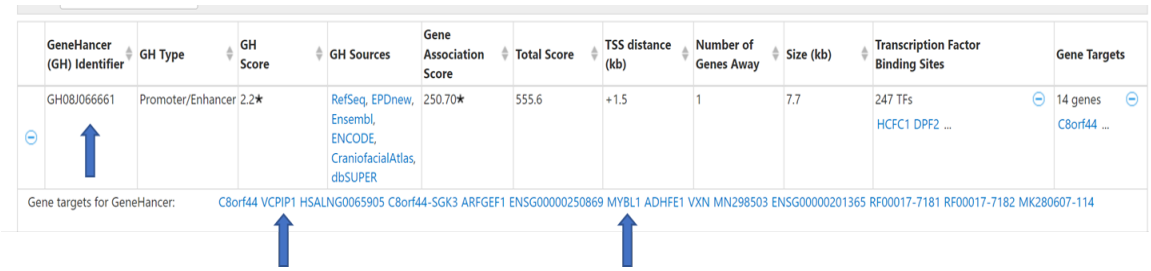


**Figure 10: GeneHancer Analyses of the VCPIP1 Promoter Region for Predicted Interactions**

Arrows point to the GeneHancer Identifier and proteins that bind to VCPIP1 promoter. VCPIP1 and MYBL1 bind with high confidence scores to the VCPIP1 promoter region (13).

We also performed experimental analysis of the genes suspected to be associated with MYBL1 gene. Our microarray data show that MYBL1 and VCPIP1 genes show a similar pattern of gene expression. MYBL1 is over-expressed in tumors and VCPIP1 gene is over-expressed (and subsequently knocked down with MYBL1 silencing). Thus far the data strongly suggests a coordinate dysregulation between MYBL1 and VCPIP1 genes. The UBXN2B, VCP, UBA1, NSFL1C and UBXN8 were examined to assess a possible relationship with MYBL1. Since MYBL1 demonstrates reliable concordance with VCPIP1, we examined the possibility that MYBL1 was involved in the Golgi / Endoplasmic Reticulum associated mitotic signaling pathway. To examine this possibility, we examined the mRNA expression of MYBL1, VCPIP1, UBXN2B, VCP, UBA1, NSFL1C, UBXN8 in cell lines (Figure 11). Data show that a similar and highly reproducible pattern of mRNA expression between MYBL1 and VCPIP1 and to some degree UBXN2B. MYBL1 and VCPIP1 results were reproducible 100% of the experimental efforts. UBXN2B were reproducible in 9 of 10 experiments. The cell lines were selected for screening purposes. Genes that consistently demonstrated the exact same

of expression in the cell lines utilized for screening, were considered suitable candidates, and considered for further studies beyond this project. Figure 11 demonstrates an example of how the transcript gene expression pattern of the genes were assessed.

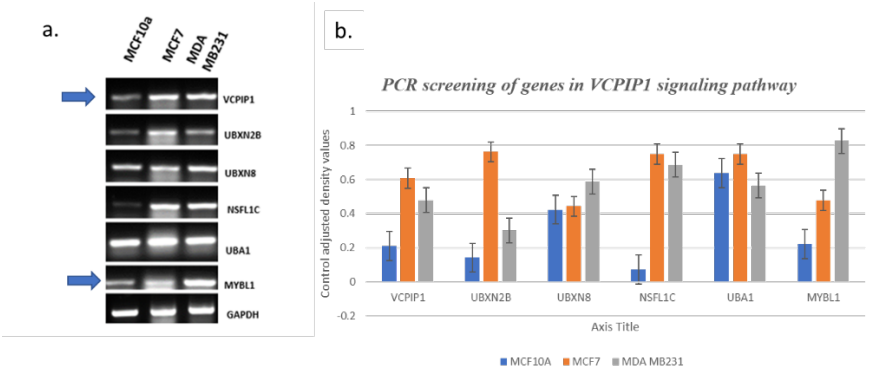


**Figure 11: PCR analyses of the mRNA Expression in Non-tumor Compared to Tumor Cell Lines of Select Genes**

MCF10A is non-tumor TNBC, MCF7 is luminal breast cancer and MDA-MB-231 is TNBC sample. (a) PCR analyses and (b) densitometer analyses of the PCR profiles.

Only genes that consistently demonstrated a reliable pattern of mRNA expression were examined via protein analyses; that included the MYBL1, VCPIP1 and UBXN2B genes. Although UBXN2B expression were reproduced in 9/10 experimental analyses, it was chosen for protein assessment (Figure 12). Like data presented for our other analyses, MYBL1 and VCPIP1 gene show higher protein expression levels in the TNBC compared to the non-tumor triple negative cell line. The UBXN2B antibody showed a pattern opposite of that observed for the mRNA in that the gene appears over-expressed in non-tumor compared to TNBC. Nearly all genes have been identified as having transcript variants. We are finding that different transcript variants display a different role in samples. Related to tumor vs non-tumor, we are frequently finding that some variants have a role

(i.e., are expressed in non-tumors) and some variants have a role (i.e., are expressed in tumors). It could be that the transcript variants probed for mRNA differ from the isoforms detected by the antibody that we chose, hence a different pattern of expression in mRNA vs the protein. To eliminate this circumstance, we suggest the same region used to generate the transcript PCR primers and immunogen for the protein antibody be considered. Sometimes this is not possible due to the limited number of commercially available antibodies.

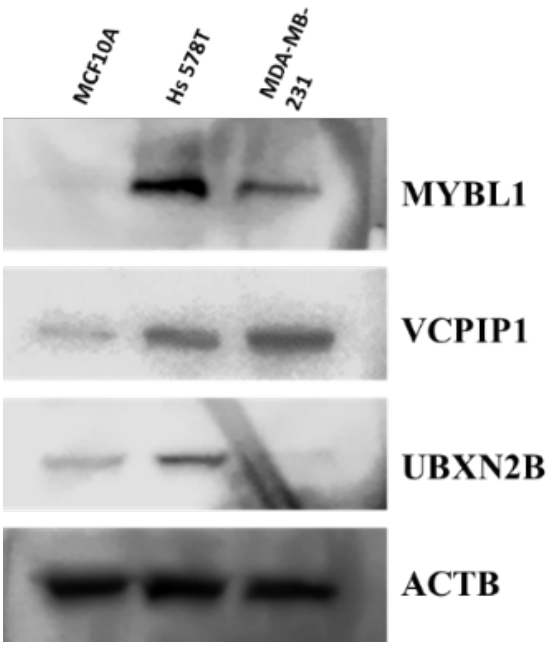


**Figure 12: Western Blot Analysis of Selection Genes**

**MCF10A is the non-tumor triple negative breast samples, compared to TNBC Hs 578T and MDA-MB-231 cell line preparations.**

# CHAPTER 5

## SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

MYBL1 is consistently dysregulated in tumor and TNBC cell lines, validating our initial research observations. MYBL1 and VCPIP1 are coregulated in tumors and TNBC cell lines and clinical patients. The same patients that demonstrate alterations in MYBL1 show alterations in VCPIP1 in 100% of the cases.

It's not clear how the genes cooperate under these conditions. There does not appear to be a direct link between the Golgi Apparatus and Endoplasmic Reticulum associated mitotic signaling pathway. This conclusion is based on the inconsistent results observed for VCP, UBA1 and UBXN8 gene expression profiles and clinical patient profiles (observed using cBioPortal).

Published data by other investigators show the 8q chromosomal loci is a high mutation region in TNBC, and our data show at least 2 additional genes might contribute to this mutation frequency: the genes being MYBL1 and VCPIP1.

## REFERENCES


1. Abdulrahman, N. M. (2021). V-Myb Avian Myel oblast Viral Oncogene Homolog Like 1 (MYBL1) Knockdown and its Role in a Triple Negative Breast Cancer. https://core.ac.uk/download/542922301.pdf
2. Attalla, S., Taifour T. Tailoring therapies to counter the divergent immune landscapes of breast cancer. Front Cell Dev Biol. 2023 Feb 22;11:1111796
3. Bertucci, F., Finetti, P., & Birnbaum, D. (2012). Basal breast cancer: a complex and deadly molecular subtype. *Current molecular medicine*, *12*(1), 96–110. https://doi.org/10.2174/156652412798376134
4. Bolcun-Filas, E., Bannister, L. A., Barash, A., Schimenti, K. J., Hartford, S. A., Eppig, J. J., Handel, M. A., Shen, L., & Schimenti, J. C. (2011). A-MYB (MYBL1) transcription factor is a master regulator of male meiosis. *Development*, *138*(15), 3319–3330. https://doi.org/10.1242/dev.067645
5. Breast Cancer Awareness Month - Stones River Regional Independent Physician Association. https://stonesriveripa.com/breast-cancer-awareness-month/
6. Cerami, E., Gao, J., Dogrusoz, U., Gross, B. E., Sumer, S. O., Aksoy, B. A., Jacobsen, A., Byrne, C. J., Heuer, M. L., Larsson, E., Antipin, Y., Reva, B., Goldberg, A. P., Sander, C., & Schultz, N. (2012). The cBio cancer genomics portal: an open platform

7. for exploring multidimensional cancer genomics data. *Cancer discovery*, *2*(5), 401–404. https://doi.org/10.1158/2159-8290.CD-12-0095

8. Charafe-Jauffret E., Ginestier C. , Monville, F., Adélaïde J. , Gene expression profiling of breast cell lines identifies potential new basal markers. Oncogene. 2006 Apr 6;25(15):2273-84.

9. Cicirò, Y., & Sala, A. (2021). MYB oncoproteins: emerging players and potential therapeutic targets in human cancer. *Oncogenesis*, *10*(2), 19. https://doi.org/10.1038/s41389-021-00309-y

10. Danforth D. N., Jr (2013). Disparities in breast cancer outcomes between Caucasian and African American women: a model for describing the relationship of biological and nonbiological factors. *Breast cancer research : BCR*, *15*(3), 208. https://doi.org/10.1186/bcr3429

11. Fu, N. Y., Nolan, E., Lindeman, G. J., & Visvader, J. E. (2020). Stem cells and the differentiation hierarchy in mammary gland development. *Physiological Reviews*, *100*(2), 489–523. https://doi.org/10.1152/physrev.00040.2018

12. Fujii, K., Murase, T., Beppu, S., Saida, K., Takino, H., Masaki, A., Ijichi, K., Kusafuka, K., Iida, Y., Onitsuka, T., Yatabe, Y., Hanai, N., Hasegawa, Y., & Inagaki, H. (2017). *myb*,*mybl1*,*mybl2* and *nfib* gene alterations and myc overexpression in salivary gland adenoid cystic carcinoma. *Histopathology*, *71*(5), 823–834. https://doi.org/10.1111/his.13281

13. genecards suite. (n.d.-a). https://www.genecards.org/cgi-bin/carddisp.pl?gene=MYBL1

14. genecards suite. (n.d.). https://www.genecards.org/cgi-bin/carddisp.pl?gene=VCPIP1

15. Harbeck, N. Penault-Llorca, F. Breast cancer. Nat Rev Dis Primers. 2019 Sep 23;5(1):66

16. Hu D., Shao W., Yuan S. Intricate crosstalk between MYB and noncoding RNAs in cancer. Cancer Cell Int. 2021 Dec 7;21(1):653.

17. Jiang, Y.-Z., Ma, D., Suo, C., Shi, J., Xue, M., Hu, X., Xiao, Y., Yu, K.-D., Liu, Y.-R., Yu, Y., Zheng, Y., Li, X., Zhang, C., Hu, P., Zhang, J., Hua, Q., Zhang, J., Hou, W., Ren, L., … Shao, Z.-M. (2019). Genomic and transcriptomic landscape of triple-negative breast cancers: Subtypes and treatment strategies. *Cancer Cell*, *35*(3). https://doi.org/10.1016/j.ccell.2019.02.001

18. Lemma, R. B., Ledsaak, M., Fuglerud, B. M., Sandve, G. K., Eskeland, R., & Gabrielsen, O. S. (2021). Chromatin occupancy and target genes of the haematopoietic master transcription factor MYB. *Scientific reports*, *11*(1), 9008. https://doi.org/10.1038/s41598-021-88516-w

19. Liu, X., Xu, Y., Han, L., & Yi, Y. (2018). Reassessing the Potential of Myb-targeted Anti-cancer Therapy. *Journal of Cancer*, *9*(7), 1259–1266. https://doi.org/10.7150/jca.23992

20. Maezawa, S., Sakashita, A., Yukawa, M., Chen, X., Takahashi, K., Alavattam, K. G., Nakata, I., Weirauch, M. T., Barski, A., & Namekawa, S. H. (2020). Super-enhancer switching drives a burst in gene expression at the mitosis-to-meiosis transition. *Nature structural & molecular biology*, *27*(10), 978–988. https://doi.org/10.1038/s41594-020-0488-3

21. Martelotto, L. G., De Filippo, M. R., Ng, C. K., Natrajan, R., Fuhrmann, L., Cyrta, J., Piscuoglio, S., Wen, H. C., Lim, R. S., Shen, R., Schultheis, A. M., Wen, Y. H., Edelweiss, M., Mariani, O., Stenman, G., Chan, T. A., Colombo, P. E., Norton, L., Vincent-Salomon, A., Reis-Filho, J. S., … Weigelt, B. (2015). Genomic landscape of adenoid cystic carcinoma of the breast. *The Journal of pathology*, *237*(2), 179–189. https://doi.org/10.1002/path.4573

22. Mi, H., Huang, X., Muruganujan, A., Tang, H., Mills, C., Kang, D., & Thomas, P. D. (2017). PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic acids research*, *45*(D1), D183–D189. https://doi.org/10.1093/nar/gkw1138

23. MYBL1 protein expression summary - the human protein atlas. (n.d.). https://www.proteinatlas.org/ENSG00000185697-MYBL1

24. Ness SA. Myb binding proteins: regulators and cohorts in transformation. Oncogene. 1999 May 13;18(19):3039-46.

25. Nsende, E. J. (2022). Experimental Validation of Gene Expression of MYBL1, MYBL2, UBXN8, and ADRM1 Genes in Triple Negative Breast Cancer Cell Lines. https://core.ac.uk/download/542922265.pdf

26. Player, A., Abraham, N., Burrell, K., Bengone, I. O., Harris, A., Nunez, L., Willaims, T., Kwende, S., & Walls, W. (2017). Identification of candidate genes associated with triple negative breast cancer. *Genes & cancer*, *8*(7-8), 659–672. https://doi.org/10.18632/genesandcancer.147

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

28. Primer3 Input (version 0.4.0). (n.d.). https://bioinfo.ut.ee/primer3-0.4.0/

29. 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.

30. 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.

31. Sadasivam, S., & DeCaprio, J. A. (2013). The DREAM complex: master coordinator of cell cycle-dependent gene expression. *Nature reviews. Cancer*, *13*(8), 585–595. https://doi.org/10.1038/nrc3556

32. D. Szklarczyk, R. Kirsch, M. A.-O. Koutrouli, K. A.-O. Nastou, F. A.-O. Mehryary, R. Hachilif. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res.* 2023 Vol. 51(D1) Pages D638-D646.

33. Suh, Y. Y., Lee, K., Shim, Y. M., Phi, J. H., Park, C. K., Kim, S. K., Choi, S. H., Yun, H., & Park, S. H. (2023). MYB/MYBL1::QKI fusion-positive diffuse glioma. *Journal of neuropathology and experimental neurology*, *82*(3), 250–260. https://doi.org/10.1093/jnen/nlac123

34. Sütçü, M. (2023). Optimal Time Intervals Between Personal Mammogram Test Decisions. International Journal of Advanced Natural Sciences and Engineering Researches. https://doi.org/10.59287/ijanser.650

35. Tang, H., & Goldberg, E. (2012). A-MYB (MYBL1) stimulates murine testis-specific LDHC expression via the camp-responsive element (CRE) site1. *Biology of Reproduction*, *86*(2). https://doi.org/10.1095/biolreprod.111.095661

36. Uchiyama, K., Jokitalo, E., Kano, F., Murata, M., Zhang, X., Canas, B., Newman, R., Rabouille, C., Pappin, D., Freemont, P., & Kondo, H. (2002). VCIP135, a novel essential factor for p97/p47-mediated membrane fusion, is required for Golgi and ER assembly in vivo. *The Journal of cell biology*, *159*(5), 855–866. https://doi.org/10.1083/jcb.200208112

37. VCPIP1 protein expression summary - the human protein atlas. (n.d.). https://www.proteinatlas.org/ENSG00000175073-VCPIP1

38. Vegiopoulos, A., García, P., Emambokus, N., & Frampton, J. (2006). Coordination of erythropoiesis by the transcription factor c-Myb. *Blood*, *107*(12), 4703–4710. https://doi.org/10.1182/blood-2005-07-2968

39. Wang, Y., Satoh, A., Warren, G., & Meyer, H. H. (2004). VCIP135 acts as a deubiquitinating enzyme during P97–P47-mediated reassembly of mitotic golgi fragments. *The Journal of Cell Biology*, *164*(7), 973–978. https://doi.org/10.1083/jcb.200401010

40. Wu, Q., Zhang, L., Xu, X., Zhang, Y., Shi, J., Lin, X., & Chen, W. (2022). Hepatitis B Virus X Protein Is Stabilized by the Deubiquitinating Enzyme VCPIP1 in a Ubiquitin-Independent Manner by Recruiting the 26S Proteasome Subunit PSMC3. *Journal of virology*, *96*(13), e0061122. https://doi.org/10.1128/jvi.00611-22

41. 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

42. Yedjou, C. G., Sims, J. N., Miele, L., Noubissi, F., Lowe, L., Fonseca, D. D., Alo, R. A., Payton, M., & Tchounwou, P. B. (2019). Health and Racial Disparity in Breast Cancer. *Advances in experimental medicine and biology*, *1152*, 31–49. https://doi.org/10.1007/978-3-030-20301-6_3

43. Zhang, B., Kirov, S., & Snoddy, J. (2005). WebGestalt: an integrated system for exploring gene sets in various biological contexts. *Nucleic acids research*, *33*(Web Server issue), W741–W748. https://doi.org/10.1093/nar/gki475

44. Zhou, Y., & Ness, S. A. (2011). Myb proteins: angels and demons in normal and transformed cells. *Frontiers in bioscience (Landmark edition)*, *16*(3), 1109–1131. https://doi.org/10.2741/3738

45. Zhu, J., Wu, Y., Yu, Y., Li, Y., Shen, J., & Zhang, R. (2022). MYBL1 induces transcriptional activation of ANGPT2 to promote tumor angiogenesis and confer sorafenib resistance in human hepatocellular carcinoma. *Cell death & disease*, *13*(8), 727. https://doi.org/10.1038/s41419-022-05180-2

46. 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.