DETERMINING THE ROLES OF MYB FAMILY TRANSCRIPTION FACTORS IN BREAST TUMORIGENESIS

Aaron R. Thorner

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Curriculum in Genetics and Molecular Biology

Chapel Hill
2010

Approved by:

Charles M. Perou
Albert S. Baldwin, Jr.
Larysa H. Pevny
W. Kimryn Rathmell
Lola M. Reid

ABSTRACT

AARON R. THORNER: Determining the Roles of MYB Family Transcription Factors in Breast Tumorigenesis
(Under the direction of Charles M. Perou)

A major advancement in the field of breast cancer research was the discovery of the breast tumor intrinsic subtypes made through the utilization of gene expression microarrays. Breast cancer can no longer be viewed as a single disease, but rather as at least five different diseases each with unique biological activity and clinical outcomes. Targeted therapy strategies are now employed to treat the different tumor types, such as estrogen receptor modulators for ER-positive disease, and HER2-inhibitors for the treatment of HER2-positive tumors. For tumors lacking therapeutic targets, patients are limited to cytotoxic chemotherapy regimens. Consequently, additional research is crucial in further elucidating the molecular pathways governing each breast tumor subtype.

Over one thousand genes are used to stratify the intrinsic molecular subtypes; however, very few of these genes have been analyzed for their direct role in tumorigenesis. This dissertation focuses on investigating two intrinsic genes, *B-Myb* and *c-Myb*, which are both members of an evolutionarily conserved gene family first identified as transforming genes in avian viruses. *B-Myb* is highly expressed in basal-like tumors, whereas *c-Myb* is highly expressed in luminal tumors. We applied *in vitro* and *in vivo* analyses to ascertain the roles of these genes within the molecular subtypes.

High *B-Myb* expression levels were predictive of poor outcomes across all breast tumors and within subtypes. Mammary epithelial cells expressing high levels of *B-Myb* were more sensitive to topoisomerase 2α inhibitors, but not other chemotherapeutics, via the induction of G2/M cell cycle genes including TOP2A itself. We identified the first published *B-Myb* germline variant causing an increased risk for basal-like disease.

We found that the *c-Myb* oncogene was behaving as a tumor suppressor in luminal breast cancer through a novel p53 stabilization pathway. These results have significant treatment implications in light of an ongoing hematologic malignancies clinical trial in which *c-Myb* is targeted for knock-down through antisense oligonucleotides. These results point to both *B-Myb* and *c-Myb* as important breast cancer biomarkers with potential clinical importance for determining disease risk and guiding treatment, and provide important insight into the roles of MYB family proteins in the etiology of breast cancer.

ACKNOWLEDGEMENTS

I would like to thank:

My advisor: Charles M. Perou

The Perou Lab, all past and present members.

My committee members:

Albert S. Baldwin, Jr.
Larysa H. Pevny
W. Kimryn Rathmell
Lola M. Reid

My wife and child: Michelle and Elsa Thorner

My parents and sister: Kathleen, Richard and Alexis Thorner

TABLE OF CONTENTS

LIST OF TABLES ix

LIST OF FIGURES x

LIST ABBREVIATIONS AND SYMBOLS xii

CHAPTER

I INTRODUCTION 1

Normal Breast Development 2

Breast Cancer Molecular Subtypes 5

Mammary Cellular Hierarchy and Relevance to Molecular Subtypes 8

MYB Family of Transcription Factors 11

Research Introduction 15

References 16

II *IN VITRO* AND *IN VIVO* ANALYSIS OF B-MYB IN BASAL-LIKE BREAST CANCER 20

Preface 20

Abstract 21

Introduction 22

Materials and Methods 24

Cell Lines 24

Western Blot Analysis 25

Cell Cycle Analysis 25

Cytotoxicity Assay 26

Microarray Analysis 26

Microarray Statistical Analyses 27

Carolina Breast Cancer Study 27

Genotyping of *B-Myb* 28

Genotype Statistical Analysis 29

Results 29

High *B-Myb* Expression in Breast Tumors Predicts Poor Outcome 29

B-Myb Germline Variant (rs2070235) Increases Risk of Basal-Like Breast Cancer 35

Ectopic Expression of *B-Myb* Increases Sensitivity to TOP2A Inhibitors 36

Gene Expression Analysis of Cell Lines Ectopically Expressing *B-Myb*.....38

Cell Cycle Profiles of *B-Myb* Overexpressing Cells.....43

Discussion.....43

Acknowledgements.....52

References.....53

III TUMOR SUPPRESSOR ROLE FOR THE C-MYB ONCOGENE IN LUMINAL BREAST CANCER 57

Preface.....57

Abstract.....58

Introduction.....59

Materials and Methods.....60

Cell Culture.....60

In Vitro Analyses 61

In Vivo Tumor Analysis.....62

Microarray Analysis.....62

Microarray Statistical Analyses.....63

Other Statistical Analyses.....63

Results....64

c-Myb High Expression Correlates with Good Prognosis....64

c-Myb Knockdown in MCF7 Cells Increases Tumorigenesis in vitro and *in vivo*....66

c-Myb Expression Signature Identifies Many Luminal/ER+ Subtype Defining Genes....71

c-Myb Gene Signature Correlates with Mature Luminal Mammary Cell Lineage....75

Discussion....78

Acknowledgements....86

References....87

IV CONCLUSION....91

References....103

APPENDIX

II Genes expressed significantly higher in doxorubicin-treated *B-Myb* overexpressingHME-CC cells versus doxorubicin-treated vector control cells....106

LIST OF TABLES

Table

2.1 High *B-Myb* expression in breast tumors (n=133; Hess *et al*., 2006) is associated with pathological complete response (pCR).....35

2.2 *B-Myb* variant genotype (G/G) increases the risk of basal-like breast cancer.....36

2.3 Top 15 significant gene ontology categories determined by EASE analysis for *B-Myb* associated genes.....41

3.1 Pathologic complete response, pCR, data of Hess *et al*., (2006) rank ordered, split into halves (high/low) based on *c-Myb* expression values, and analyzed by chi-square.....66

3.2 IC50 doses (72h) of tamoxifen on c-Myb knock-down (shMYB) versus control (shGFP) 67

3.3 c-Myb, ER, and GATA3 gene signature define many luminal cluster genes.....73

LIST OF FIGURES

Figure

1.1 Immunohistochemistry of a normal, human mammary gland 4

1.2 Breast tumors hierarchical clustered using the intrinsic gene list 6

1.3 Kaplan-Meier plot depicting overall survival for each subtype 7

1.4 Current model of mammary cell differentiation 10

1.5 MYB family proteins 12

2.1 *B-Myb* expression across breast cancer subtypes 30

2.2 High expression of *B-Myb* correlates with poor survival 32

2.3 High expression of *B-Myb* correlates with poor outcome 33

2.4 Western blot analysis of ectopic expression of *B-Myb* or *B-Myb* variant in basal-like mammary cell lines 37

2.5 Drug sensitivities in *B-Myb* overexpressing cell lines 39

2.6 Enrichment of G2/M phase genes in doxorubicin-treated *B-Myb* overexpressing HME-CC cells 42

2.7 Cell cycle profile of HME-CC cells stably expressing *B-Myb* and treated with doxorubicin 44

2.8 Cell cycle profile of HME-CC cells stably expressing *B-Myb* and treated with etoposide 44

2.9 *B-Myb* expression may increase anti-proliferative effects of TOP2A inhibitors 49

3.1 *c-Myb* high expression correlates with luminal breast cancer, small tumor size and low tumor grade 65

3.2 *c-Myb* high expression correlates with good prognosis in all breast subtypes combined and luminal tumors 68

3.3 c-Myb knock-down increases luminal tumor growth both *in vitro* and *in vivo* 69

3.4 Luminal gene cluster....72

3.5 *c-Myb* gene signature correlates with mature (differentiated) luminal mammary cells....76

3.6 c-Myb induces p53 stabilization and activation in a Hep27 dependent manner....81

3.7 Potential c-Myb-Hep27-Mdm2-p53 pathway in luminal breast cancer 82

LIST OF ABBREVIATIONS AND SYMBOLS

5-FU	5-fluorouracil
95% CI	95% confidence interval
ALDH1A1	aldehyde dehydrogenase 1 family, member A1
ANOVA	analysis of variance
BRCA1	breast cancer 1, early onset
BSA	bovine serum albumin
CBCS	Carolina Breast Cancer Study
CDK2	cyclin-dependent kinase 2
CK 5/6	cytokeratins 5/6
CK 8/18	cytokeratins 8/18
CK17	cytokeratin 17
CNS	central nervous system
Cy3 or 5	cyanine dye 3 or 5
dMYB	*Drosophila melanogaster* MYB gene
EASE	expression analysis systematic explorer
EpCAM	epithelial cell adhesion molecule
ER	estrogen receptor alpha
ESR1	estrogen receptor alpha gene
FACS	fluorescence activated cell sorting
FDR	false discovery rate
GEO	gene expression omnibus

HER1	epidermal growth factor receptor (erythroblastic leukemia viral (v-erb-b) oncogene homolog
HER2/ERBB2	v-erb-b2 erythroblastic leukemia viral oncogene homolog 2
HMEC	human mammary epithelial cell
hTERT	human telomerase reverse transcriptase
IACUC	institutional animal care and use committee
IC50	inhibitory concentration 50%
LP	luminal progenitor mammary cell population
LRT	likelihood ratio test
MaSC	mammary stem cell
ML	mature luminal mammary cell population
MMTV	murine mammary tumor virus
MTT	mitochondrial dye conversion assay [3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide]
NKI	Netherlands Cancer Institute
OR	odds ratio
OS	overall survival
PAM	prediction analysis of microarrays
PBS	phosphate buffered saline
pCR	pathologic complete response
PR	progesterone receptor
RB	retinoblastoma protein
RFS	relapse-free survival
SAM	significance analysis of microarrays
shGFP	non-specific short hairpin RNA directed against GFP

shMYB	short hairpin RNA directed against c-Myb
shRNA	short hairpin RNA
SMA	α-smooth muscle actin
T/FAC	5-fluorouracil-adriamycin-cyclophosphamide
TDLU	terminal duct lobular unit
TOP1	DNA topoisomerase I
TOP2A	DNA topoisomerase 2α
Tx	treatment
UMD	University of North Carolina at Chapel Hill Microarray Database
UNC	University of North Carolina at Chapel Hill

CHAPTER I

INTRODUCTION

Breast cancer is the second leading cause of cancer death worldwide and approximately one in eight women in the United States will develop invasive breast cancer in her lifetime. The incidence of breast cancer over the past several years has been increasing, partially due to breast cancer awareness campaigns promoting frequent screenings, and increased longevity; however, the mortality rate has decreased thanks to dramatic improvements in patient care and early detection. One of the major treatment advances has been in the identification of the estrogen receptor and drugs that target this protein/pathway; specifically, patients with estrogen receptor positive disease are treated with selective estrogen receptor modulators, such as tamoxifen, to prevent recurrence, and/or aromatase inhibitors, which are a class of drug used to reduce circulating estrogen. There have also been significant improvements in the treatment of estrogen receptor negative breast cancer using targeted therapies, such as trastuzumab, which is a monoclonal antibody directed against HER2-positive breast cancer. Lastly, cytotoxic chemotherapy is still the standard of care for many patients today. Despite these treatment advancements approximately 40,000 women in the U.S. die of breast cancer each year; therefore additional studies and improvements are still needed.

A better understanding of the molecular etiology of breast cancer is of utmost importance in order to discover novel therapeutic targets. Therefore, my research has

focused on the MYB family of transcription factors, specifically the genes *B-Myb* and *c-Myb*. These genes have implications in other tumor types, and are highly expressed in distinct subsets of breast tumors. Identifying the contributions of both B-Myb and c-Myb to mammary tumorigenesis may improve upon our current understanding of breast cancer disease progression and could lead to improvements in both disease prediction and patient treatment.

Normal Breast Development

To better understand mammary carcinogenesis, comprehension of mammary development is imperative. The human breast is a dynamic organ that undergoes several dramatic hormone and growth factor-directed changes in shape, size, and function throughout a woman's lifetime (Howard and Gusterson, 2000; Russo and Russo, 2004). The fact that mammary cells are in this continuous state of flux increases the likelihood of tumorigenesis.

The mammary gland arises early *in utero* from an epithelial ectodermal bud that invades into the dermis (Parmar *et al.*, 2002). At the time of birth, the infant breast consists of a primitive, blunt-ended, ductal system comprised of one or two inner layers of luminal (epithelial) cells, and one layer of myoepithelial cells surrounding the lumen (Russo and Russo, 2004). The postnatal mammary gland, no longer influenced by the mother's hormones, undergoes differentiation and involution by the age of two, leaving small tubular structures surrounded by fibroblastic stroma (Anbazhagan *et al.*, 1991). The prepubescent gland remains relatively unchanged until the start of puberty.

Both the ductal epithelium and the surrounding stroma (a support structure consisting of fibroblasts, endothelial cell, lymphocytes, and adipocytes) transform during puberty. At

this time, there is a dramatic increase in adipose and fibroblastic tissue, and the ducts elongate and branch, terminating with the future β-casein (milk)-producing cells in the terminal duct lobular units (TDLUs), or "alveolar buds" (Villadsen, 2005). The adult mammary gland consists of a main lactiferous duct attached to the nipple that branches into 15-25 milk ducts, which further branch into many sub-segmental ducts. It is continually influenced by hormones, with each ovulation promoting a small amount of additional budding.

Three major cell lineages comprise the ducts: the luminal epithelial cells lining the inside of the duct, the myoepithelial cells, which form the basal layer of the duct and surround the luminal epithelial cells, and the milk-producing luminal-alveolar epithelial cells localized to the TDLUs. Several markers are commonly used to distinguish the mature luminal and myoepithelial cells by immunohistochemistry including α-smooth muscle actin (SMA), and cytokeratins 5/6 and 17 for the myoepithelial cells, and for luminal cells, estrogen receptor alpha (ERα), GATA-3, and cytokeratins 8/18 (Figure 1.1).

During pregnancy there occurs a substantial increase in the number of alveoli as well as novel formation of lobules by budding from TDLUs (Villadsen *et al.*, 2007). After birth, milk is produced by the alveolar cells in the TDLUs, which is then secreted into the ducts and flows to the nipple. After weaning, the postlactational TDLUs and acini (clusters of milk-producing lobules) undergo apoptosis during involution. Finally, during postmenopausal involution, the number of both ducts and lobules are reduced (Howard and Gusterson, 2000). The extensive remodeling this organ undergoes throughout a woman's lifespan suggests the existence of a mammary stem cell.

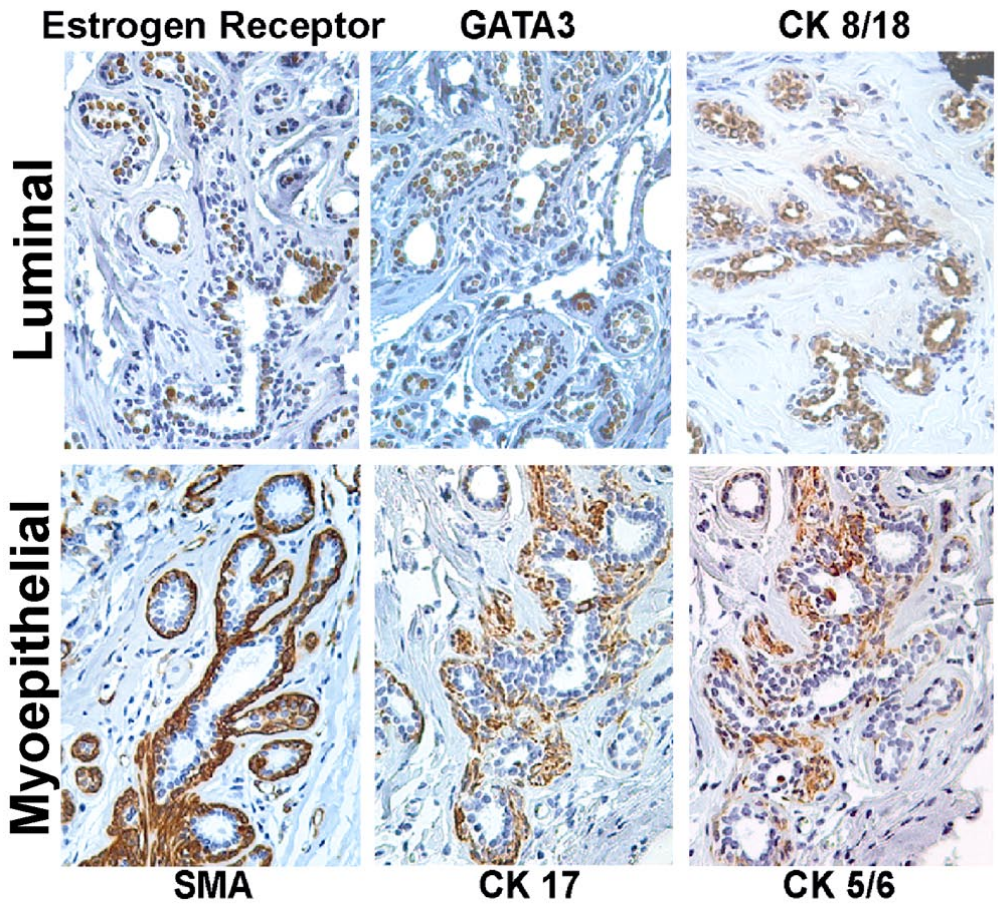


Figure 1.1. Immunohistochemistry of a normal, human mammary gland. Tissue stained for proteins uniquely expressed in luminal epithelial cells (ER, GATA3, CK 8/18), or myoepithelial cells (SMA, CK17, CK 5/6). ER: estrogen receptor alpha, CK: cytokeratin, SMA: alpha-smooth muscle actin

Breast Cancer Molecular Subtypes

Breast cancer is a heterogeneous disease pathologically, molecularly, and clinically. Over the past decade, the use of gene expression microarrays by our lab and others has revealed the existence of at least six molecular subtypes of breast cancer (Herschkowitz *et al.*, 2007; Hu *et al.*, 2006; Parker *et al.*, 2009; Perou *et al.*, 2000; Sorlie *et al.*, 2003). These subtypes include luminal A, luminal B, normal breast-like, HER2-enriched, basal-like, and claudin-low, each with unique gene expression profiles (Figure 1.2) and distinct clinical outcomes (Figure 1.3). The molecular subtypes were initially identified using a 496 "intrinsic" gene set; these were genes whose expression varied greatly between different tumors, but showed little variation between paired samples of the same tumor (Perou *et al.*, 2000). More recently, our lab has created a 50-gene subtype predictor (PAM50) which adds significant prognostic and predictive information to the commonly used clinical parameters for breast cancer patients (Parker *et al.*, 2009).

The foremost determinant of subtype is the presence or absence of estrogen receptor-alpha (ESR1/ER) expression. Both luminal A and luminal B tumors are ER-positive and comprise the majority (60-80 percent) of breast cancer cases. The luminal A tumors have markedly high gene expression of the transcription factors GATA3, XBP1, FOXA1, (Lacroix and Leclercq, 2004) and cytokeratins 8 and 18. The luminal B tumors also express these genes, but at lower levels, and additionally have high expression of the "proliferation gene signature"; this is a dominant gene signature that is a marker of cell proliferation rates across of variety of tumor types (Hu *et al.*, 2006; Whitfield *et al.*, 2006). Normal breast-like tumors cluster with the ER-positive subtype and have similar gene expression patterns to normal

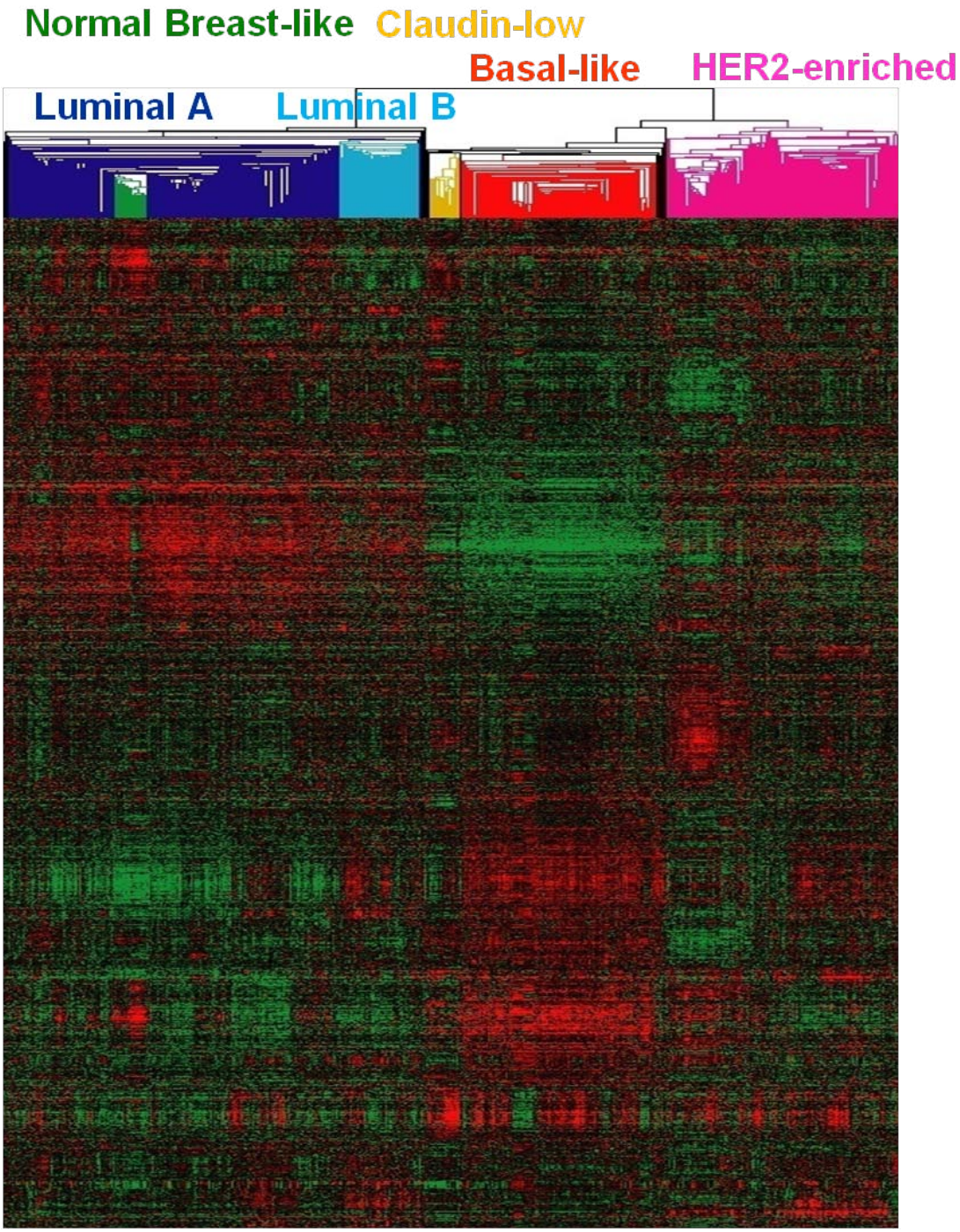


Figure 1.2. Breast tumors hierarchical clustered using the intrinsic gene list. A combined data set of 324 tumor samples collected at UNC was clustered using the intrinsic gene set (Hu *et al*., 2006). Clustering identified the six major intrinsic subtypes: luminal A, luminal B, normal breast-like, basal-like and HER2-enriched, and claudin-low.

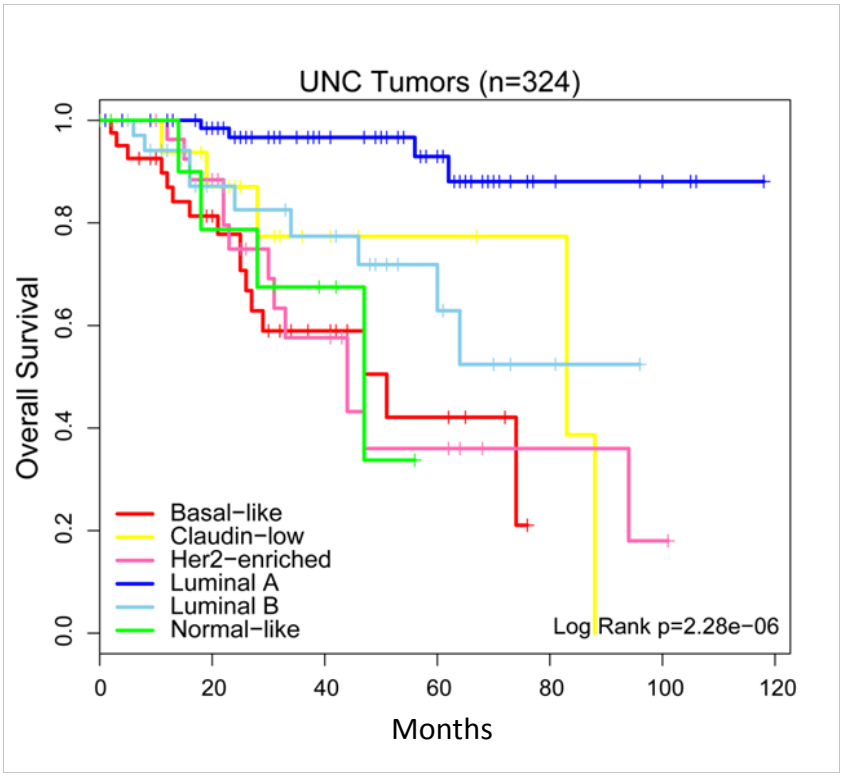


Figure 1.3. Kaplan-Meier plot depicting overall survival for each subtype.

tissue, most likely due to the tumor specimens containing high amounts of "contaminating" normal mammary tissue.

The ER-negative subtypes consist of the HER2/ERBB2-enriched, basal-like, and claudin-low tumors. The HER2-enriched tumors overexpress HER2 and are typically amplified at this locus at the DNA level. Basal-like disease makes up 10-15 percent of breast carcinomas and was named "basal-like" due to the observed high expression of cytokeratins 5 and 17; cell surface markers expressed in normal mammary basal/myoepithelial cells. However, it is likely that this tumor type is derived from luminal epithelium and not myoepithelium, as malignant myoepithelioma of the breast is known to arise from the myoepithelial cell population and is an extremely rare tumor (Sauer, 2007). Basal-like tumors are sometimes referred to clinically as "triple-negative" tumors due to the absence of

immunohistochemical staining for ER, progesterone receptor (PR), and HER2, although this is not true of all basal-like tumors (Carey *et al.*, 2007). Tumors of this subtype also have high expression of the proliferation gene signature. The recently identified claudin-low subtype is ER-negative and is defined by low expression of tight junction and cell-to-cell adhesion proteins including claudins 3, 4, 7, occludin, and E-cadherin (Herschkowitz *et al.*, 2007). Claudin-low tumors share features with both normal mammary stem cells, including low expression of EpCAM and mucin-1 (Hennessy *et al.*, 2009), and also mammary tumor initiating cells, which are described by the cell surface markers $CD24^{lo}$, $CD44^{+}$ (Al-Hajj *et al.*, 2003).

Mammary Cellular Hierarchy and Relevance to Molecular Subtypes

Breast cancer was studied as one disease for many decades. However, data collected over the past several decades, notably gene expression analysis, has revealed that breast cancer not one disease, but rather a group of complex diseases described above. The molecular subtyping of breast cancer also allows us to study the etiology of each tumor type. The mammary epithelial cell population most likely exists as a hierarchy of cell types, from undifferentiated stem cells, to mature luminal cells, and the molecular subtypes may reveal clues as to the cell-of-origin for each tumor type.

Many, if not all, human organs are made up of a cellular hierarchy, from undifferentiated, tissue-specific stem cells, through lineage-specific progenitors, and finally mature, differentiated cells that make up the bulk of the organ. Current evidence suggests this to also be true for the mammary gland epithelium. It has been hypothesized that an ER-negative stem cell can self-renew or undergo differentiation leading to committed progenitor populations (it is unknown how many progenitor types) that will eventually give rise to the

myoepithelial, ductal epithelial, and alveolar cell populations that comprise the epithelial component of the breast (Dontu *et al.*, 2004). The mammary stem cell was also reported as not expressing the progesterone receptor, or keratins 8/18 (luminal markers), but did express keratins 5/6 and 14 (basal markers). This is in contrast to the luminal progenitor population which expresses both basal keratins 5/6 and luminal keratins 8/18 (Figure 1.4) (Boecker and Buerger, 2003; Clarke, 2005; Livasy *et al.*, 2005; Woodward *et al.*, 2005). Finally, the mature luminal cells express ER, GATA3 and luminal keratins 8/18, with the luminal alveolar cells expressing milk protein in addition to these genes.

It has been reported that many cancers contain, and are driven by, tumor stem cells; a rare cell population within the bulk of the tumor with the capacity for limitless self-renewal and the ability to promote tumor formation (Clarke *et al.*, 2006; Jordan *et al.*, 2006; Reya *et al.*, 2001; Wicha *et al.*, 2006). Human mammary tumor stem cells have been isolated using $CD44^+$ and $CD24^-$ (Al-Hajj *et al.*, 2003; Shipitsin *et al.*, 2007). Also, ALDH1A1 (aldehyde dehydrogenase 1 family, member A1)-positive cells from the normal human mammary gland were capable of mammosphere formation (an *in vitro* assay used to determine stem cell potential) as well as mammary structure formation in a humanized mouse mammary gland, suggesting ALDH1A1 is expressed in the normal and tumor stem cell populations (Ginestier *et al.*, 2007).

Based on cell surface marker expression as well as recent complex gene expression analyses comparing gene expression signatures of normal mammary epithelial populations to breast tumor gene expression patterns has led to a hypothetical model of breast tumor cell-of-origin (Figure 1.4) (Lim *et al.*, 2009; Prat and Perou, 2009). However, it is possible that

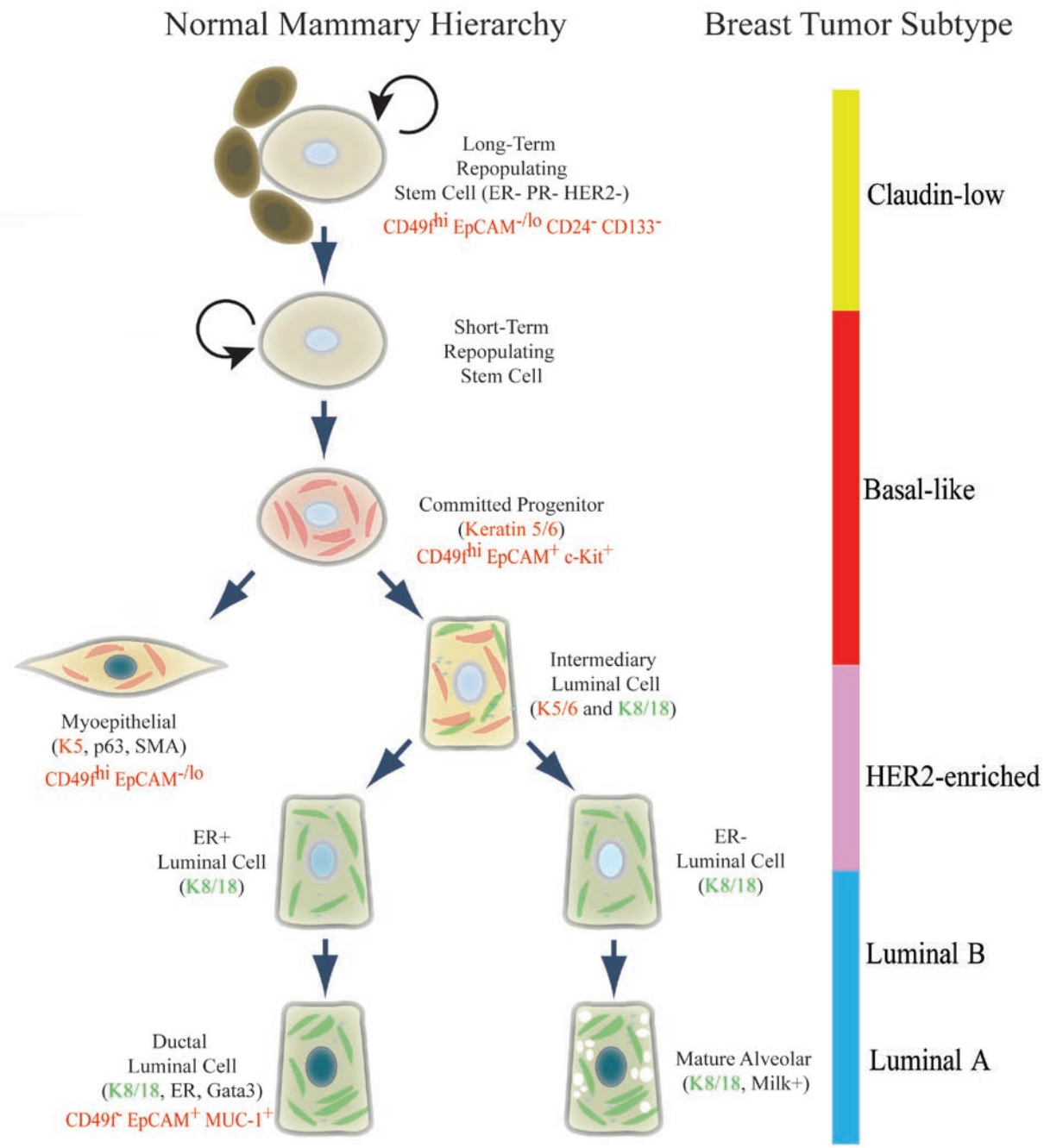


Figure 1.4. Current model of mammary cell differentiation. Immunohistochemical markers known to stain positive are in parenthesis for each cell type. Cell surface markers used in isolation of various lineages are noted in red. Hypothetical tumor cell-of-origin noted by bar on right. Cell images were created by Jason Herschkowitz. Figure was also adapted from Visvader, 2009 and Pratt and Perou, 2009.

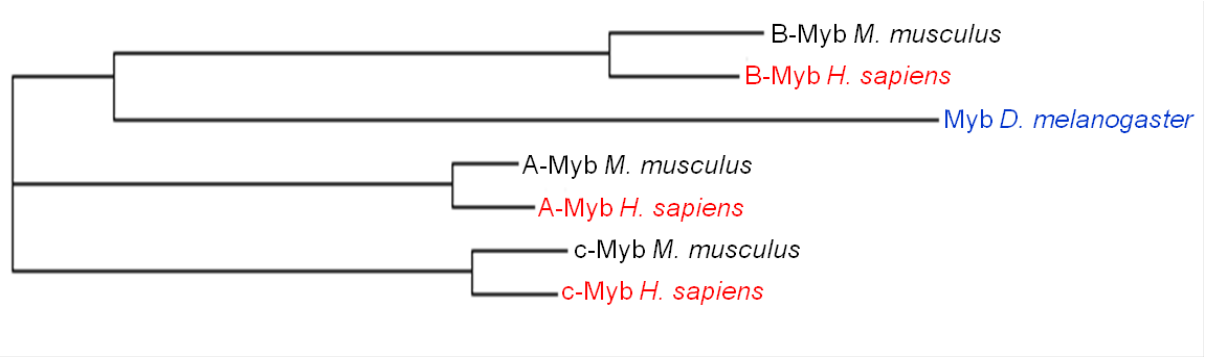
tumor gene expression may not reflect that of the cell-of-origin, and it is still necessary to validate these correlative findings experimentally.

MYB Family of Transcription Factors

Many genes are differentially expressed across the breast tumor molecular subtypes, however, the contribution of the vast majority of these genes to tumorigenesis is unknown. The vertebrate MYB family of nuclear transcription factors consists of three members: MYB (*c-Myb*), MYBL1 (*A-Myb*), and MYBL2 (*B-Myb*) (Figure 1.5A), all of which are found in the intrinsic gene set (Hu *et al.*, 2006) and all of which show distinctive expression signatures. It is these distinctive signatures that and previously known importance of this gene family that caused us to study these genes in depth in order to better define their roles in breast development and tumorigenesis.

The *c-M*yb proto-oncogene was first identified as the mammalian homolog of *v-myb*, which is the transforming gene transmitted by the avian myeloblastosis and E26 retroviruses that causes acute leukemia in birds (Klempnauer *et al.*, 1982; Leprince *et al.*, 1983). *A*- and *B-Myb* were later discovered during low stringency screening of human cDNA libraries with a *c-Myb* probe (Nomura *et al.*, 1988). Found in the genomes of both plants and animals, the MYB genes are conserved throughout evolution and control processes from flavonoid production to cellular proliferation (Ito *et al.*, 2001; Rosinski and Atchley, 1998). In contrast to vertebrates, invertebrates contain only one MYB protein, which in Drosophila (dMYB) is phylogenetically and functionally complementary to vertebrate B-Myb, suggesting B-Myb to be the most ancient family member (Figure 1.5A) (Davidson *et al.*, 2005).

(A)



(B)

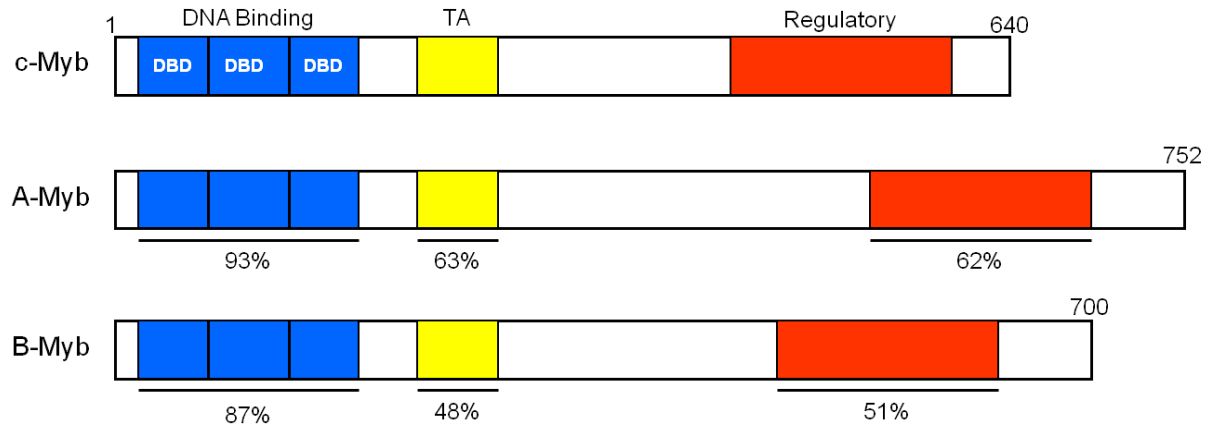


Figure 1.5. MYB family proteins. (A) Phylogram depicting the evolutionary relationships between drosophila, human, and mouse MYB proteins (ClustalW v.1.83). (B) Protein structural homologies between MYB family members. Color bars indicate conserved domains between the three family members. Percentages represent sequence homology relative to c-Myb. The DNA binding region consists of three DNA binding domains (DBD) recognizing the consensus sequence PyAAC(G/T)G. TA: transcriptional activation domain. Figure adapted from (Ness, 2003; Oh and Reddy, 1999; Rosinski and Atchley, 1998; Weston, 1998).

All three human MYB protein sequences are highly conserved, with over 60% amino acid sequence identity (Figure 1.5B). The protein structure can be divided into three major domains: DNA-binding, transcriptional activation, and regulatory (Ness, 2003). The highest sequence homology (~90%) between the three members is found near the amino-terminus of the protein (residues 1-200), which contains three "MYB DNA binding domains". Each domain consists of approximately 50 amino acids with the ability to recognize and bind to the DNA consensus sequence PyAAC(G/T)G (Rosinski and Atchley, 1998). The carboxyl-terminal end of the protein contains the regulatory domain and shares the second highest, albeit small, region of homologous sequence. This domain contains residues which, upon phosphorylation or other modification, confer a conformational change allowing for transactivation (Weston, 1998). Finally, the central region of the MYB family proteins contains the transcriptional activation domain. This domain has the least sequence homology between the three family members and may allow for protein-protein interactions with unique transcriptional co-factors to alter transcriptional activity (Sala, 2005). For example, B-Myb is induced during late G1/S-phase and the regulatory domain is phosphorylated by CDK2 and mitotic cyclins to activate its ability to interact with DNA and promote transcription of cell cycle genes (Sala, 2005).

In vitro studies have shown that each family member has the ability to transactivate genes containing the MYB DNA consensus sequence; however, varying tissue-specific expression patterns of each protein suggests distinct biological roles. Until recently, the expression of *c-Myb* was largely observed to occur in early lineage hematopoietic cells. Mice homozygous null for *c-Myb* die by embryonic day 15 (E15) due to severe anemia caused by improper erythropoiesis (Mucenski *et al.*, 1991). *A-Myb* expression is also

confined to specific tissue types, namely germinal B lymphocytes, the developing central nervous system, the male reproductive system, and the proliferating mammary gland (Ness, 2003; Trauth *et al.*, 1994). *A-Myb* knockout mice are viable; however, males are sterile due a defect in spermatogenesis, and females cannot feed their pups due to a lack of ductal proliferation in the mammary gland during pregnancy (Toscani *et al.*, 1997). In contrast to *c-Myb* and *A-Myb*, the expression of *B-Myb* is ubiquitous to virtually all proliferating cells and plays an essential role in vertebrate development (Sala, 2005). Knocking out murine *B-Myb* causes early embryonic lethality (E4.5-6.5) resulting from unsuccessful inner cell mass formation, suggesting a possible role for B-Myb in embryonic stem cells (Tanaka *et al.*, 1999).

Since MYB transcription factors play significant roles in cell proliferation, and in fact one MYB protein is a viral transforming protein, the altered expression of these proteins has been observed and implicated in a variety of tumors. Amplification of the *B-Myb* locus (20q13) is frequently found in breast, liver, ovarian, and prostate tumors (Sala, 2005). This gene is listed in the top 5 "markers of treatment failure" of embryonal tumors of the central nervous system (CNS) and high expression of *B-Myb* in cancer generally correlates with poor patient outcomes (Pomeroy *et al.*, 2002). However, the roles of B-Myb in breast tumor progression, as well as its mammary transcriptional targets, and the significance of its expression during treatment, are still poorly understood.

Amplification of the *c-Myb* oncogene locus at 6q22-23 is observed in a variety of tumors including T-cell leukemia, acute myelogenous leukemia, colon cancer, and adenoid cystic carcinomas of the head, neck and breast (Oh and Reddy, 1999; Persson *et al.*, 2009). *c-Myb* has been implicated in progenitor cell maintenance in the hematopoietic system,

neuroblastoma cell cultures, and colonic crypts (Malaterre *et al.*, 2007; Maurice, 2007; Thiele *et al.*, 1988). In breast cancer it is reported that *c-Myb* expression is increased in ERα-positive tumors and its locus is amplified in some (29%) hereditary, *BRCA1*-mutant breast tumors (Guérin M, 1990; Kauraniemi *et al.*, 2000). *c-Myb* was observed by us, and others, to be highly expressed in virtually all ER+/luminal tumors and is sporadically found highly expressed in a few basal-like tumors, consistent with previously reported data for *BRCA1*-associated tumors (which tend to be basal-like). The function of the *c-Myb* oncogene in breast cancer is not well studied. Understanding the contribution of c-Myb to the ER+/luminal tumors may lead to better treatment options for patients with this disease.

Research Introduction

A great deal of direct and indirect evidence suggests that the altered expression/function of MYB family of proteins can contribute to tumorigenesis. In breast tumors, we have observed distinct expression patterns for each MYB family member with different tumor subtypes showing very high expression is some subtypes and low to absent expression in others. We hypothesize that altered expression levels of *B-Myb* or *c-Myb* in the mammary gland *in vivo* may contribute to faster tumor formation, or decreased tumor latency, and therefore, we tested these hypotheses using a combination of *in vitro* and *in vivo* methods to determine the role of the MYB family members in breast tumorigenesis. Chapters two and three are devoted to evaluating the *in vitro* and *in vivo* roles of altered expression of *B-Myb* and *c-Myb*, respectively.

REFERENCES

Al-Hajj M, Wicha MS, Benito-Hernandez A, Morrison SJ, Clarke MF (2003). Prospective identification of tumorigenic breast cancer cells. *PNAS* **100:** 3983-3988.

Anbazhagan R, Bartek J, Monaghan P, Gusterson BA (1991). Growth and development of the human infant breast. *Am J Anat.* **192:** 407-417.

Boecker W, Buerger H (2003). Evidence of progenitor cells of glandular and myoepithelial cell lineages in the human adult female breast epithelium: a new progenitor (adult stem) cell concept. *Cell Proliferation* **36:** 73-84.

Carey LA, Dees EC, Sawyer L, Gatti L, Moore DT, Collichio F *et al* (2007). The Triple Negative Paradox: Primary Tumor Chemosensitivity of Breast Cancer Subtypes. *Clin Cancer Res* **13:** 2329-2334.

Clarke MF, Dick JE, Dirks PB, Eaves CJ, Jamieson CHM, Jones DL *et al* (2006). Cancer Stem Cells--Perspectives on Current Status and Future Directions: AACR Workshop on Cancer Stem Cells. *Cancer Res* **66:** 9339-9344.

Clarke RB (2005). Isolation and characterization of human mammary stem cells. *Cell Proliferation* **38:** 375-386.

Davidson CJ, Tirouvanziam R, Herzenberg LA, Lipsick JS (2005). Functional Evolution of the Vertebrate Myb Gene Family: B-Myb, but Neither A-Myb nor c-Myb, Complements Drosophila Myb in Hemocytes. *Genetics* **169:** 215-229.

Dontu G, El-Ashry D, Wicha MS (2004). Breast cancer, stem/progenitor cells and the estrogen receptor. *TRENDS in Endocrinology and Metabolism* **Vol.15**.

Ginestier C, Hur MH, Charafe-Jauffret E, Monville F, Dutcher J, Brown M *et al* (2007). ALDH1 Is a Marker of Normal and Malignant Human Mammary Stem Cells and a Predictor of Poor Clinical Outcome. *Cell Stem Cell* **1:** 555-567.

Guérin M SZ, Andrieu N, Riou G. (1990). Strong association between c-myb and oestrogen-receptor expression in human breast cancer. *Oncogene* **5:** 131-5.

Hennessy BT, Gonzalez-Angulo A-M, Stemke-Hale K, Gilcrease MZ, Krishnamurthy S, Lee J-S *et al* (2009). Characterization of a Naturally Occurring Breast Cancer Subset Enriched in Epithelial-to-Mesenchymal Transition and Stem Cell Characteristics. *Cancer Res* **69:** 4116-4124.

Herschkowitz J, Simin K, Weigman V, Mikaelian I, Usary J, Hu Z *et al* (2007). Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. *Genome Biology* **8:** R76.

Howard BA, Gusterson BA (2000). Human Breast Development. *Journal of Mammary Gland Biology and Neoplasia* **V5:** 119.

Hu Z, Fan C, Oh D, Marron J, He X, Qaqish B *et al* (2006). The molecular portraits of breast tumors are conserved across microarray platforms. *BMC Genomics* **7:** 96.

Ito M, Araki S, Matsunaga S, Itoh T, Nishihama R, Machida Y *et al* (2001). G2/M-Phase-Specific Transcription during the Plant Cell Cycle Is Mediated by c-Myb-Like Transcription Factors. *Plant Cell* **13:** 1891-1905.

Jordan CT, Guzman ML, Noble M (2006). Cancer Stem Cells. *N Engl J Med* **355:** 1253-1261.

Kauraniemi P, Hedenfalk I, Persson K, Duggan DJ, Tanner M, Johannsson O *et al* (2000). MYB Oncogene Amplification in Hereditary BRCA1 Breast Cancer. *Cancer Res* **60:** 5323-5328.

Klempnauer K-H, Gonda TJ, Bishop MJ (1982). Nucleotide sequence of the retroviral leukemia gene v-myb and its cellular progenitor c-myb: The architecture of a transduced oncogene. *Cell* **31:** 453-463.

Lacroix M, Leclercq G (2004). About GATA3, HNF3A, and XBP1, three genes co-expressed with the oestrogen receptor-alpha gene (ESR1) in breast cancer. *Mol Cell Endocrinol* **219:** 1 - 7.

Leprince D, Gegonne A, Coll J, de Taisne C, Schneeberger A, Lagrou C *et al* (1983). A putative second cell-derived oncogene of the avian leukaemia retrovirus E26. *Nature* **306:** 395.

Lim E, Vaillant F, Wu D, Forrest NC, Pal B, Hart AH *et al* (2009). Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. *Nat Med* **15:** 907-913.

Livasy CA, Karaca G, Nanda R, Tretiakova MS, Olopade OI, Moore DT *et al* (2005). Phenotypic evaluation of the basal-like subtype of invasive breast carcinoma. *Mod Pathol* **19:** 264.

Malaterre J, Carpinelli M, Ernst M, Alexander W, Cooke M, Sutton S *et al* (2007). c-Myb is required for progenitor cell homeostasis in colonic crypts. *PNAS* **104:** 3829-3834.

Maurice D, Joel Hooper, Georgina Lang and Kathleen Weston (2007). c-Myb regulates lineage choice in developing thymocytes via its target gene Gata3. *The EMBO Journal*.

Mucenski ML, McLain K, Kier AB, Swerdlow SH, Schreiner CM, Miller TA *et al* (1991). A functional c-myb gene is required for normal murine fetal hepatic hematopoiesis. *Cell* **65:** 677.

Ness SA (2003). Myb protein specificity: evidence of a context-specific transcription factor code. *Blood Cells, Molecules, and Diseases* **31:** 192.

Nomura N, Takahashi M, Matsui M, Ishii S, Date T, Sasamoto S *et al* (1988). Isolation of human cDNA clones of myb-related genes, A-myb and B-myb. *Nucl. Acids Res.* **16:** 11075-11089.

Oh IH, Reddy EP (1999). The *myb* gene family in cell growth, differentiation and apoptosis. *Oncogene* **18:** 3017-33.

Parker JS, Mullins M, Cheang MCU, Leung S, Voduc D, Vickery T *et al* (2009). Supervised Risk Predictor of Breast Cancer Based on Intrinsic Subtypes. *J Clin Oncol* **27:** 1160-1167.

Parmar H, Young P, Emerman JT, Neve RM, Dairkee S, Cunha GR (2002). A Novel Method for Growing Human Breast Epithelium in Vivo Using Mouse and Human Mammary Fibroblasts. *Endocrinology* **143:** 4886-4896.

Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA *et al* (2000). Molecular portraits of human breast tumours. *Nature* **406:** 747.

Persson M, Andren Y, Mark J, Horlings HM, Persson F, Stenman G (2009). Recurrent fusion of MYB and NFIB transcription factor genes in carcinomas of the breast and head and neck. *Proceedings of the National Academy of Sciences***:** -.

Pomeroy SL, Tamayo P, Gaasenbeek M, Sturla LM, Angelo M, McLaughlin ME *et al* (2002). Prediction of central nervous system embryonal tumour outcome based on gene expression. *Nature* **415:** 436.

Prat A, Perou CM (2009). Mammary development meets cancer genomics. *Nat Med* **15:** 842-844.

Reya T, Morrison SJ, Clarke MF, Weissman IL (2001). Stem cells, cancer, and cancer stem cells. *Nature* **414:** 105.

Rosinski JA, Atchley WR (1998). Molecular Evolution of the Myb Family of Transcription Factors: Evidence for Polyphyletic Origin. *Journal of Molecular Evolution* **46**.

Russo J, Russo IH (2004). Development of the human breast. *Maturitas* **49:** 2-15.

Sala A (2005). B-MYB, a transcription factor implicated in regulating cell cycle, apoptosis and cancer. *European Journal of Cancer* **41:** 2479-2484.

Sauer T (2007). Cytologic findings in malignant myoepithelioma: a case report and review of the literature. *CytoJournal* **4**.

Shipitsin M, Campbell LL, Argani P, Weremowicz S, Bloushtain-Qimron N, Yao J *et al* (2007). Molecular Definition of Breast Tumor Heterogeneity. *Cancer Cell* **11:** 259.

Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A *et al* (2003). Repeated observation of breast tumor subtypes in independent gene expression data sets. *PNAS* **100:** 8418-8423.

Tanaka Y, Patestos NP, Maekawa T, Ishii S (1999). B-myb is Required for Inner Cell Mass Formation at an Early Stage of Development. *J. Biol. Chem.* **274:** 28067-28070.

Thiele CJ, Cohen PS, Israel MA (1988). Regulation of c-myb expression in human neuroblastoma cells during retinoic acid-induced differentiation. *Mol. Cell. Biol.* **8:** 1677-1683.

Toscani A, Mettus RV, Coupland R, Simpkins H, Litvin J, Orth J *et al* (1997). Arrest of spermatogenesis and defective breast development in mice lacking A-myb. *Nature* **386:** 713.

Trauth K, Mutschler B, Jenkins N, Gilbert D, Copeland N, Klempnauer K (1994). Mouse A-myb encodes a trans-activator and is expressed in mitotically active cells of the developing central nervous system, adult testis and B lymphocytes. *The EMBO Journal* **13:** 5994–6005.

Villadsen R (2005). In search of a stem cell hierarchy in the human breast and its relevance to breast cancer evolution. *APMIS* **113:** 903-921.

Villadsen R, Fridriksdottir AJ, Ronnov-Jessen L, Gudjonsson T, Rank F, LaBarge MA *et al* (2007). Evidence for a stem cell hierarchy in the adult human breast. *J. Cell Biol.* **177:** 87-101.

Weston K (1998). Myb proteins in life, death and differentiation. *Current Opinion in Genetics & Development* **8:** 76.

Whitfield ML, George LK, Grant GD, Perou CM (2006). Common markers of proliferation. *Nat Rev Cancer* **6:** 99-106.

Wicha MS, Liu S, Dontu G (2006). Cancer Stem Cells: An Old Idea--A Paradigm Shift. *Cancer Res* **66:** 1883-1890.

Woodward WA, Chen MS, Behbod F, Rosen JM (2005). On mammary stem cells. *J Cell Sci* **118:** 3585-3594.

CHAPTER II

IN VITRO AND *IN VIVO* ANALYSIS OF B-MYB IN BASAL-LIKE BREAST CANCER

PREFACE

This work was previously published. My role in this project included initiating, organizing, and designing the study. I performed the microarray data analysis, *B-Myb* over-expression in mammary cell lines experiments including cell line chemotherapy treatments, and flow cytometry assays. I initiated all aspects of manuscript preparation, including writing, figure preparation, and revisions. Katherine A. Hoadley aided in study design and contributed writing of the manuscript. Joel S. Parker assisted with microarray analysis. Robert C. Millikan and Scott Winkel were essential in the collection and analysis of the Carolina Breast Cancer Study patient samples. Charles M. Perou was the principal investigator, conceived the project, and contributed to writing the manuscript.

Thorner AR, Hoadley KA, Parker JS, Winkel S, Millikan RC, Perou CM (2009). In vitro and in vivo analysis of B-Myb in basal-like breast cancer. Oncogene **28:** 742-51.

ABSTRACT

A defining feature of basal-like breast cancer, a breast cancer subtype with poor clinical prognosis, is the high expression of "proliferation signature" genes. We identified *B-Myb*, a MYB family transcription factor that is often amplified and overexpressed in many tumor types, as being highly expressed in the proliferation signature. However, the roles of B-Myb in disease progression, and its mammary-specific transcriptional targets, are poorly understood. Here, we demonstrated that *B-Myb* expression is a significant predictor of survival and pathological complete response to neoadjuvant chemotherapy in breast cancer patients. We also identified a significant association between the G/G genotype of a nonsynonymous *B-Myb* germline variant (rs2070235, S427G) and an increased risk of basal-like breast cancer [OR 2.0, 95% CI (1.1-3.8)]. In immortalized, human mammary epithelial cell lines, but not basal-like tumor lines, cells ectopically expressing wild-type *B-Myb* or the S427G variant showed increased sensitivity to two DNA topoisomerase IIα inhibitors, but not to other chemotherapeutics. In addition, microarray analyses identified many G2/M genes as being induced in *B-Myb* overexpressing cells. These results confirm that B-Myb is involved in cell cycle control, and that dysregulation of *B-Myb* may contribute to increased sensitivity to a specific class of chemotherapeutic agents. These data provide insight into the influence of *B-Myb* in human breast cancer, which is of potential clinical importance for determining disease risk and for guiding treatment.

INTRODUCTION

Breast cancer is not one disease, but rather, represents at least six subtypes (Hu *et al.*, 2006; Perou *et al.*, 2000; Sorlie *et al.*, 2003). These include luminal A, luminal B, normal-like, HER2-enriched, claudin-low and basal-like, each with unique gene expression profiles and distinct clinical outcomes. The majority of breast cancer cases (60-80 percent) comprise the luminal/estrogen receptor-α positive (ER+) tumors, while basal-like breast carcinoma accounts for 10-15 percent of all cases. The basal-like subtype, often clinically observed as "triple-negative" tumors (negative for ERα, PR, and HER2), is of particular interest because treatment options are limited to chemotherapy only and patients with this disease typically have poor outcomes. One contribution to the poor outcome of basal-like breast cancer patients may be their high cellular proliferation rates, which is manifested transcriptionally as the high expression of the "proliferation signature"; this is a dominant gene signature that is a marker of cell proliferation rates across multiple tumor types (Hu *et al.*, 2006; Whitfield *et al.*, 2006). *B-Myb*, a gene with known cell cycle control functions and implications in tumorigenesis (Sala, 2005), is one of approximately 100 genes that define the proliferation signature.

B-Myb is a member of the vertebrate MYB family of nuclear transcription factors. In humans this family is comprised of *A-Myb* (MYBL1), *B-Myb* (MYBL2), and *c-Myb* (MYB). Each family member is able to recognize and bind to the same DNA consensus sequence (PyAAC(G/T)G) to promote gene transcription; however, varying tissue-specific expression patterns, as well as protein-protein interactions with unique co-factors, suggests that distinct biological roles exist for each MYB family member (Rosinski and Atchley, 1998; Sala, 2005). Found in the genomes of both plants and animals, MYB proteins are conserved

throughout evolution and control processes from flavonoid production to cellular proliferation (Ito *et al.*, 2001; Rosinski and Atchley, 1998). In contrast to vertebrates, invertebrates contain only one MYB protein, which in Drosophila (dMYB) is phylogenetically and functionally complementary to vertebrate B-Myb, suggesting B-Myb to be the most ancient family member (Davidson *et al.*, 2005). The expression of *B-Myb*, unlike *c-Myb* and *A-Myb* (Malaterre *et al.*, 2007; Mucenski *et al.*, 1991; Ness, 2003; Toscani *et al.*, 1997; Trauth *et al.*, 1994), is ubiquitously expressed in virtually all proliferating cells as a regulator of cell cycle progression and plays an essential role in vertebrate development; knocking out murine *B-Myb* causes early embryonic lethality (E4.5-6.5) resulting from unsuccessful inner cell mass formation (Tanaka *et al.*, 1999).

MYB family members have been implicated in tumorigenesis for several decades. The *c-Myb* proto-oncogene was first identified as the mammalian homolog of *v-myb*, which is the transforming gene transmitted by the avian myeloblastosis and E26 retroviruses causing acute leukemia in birds (Klempnauer *et al.*, 1982; Leprince *et al.*, 1983). *A-* and *B-Myb* were later discovered during low stringency screening of human cDNA libraries (Nomura *et al.*, 1988). The *B-Myb* chromosomal locus, 20q13, is amplified and/or highly expressed in a variety of tumor types including breast, prostate, liver and ovarian carcinomas, and in most cases this high expression portends a poor prognosis (Sala, 2005). *B-Myb* is also an important marker of poor outcome in embryonal tumors of the central nervous system (CNS) (Pomeroy *et al.*, 2002). Recently, a nonsynonymous *B-Myb* germline variant (rs2070235) causing a serine to glycine amino acid change (S427G) was linked to a decrease in overall cancer risk for neuroblastomas, chronic myelogenous leukemia, and colon cancers in a combined dataset of cases and controls (Schwab *et al.*, 2007). However, the molecular

roles of *B-Myb* in disease progression, as well as its transcriptional target genes in the mammary gland, are still poorly understood. To gain insight into *B-Myb* and its involvement in breast cancer, we analyzed the expression of *B-Myb* across the breast cancer subtypes, examined its relationship to survival and pathological complete response and the correlation of variant rs2070235 to disease risk. We also manipulated the expression of *B-Myb* and the S427G variant in normal and tumor derived mammary cell lines and observed alterations in drug sensitivity and cell cycle profiles.

MATERIALS AND METHODS

Cell Lines. *hTERT*-immortalized, human mammary epithelial cell lines (HME-CC, ME16C), and basal-like breast cancer-derived lines (SUM102, SUM149) were cultured as described (Hoadley *et al.*, 2007; Troester *et al.*, 2004). Full-length, human *B-Myb* cDNA (GenBank NM_002466) was cloned into the pBabe.puro.GWrfA (Gateway Reading Frame A) vector using Gateway® Cloning Technology (Invitrogen, Carlsbad, CA, USA). Retrovirus was produced in Phoenix 293T cells by transfecting with ten micrograms of vector using Lipofectamine 2000 (Invitrogen), as per manufacturer's instructions. Media was changed 24 hours post-transfection and supernatants collected 12 hours later. Seventy-five micrograms of polybrene were added to the collected supernatants and applied to the mammary cell lines. Stable populations were selected by culturing in 1 ug/mL puromycin for HME-CC and ME16C, or 0.5 ug/mL puromycin for SUM102 and SUM149.

B-Myb variant (S427G) was created in the pBabe.puro.B-Myb expression vector using QuikChange® XL Site-Directed Mutagenesis Kit (Stratagene, Los Angeles, CA, USA).

Western Blot Analysis. Cells were grown in 10 cm tissue culture-treated dishes until 80 percent confluence, followed by harvest, protein isolation and quantification as previously described (Troester *et al.*, 2004). Membranes were probed for B-Myb (sc-725; Santa Cruz Biotechnology, Inc., Santa Cruz, CA, USA), and β-actin (AC-15; Abcam, Cambridge, MA, USA), followed by anti-rabbit or anti-mouse IgG horseradish peroxidase-linked whole antibody (Amersham Biosciences) and detected using SuperSignal West Pico Chemiluminescent Substrate (Pierce, Rockford, IL, USA).

Cell Cycle Analysis. Cells were treated with a range of doxorubicin (0-70 nM) or etoposide (0-2 nM) doses and DNA content was analyzed using a modified propidium iodide staining assay. Briefly, one million cells were collected by trypsinization, washed in 1x PBS, and fixed in 70% ethanol at 4C overnight. Cells were washed with PBS/0.2% BSA and resuspended in 500 microliters of PBS/0.2% BSA and 100 micrograms RNaseA. Propidium iodide was added to a final volume of 50 micrograms, and cells incubated at 37C for 30 minutes. DNA content analysis was performed using a FACScan flow cytometer (Becton Dickinson, Franklin Lakes, NJ, USA). To determine fractions of cells in G1 and G2/M phases, histograms of DNA content were analyzed using Summit v4.3 software (Dako, Carpinteria, CA, USA) gating around 2N and 4N DNA content. Analyses were done on three separate days and mean values were calculated.

Cytotoxicity Assay. Cell line sensitivities to drugs were assessed using a modified mitochondrial dye conversion assay (Cell-Titer 96, Promega #G4100, Madison, WI, USA) as described (Hoadley *et al.*, 2007; Troester *et al.*, 2004). Chemotherapeutics (carboplatin, doxorubicin, 5-fluorouracil, paclitaxel, etoposide, camptothecin) were purchased from Sigma (St. Louis, MO, USA). The 72-hour inhibitory concentration that caused a 50% reduction in MTT dye conversion (IC50) was determined using nonlinear regression (SAS Statistical Software, Cary, NC, USA) (Vanewijk and Hoekstra, 1993). Differences in the IC50 estimates were tested by a traditional ANOVA test of nested models was performed using the R system for statistical computing (R Development Core Team, 2006 http://www.R-project.org).

Microarray Analysis. Five replicates each of HME-CC *B-Myb*-overexpressing and vector control cell lines were treated with the 72-hour IC50 dose of doxorubicin. Poly-A(+) RNA was collected (Micro-FastTrack2.0 mRNA Isolation Kit, Invitrogen) from treated control cells and *B-Myb* overexpressing lines, reverse transcribed and labeled using the Agilent Low RNA Input Linear Amplification Kit (Agilent Technologies, Santa Clara, CA, USA), and hybridized to Agilent Human 44K Custom Oligo microarrays as described (Hu *et al.*, 2005). An untreated HME-CC cell line reference was co-hybridized to all arrays (e.g. untreated HME-CC vs. doxorubicin-treated empty vector HME-CC). Microarrays were scanned on an Axon Genepix 4000B microarray scanner and analyzed using GenePix Pro 5.1 software (Molecular Devices, Sunnyvale, CA, USA). Data was normalized using Lowess normalization on the Cy3 and Cy5 channels. Microarray data is available at the UNC

Microarray Database [http://genome.unc.edu] and at the Gene Expression Omnibus (GSE11429).

Microarray Statistical Analyses. Supervised microarray analysis was performed by selecting genes with an absolute signal intensity of at least 30 units in both dye channels and data present in at least 70% of experimental samples. A two-class, unpaired Significance Analysis of Microarrays (SAM) was performed to identify significant genes associated with *B-Myb* expression with a false discovery rate (FDR) of less than 3% (Tusher *et al.*, 2001).

The Netherlands Cancer Institute breast cancer dataset (NKI-295, n=295) was used for analysis of *B-Myb* expression across breast cancer subtypes (van de Vijver *et al.*, 2002); however, only locally-treated tumors (no chemotherapy) were used in survival analyses (n=165). Association of *B-myb* expression states (rank ordered and split in halves: low/high) relative to survival was tested using the Cox-Mantel log-rank test and results visualized using Kaplan-Meier survival plots (WinSTAT v.2007.1). Testing the association of *B-myb* expression versus subtypes was performed using ANOVA. Three other published datasets were analyzed, as above, for survival (Miller *et al.*, 2005; Wang *et al.*, 2005) or pathological response (Hess *et al.*, 2006) by chi-square using the R system for statistical computing.

Carolina Breast Cancer Study. The Carolina Breast Cancer Study (CBCS) is a population-based, case-control study of breast cancer conducted in 24 counties of central and eastern North Carolina between 1993 and 2001 (Millikan *et al.*, 2003; Newman *et al.*, 1995). Incident cases of primary invasive and in situ breast cancer were identified using the North

Carolina Central Cancer Registry, with over-sampling of African-American and younger women. Controls were frequency-matched to cases based upon age (± 5 years) and self-reported race using randomized recruitment. Cases were subtyped using a panel of immunohistochemistry markers that identified the "intrinsic" breast cancer subtypes (Carey et al., 2006; Millikan et al., 2008). Definitions of subtypes were luminal A (ER+ PR+ HER2-), luminal B (ER+PR+HER2+), basal-like (ER-PR-HER2-, ck5/6+ and/or HER1+), HER2+ (ER-PR-HER2+), and unclassified (negative for all five markers). Procedures for recruiting and enrolling study participants and obtaining biologic samples were approved by the Institutional Review Board of the UNC School of Medicine, and informed consent was obtained from each participant.

Genotyping of B-Myb. Genotyping for the B-Myb polymorphism at codon 427, Serine (AGC) to Glycine (GGC) (rs2070235), was conducted using DNA extracted from peripheral blood lymphocytes. The ABI 7700 Sequence Detection System, or "Taqman"™ assay (Applied Biosystems, ABI, Foster City, CA, USA) was employed, with polymerase chain reaction (PCR) primers and probes designed using Primer Express TM software (ABI) and assay conditions based on the allelic discrimination protocol from Applied Biosystems. PCR primer and probe sequences are available upon request. Out of a total of 3862 samples genotyped, the failure rate was 0.4%, and complete agreement was obtained on a 10% random sample.

There were no significant departures from Hardy-Weinberg equilibrium among African-American cases (p=0.80), African-American controls (p=0.77), Caucasian cases (p=0.27), or Caucasian controls (p=0.51).

Cell line genotyping was performed by Polymorphic DNA Technologies, Inc (Alameda, CA, USA).

Genotype Statistical Analysis. B-Myb genotypes were determined for 1256 cases with subtype information (500 African-American, 756 Caucasian) and 1814 controls (679 African-American, 1135 Caucasian). Odds ratios (ORs) and 95% Confidence Intervals (CIs) for each breast cancer subtype versus all controls were calculated using unconditional logistic regression. The PROC GENMOD statement in SAS (version 8.2; SAS Institute, Cary, NC, USA) was used to incorporate offsets derived from sampling probabilities used to identify eligible study participants, race (African-American, Caucasian), and age (11-level ordinal variable representing 5-year age categories). Trend tests were conducted by calculating the p-value for B-Myb genotype coded as an ordinal variable. Likelihood ratio tests (LRTs) were used to test for modification of ORs by race.

RESULTS

High *B-Myb* Expression in Breast Tumors Predicts Poor Outcome. To assess the relevance of *B-Myb* gene expression across the breast cancer subtypes, breast tumor microarray data from the Netherlands Cancer Institute (NKI-295, n=295, (van de Vijver *et al.*, 2002)) was analyzed. Tumor samples were classified into five breast cancer subtypes using a single sample centroid-based predictor as described (Hu *et al.*, 2006). An ANOVA analysis performed on these stratified samples showed that *B-Myb* expression differs significantly across the subtypes and was highest in basal-like tumors (Figure 2.1).

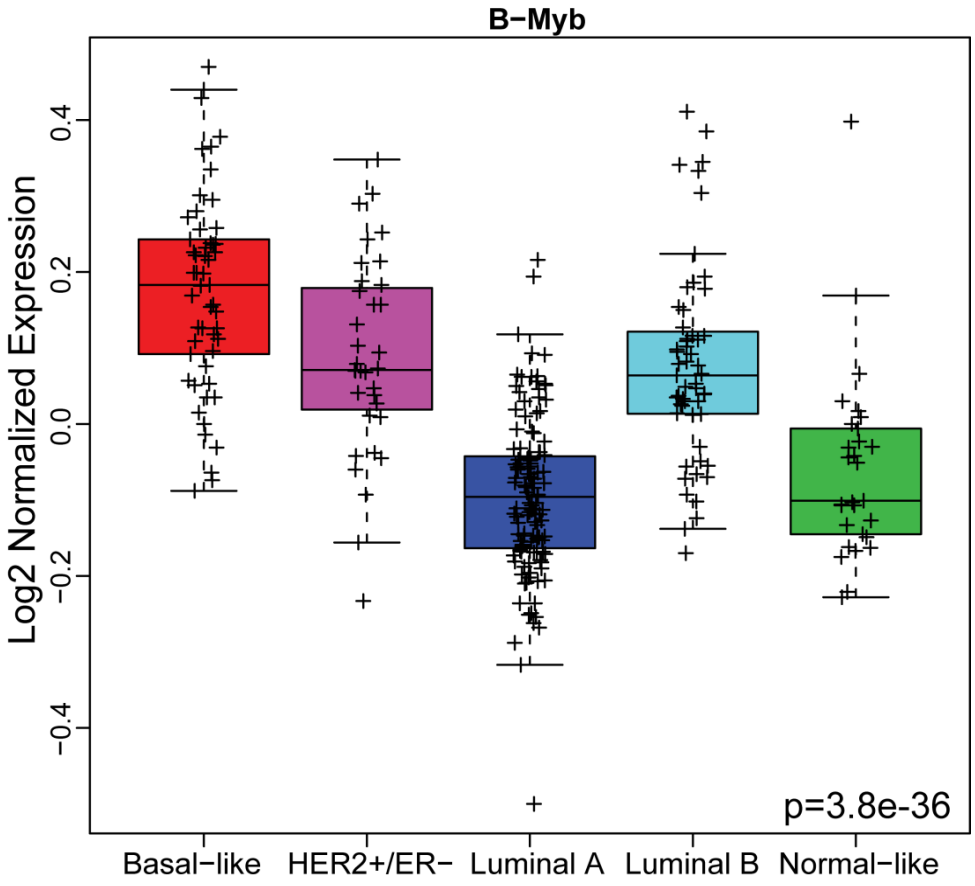


Figure 2.1. *B-Myb* expression across breast cancer subtypes. The NKI breast tumor microarray dataset (n=295) was classified into the five intrinsic subtypes and box plots used to visualize *B-Myb* expression according to breast cancer subtypes. Statistical significance was calculated by ANOVA.

To test for correlations between *B-Myb* mRNA expression alone and patient outcome, we analyzed the NKI patients not receiving adjuvant systemic treatment (i.e. local treatment only; n=165). This allowed us to better identify the prognostic abilities of *B-Myb* without the confounding data of treatment response. The NKI "local-only" tumors were rank ordered into halves (low/high) based on their *B-Myb* expression levels and analyzed for overall survival (OS) and relapse free survival (RFS) by Kaplan-Meier analysis. Poor OS and RFS were highly correlated ($p<0.001$) with *B-Myb* high expression levels in these NKI samples (Figure 2.2A, and RFS data not shown). *B-Myb* expression alone was also able to significantly predict OS on local-only treated luminal A subtype tumors (n=72) (Figure 2.2B), luminal B (n=26) (Figure 2.2C), HER2+/ER- (n=21) (Figure 2.2D), but not basal-like tumors (n=30) (Figure 2.3A). We then evaluated the prognostic ability of *B-Myb* using two other published breast tumor microarray datasets (Miller *et al.*, 2005; Wang *et al.*, 2005). Wang *et al.*, 2005 (n=286) consisted of microarrays on untreated, lymph-node-negative primary ER+ and ER- breast cancers with relapse data, and *B-Myb* was capable of predicting RFS in these patients (Figure 2.2E). On this same dataset, *B-Myb* also predicted RFS in the ER+ patient subset (n=209), but not the ER- subset (n=77) (Figures 2.3B, C). Another dataset consisting of primary invasive tumors (Miller *et al.*, 2005) (n=234) was tested and similar results were found (Figure 2.2F).

To determine if *B-Myb* expression was involved with pathologic complete response (pCR), we used the data of Hess *et al.*, 2006, where microarrays were performed on pre-treatment breast tumors from patients receiving neoadjuvant paclitaxel, followed by 5FU-Adriamycin-Cyclophosphamide (T/FAC; n=133). Again, samples were split into two groups

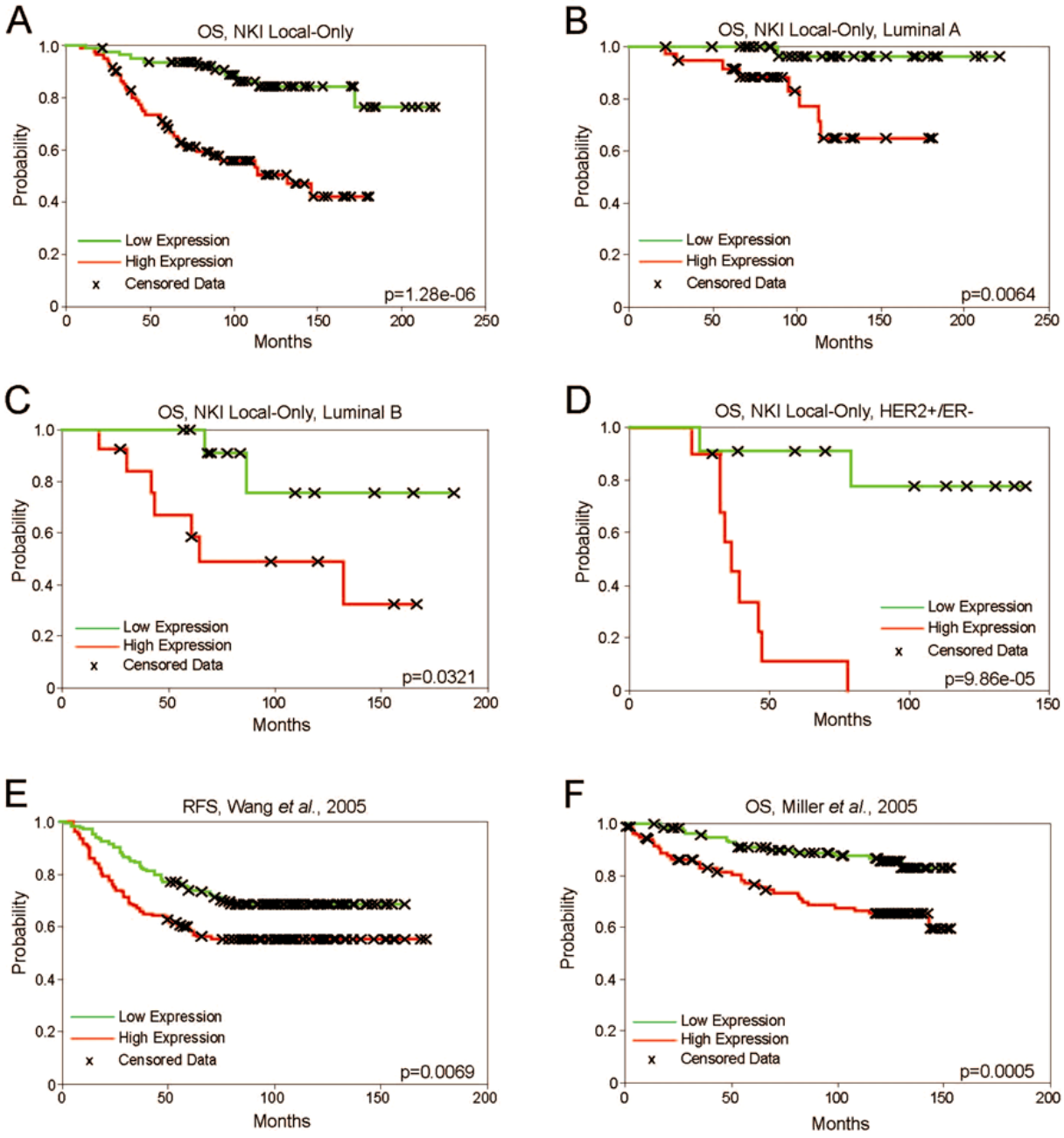


Figure 2.2. High expression of B-Myb correlates with poor sruvival. Kaplan-Meier survival analyses based on B-Myb expression values rank ordered into halves (low/high). (A-D) Overall survival (OS) of locally treated NKI tumor samples: (A) All subtypes combined (n=165), (B) Luminal A (n=72), (C) Luminal B (n=26), (D) HER2+/ER- (n=21). (E) RFS, Wang *et al*., 2005 (n=286), a locally treated, lymph-node-negative tumor microarray dataset. (F) Miller *et al*., 2005 breast tumor microarray dataset (n=234).

Figure 2.3.

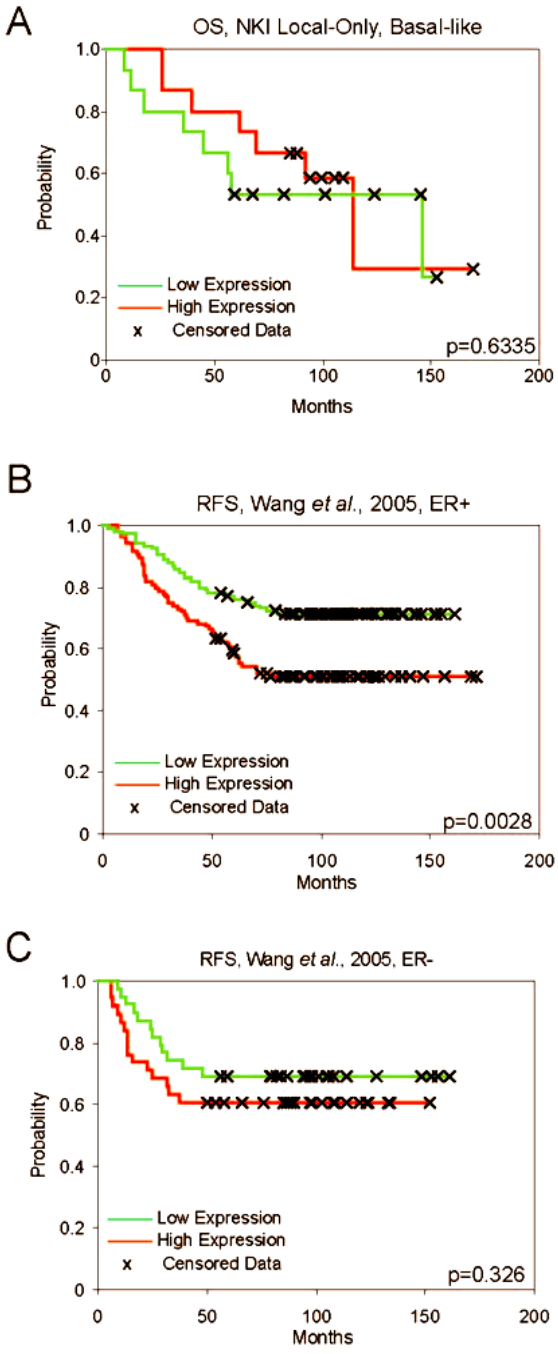

A
OS, NKI Local-Only, Basal-like
Probability
1.0
0.8
0.6
0.4
0.2
0
Low Expression
High Expression
Censored Data
p=0.6335
0
50
100
150
200
Months
B
RFS, Wang et al., 2005, ER+
Probability
1.0
0.8
0.6
0.4
0.2
0
Low Expression
High Expression
Censored Data
p=0.0028
0
50
100
150
200
Months
C
RFS, Wang et al., 2005, ER-
Probability
1.0
0.8
0.6
0.4
0.2
0
Low Expression
High Expression
Censored Data
p=0.326
0
50
100
150
200
Months

Figure 2.3. High expression of B-Myb correlates with poor outcome. Kaplan-Meier survival analysis based on B-Myb expression values rank ordered and split into halves (low/high). (A) NKI, locally treated, basal-like (n=30). (B) Wang *et al*., 2005, ER+ samples (n=209). (C) Wang *et al*., 2005, ER- samples (n=77).

based on *B-Myb* expression (low/high). *B-Myb* high expression was again associated subtype (data not shown) and with pCR, as calculated by chi-square test (p=0.008; Table 2.1).

Table 2.1. High *B-Myb* expression in breast tumors (n=133; Hess *et al.*, 2006) is associated with pathological complete response (pCR).

***B-Myb* Expression**	**No pCR**	**pCR**
High	42	24
Low	57	10

χ^2=6.9433, df=1, p-value=0.008.

***B-Myb* Germline Variant (rs2070235) Increases Risk of Basal-Like Breast Cancer.** A nonsysnonymous germline *B-Myb* variant exists that causes a serine to glycine substitution (rs2070235, S427G). This non-conservative change prompted us to look for correlations between this variant and baseline susceptibility risk in the population-based Carolina Breast Cancer Study/CBCS (Newman *et al.*, 1995). Odds ratios for the *B-Myb* genotype and all breast cancer cases, luminal A, and basal-like cases, versus controls are presented in Table 2.2. There was no association between *B-Myb* genotype and all breast cancer (p=0.71); however a statistically significant association was observed for basal-like breast cancer (p=0.047), but not luminal A (p=0.14). No association was observed for *B-Myb* genotype and the other breast cancer subtypes (luminal B, HER2+/ER- and unclassified: data not shown). Odds ratios were similar in African-Americans and Caucasians (likelihood-ratio tests/LRTs for interaction with race were not statistically significant). Among controls in the CBCS, allele frequencies for the *B-Myb* G allele were higher in African-Americans (0.27) than Caucasians (0.08).

Table 2.2. *B-Myb* variant genotype (G/G) increases the risk of basal-like breast cancer.

B-Myb Genotype	Controls (N=1814)	Basal-like cases (N=206)	OR (95% CI)[a]	Luminal A cases (N=698)	OR (95% CI)[a]	All cases (N=1256)	OR (95% CI)[a]
A/A	1319	131	Referent	531	Referent	912	Referent
A/G	436	60	1.2 (0.8-1.7)	149	0.8 (0.7-1.1)	301	1.0 (0.8-1.2)
G/G	59	15	**2.0 (1.1-3.8)**	18	0.8 (0.5-1.4)	43	1.0 (0.6-1.5)
Trend test			**p=0.047**		p=0.14		p=0.71

Abbreviations: A, wild-type; G, variant; OR, odds ratio. Odds ratios for the *B-Myb* genotype, and basal-like, luminal A and all breast cancer cases. Patient saples were genotyped for the *B-Myb* polymorphism at codon 427, rs2070235.
[a]Adjusted for offsets, age and race.

Ectopic Expression of *B-Myb* Increases Sensitivity to TOP2A Inhibitors. Given *B-Myb*'s expression within the proliferation signature, evidence suggesting that rapidly growing tumors may be more chemotherapy sensitive, and the correlation between *B-Myb* expression and pathological complete response, we sought to determine if ectopic expression of *B-Myb* *in vitro* had an effect on sensitivity to chemotherapeutics. *B-Myb*, and the *B-Myb* S427G variant, were overexpressed in two *hTERT*-immortalized human mammary epithelial cell lines (HME-CC, ME16C) and two basal-like tumor derived lines (SUM102, SUM149) (Figure 2.4). Low endogenous B-Myb levels were detectable by western blot in the tumor lines, and in all lines by microarray analysis for mRNA levels (Figure 2.4 and data not shown). All cell lines were also genotyped for rs2070235 and identified as homozygous for the major allele. It is important to note that the two normal tissue derived cell lines have a basal-like phenotype when assessed by gene expression analysis (Troester *et al.*, 2004), thus, these lines represent appropriate counterparts to the two basal-like tumor lines.

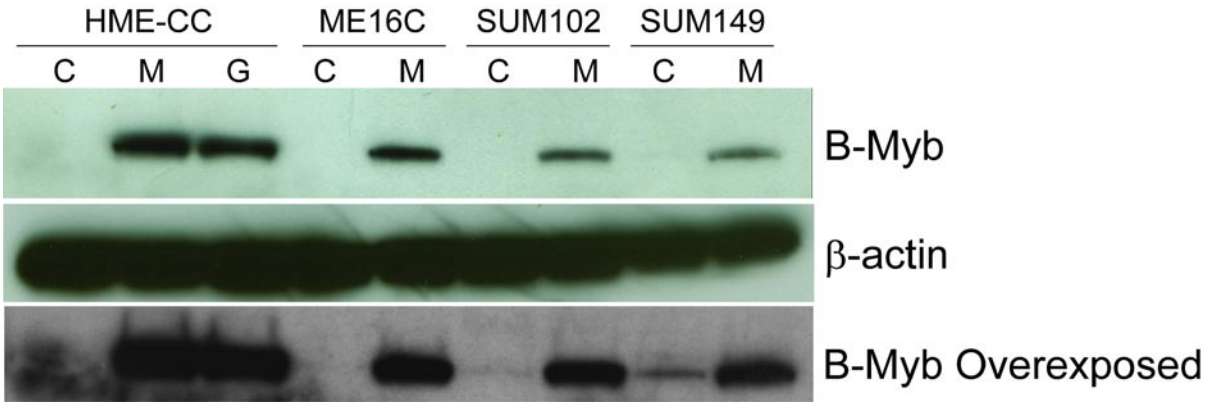


Figure 2.4. Western blot analysis of ectopic expression of *B-Myb* or *B-Myb* variant in basal-like mammary cell lines. Two *hTERT*-immortalized HMEC lines (HME-CC and ME16C), and two basal-like tumor derived lines (SUM102 and SUM149) were transduced with an empty vector control (C) or B-Myb (M) cDNA expression construct. For one cell line (HME-CC), the B-Myb S427G variant (G) was also introduced and analyzed by immunoblotting. To visualize any detectable endogenous B-Myb protein, the B-Myb blot was overexposed ("B-Myb Overexposed"; >30 minutes).

Cells ectopically expressing *B-Myb*, or S427G variant, were treated with a panel of chemotherapy agents including two DNA topoisomerase IIα (TOP2A) inhibitors (doxorubicin, etoposide), a DNA topoisomerase I inhibitor (camptothecin), a microtubule stabilizer (paclitaxel), a DNA alkylating agent (carboplatin), and an antimetabolite (5-flurouracil), most of which are commonly used in breast cancer treatment. HME-CC cells overexpressing *B-Myb*, or variant, were approximately twice as sensitive to the two DNA topoisomerase IIα inhibitors, based on IC50 assays, compared to the parental cell lines, but showed no significant change in sensitivity to the other drugs (Figure 2.5A). To further test *B-Myb* effects on chemotherapy sensitivity, another immortalized HMEC line (ME16C) and two basal-like tumor derived cell lines (SUM102, SUM149) were tested. The ME16C cell

line was also sensitized to TOP2A inhibitors by *B-Myb* expression, but not to treatment with other chemotherapeutics (Figure 2.5B and data not shown); however, this sensitivity profile was not observed in either of the two basal-like tumor derived lines (Figures 2.5C and D), where *B-Myb* expression had no apparent effect. The *B-Myb* S427G variant was also tested in each cell line for chemosensitivity and behaved similarly to the cells overexpressing wild-type *B-Myb* (Figure 2.5A and data not shown).

Gene Expression Analysis of Cell Lines Ectopically Expressing *B-Myb*. To look for gene expression changes in cells overexpressing *B-Myb*, and to further assess the chemotherapy phenotype, microarrays were performed on the HME-CC cell lines. Under normal, non-confluent conditions, the only statistically significant expression difference (Significance Analysis of Microarrays/SAM analysis (Tusher *et al.*, 2001) with a 3% false discovery rate/FDR) was *B-Myb* itself; also, no gene expression differences were observed between the *B-Myb* and *B-Myb* S427G-expressing cell lines. Since a chemotherapy-related phenotype was observed with TOP2A inhibitors, we tested the cell lines after treatment with the 72 hour IC50 dose of doxorubicin. In a two-class SAM analysis (i.e. doxorubicin-treated HME-CC control vs. doxorubicin-treated HME-CC+*B-Myb*), 217 genes were identified (FDR <3%; Appendix II).

An EASE (Expression Analysis Systematic Explorer) analysis (Hosack *et al.*, 2003) was performed and many cell cycle related gene ontology categories were identified as being significantly enriched (Table 2.3). Therefore, in doxorubicin treated cells overexpressing *B-Myb* there was significantly higher expression of many cell cycle genes compared to the treated control.

Figure 2.5

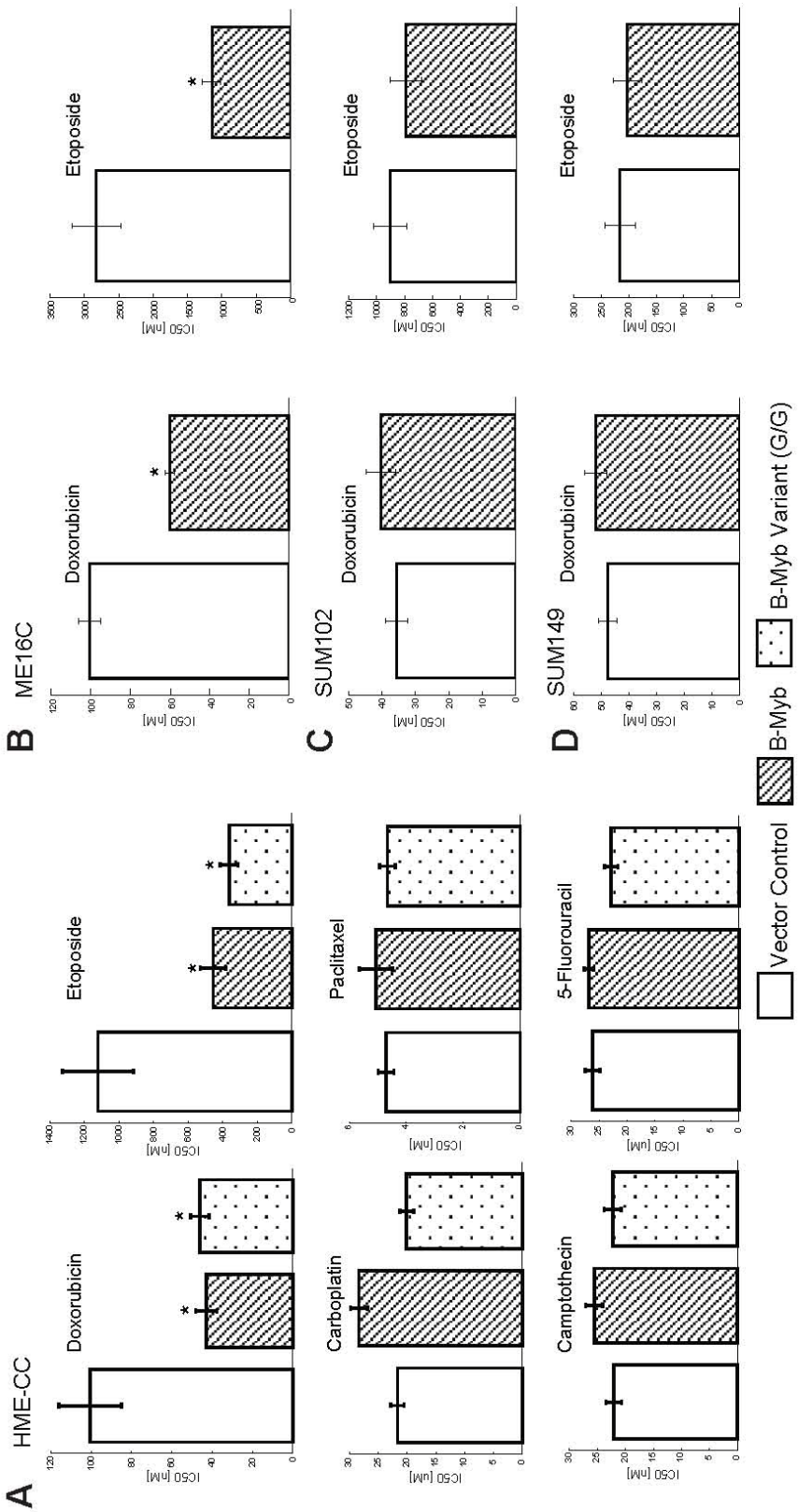

A
HME-CC
Doxorubicin
Etoposide
Carboplatin
Paclitaxel
Camptothecin
5-Fluorouracil
B
ME16C
C
SUM102
D
SUM149
IC50 [nM]
IC50 [uM]
Vector Control
B-Myb
B-Myb Variant (G/G)

Figure 2.5. Drug sensitivities in *B-Myb* overexpressing cell lines. IC50 doses (72h) of chemotherapy on cell lines stably expressing vector control, *B-Myb*, or *B-Myb* S427G variant. Each MTT experiment was performed in triplicate and error bars represent 95% confidence intervals (*$p<0.001$ relative to vector control). (A) *hTERT*-immortalized HMEC line HME-CC. (B) *hTERT*-immortalized HMEC line ME16C. (C) Basal-like tumor derived cell line SUM102 and (D) SUM149.

B-Myb is thought to be a transcriptional regulator of G2/M genes (Zhu *et al.*, 2004). To determine if the 217-gene SAM list was enriched for genes within a particular phase of the cell cycle, each gene was assigned to a specific phase by comparing them to a known list derived from a precise cell cycle microarray time course experiment (Whitfield *et al.*, 2002), or by literature search. Out of 217 significant genes, 101 genes were identified as being specifically induced during the cell cycle and 60/101 genes were assigned as G2/M-specific genes (Figure 2.6). In addition, previous known B-Myb target genes were present on the SAM list including *CDC2*, *Cyclin B1*, *BIRC5*, and the B-Myb binding partner, *LIN-9*.

Table 2.3. Top 15 significant gene ontology categories determined by EASE analysis for *B-Myb* associated genes**.**

Doxorubicin-induced genes:					
Gene Ontology Category	**List Hits**	**List Total**	**Population Hits**	**Population Total**	**Bonferroni p-value**
Mitotic cell cycle	58	138	352	13248	2.59E-51
Cell cycle	65	138	745	13248	8.60E-41
M phase	37	138	174	13248	1.22E-34
Nuclear division	35	138	167	13248	2.84E-32
Mitosis	32	138	131	13248	2.31E-31
M phase of mitotic cell cycle	32	138	133	13248	3.87E-31
Cell proliferation	65	138	1116	13248	4.33E-30
DNA replication and chromosome cycle	29	138	199	13248	2.21E-21
DNA metabolism	36	138	579	13248	4.60E-15
DNA replication	20	138	156	13248	1.22E-12
S phase of mitotic cell cycle	20	138	158	13248	1.54E-12
Spindle	17	135	97	12954	1.66E-12
Regulation of cell cycle	28	138	406	13248	4.68E-12
Cytokinesis	16	138	112	13248	3.14E-10
Cell growth and/or maintenance	81	138	3996	13248	3.44E-09

Abbreviations: EASE, Expression Analysis Systematic Explorer. Significant genes, as determined by SAM, for doxorubicin-treated *B-Myb* overexpressing HME-CC cells versus doxorubicin-treated controls were input to EASE and analyzed for enriched gene ontology categories.

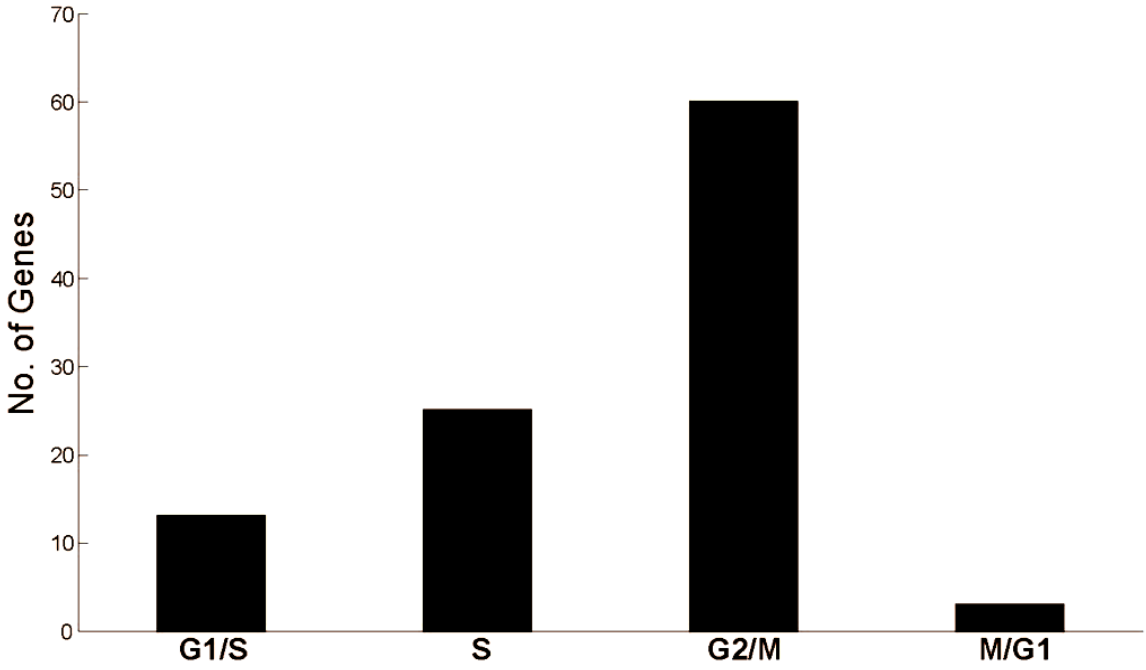


Figure 2.6. Enrichment of G2/M phase genes in doxorubicin-treated *B-Myb* overexpressing HME-CC cells. Significance Analysis of Microarray was used to identify 217 significant genes whose high expression was present in *B-Myb* overexpressing cells. These genes were then assigned to a specific phase of the cell cycle by comparing them to Whitfield *et al.*, 2002, which identified 101/217 genes as being specifically induced during the cell cycle. The graph shows to which phase of the cell cycle these 101 genes mapped.

Cell Cycle Profiles of *B-Myb* Overexpressing Cells. Since *B-Myb* expressing, doxorubicin treated cells produced a significant G2/M-enriched gene list, we hypothesized that the cell cycle profiles of these cells may be different than treated controls. *B-Myb* expressing HME-CC and empty vector controls were both treated with a range of doxorubicin doses for 48 hours and their cell cycle profiles analyzed for DNA content using flow cytometry. At zero dose or high dose of doxorubicin, the cell cycle profiles for both *B-Myb* overexpressing cells and controls were identical in terms of the percentage of cells in G1 and G2/M phase (Figure 2.7). However, at low and intermediate concentrations (10-35 nM) of doxorubicin there was a significant difference in the number of cells in G1 or G2/M with a larger percentage of *B-Myb* overexpressing cells in G1 versus controls, and a lower percentage of *B-Myb* overexpressing cells in G2/M (Figure 2.7A and 2.7B, respectively); very similar results were obtained with etoposide treatment (Figure 2.8). At high doses of doxorubicin or etoposide, regardless of *B-Myb* expression, the majority of cells arrested in G2/M.

DISCUSSION

Clinically defined as ER-, PR-, and HER2 not amplified, the basal-like subtype of breast cancer portends a poor prognosis. Basal-like breast cancer has an inherently high proliferation rate, which by microarray is identified by high expression of the "proliferation signature" genes; this signature has been identified in many publications, in many tumor types, and is highly enriched for cell cycle regulated genes (Whitfield *et al.*, 2006). Here, we find that *B-Myb*, a gene highly expressed within the proliferation cluster, plays an important role in regulating cell cycle progression, which likely has effects on patient prognosis and response to chemotherapy.

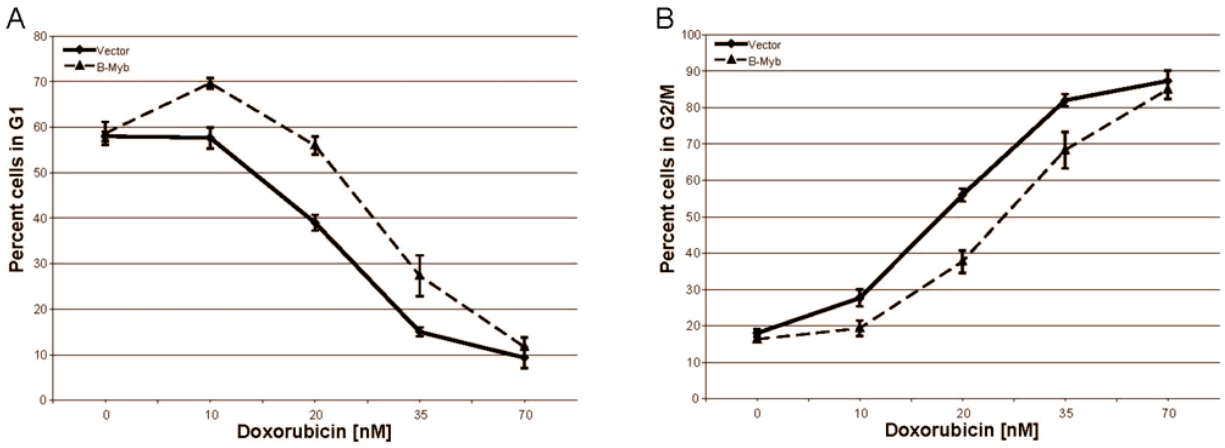


Figure 2.7. Cell cycle profile of HME-CC cells stably expressing *B-Myb* and treated with doxorubicin. Cell cultures were treated with a range of doses of doxorubicin for 48 hours followed by propidium iodide DNA content analysis. Percentage of cells in (A) G1 phase and (B) G2/M were calculated by gating based on DNA content. Error bars indicate standard deviations between three independent experiments.

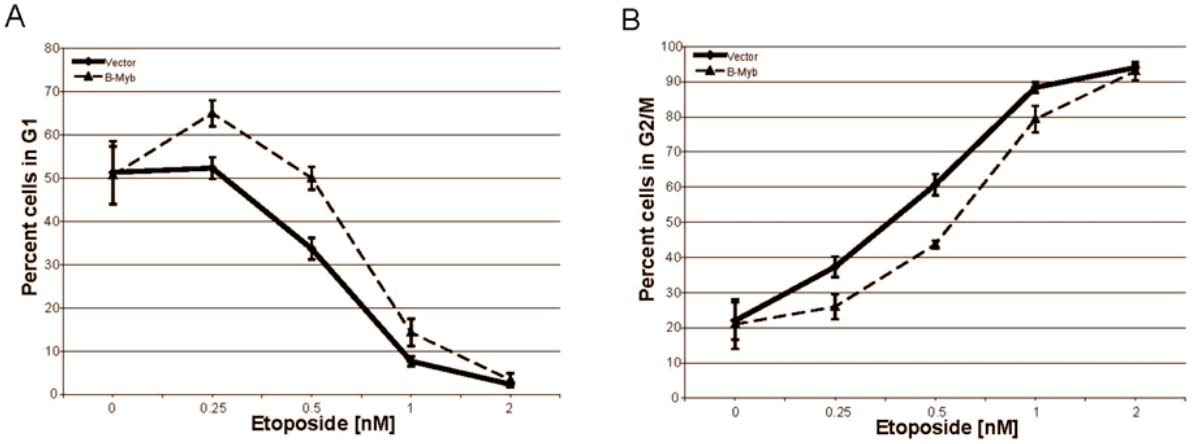


Figure 2.8. Cell cycle profile of HME-CC cells stably expressing *B-Myb* and treated with etoposide. Cell cultures were treated with a range of doses of etoposide for 48 hours followed by propidium iodide DNA content analysis. Percentage of cells in (A) G1 phase and (B) G2/M were calculated by gating based on DNA content. Error bars indicate standard deviations between three independent experiments.

We showed that *B-Myb* high expression was significantly associated with the poor outcome, basal-like breast cancer subtype and that *B-Myb* gene expression levels alone predicted poor outcomes in the absence of therapy (i.e. prognosis, Figure 2.2) and are correlated with achieving a pCR (i.e. prediction, Table 2.1). Stratification of *B-Myb* expression, even within the luminal A, B, and HER2+/ER- subtypes, or ER+ tumors, was significantly associated with survival. Stratification of *B-Myb* expression *within* basal-like tumors did not predict outcomes, nor was outcome predicted in the ER- subset of the Wang *et al.*, 2005 dataset (which includes basal-like tumors); however, these are inherently poor outcome tumors that significantly trend towards high *B-Myb* expression (Figure 2.3).

The *B-Myb* chromosomal locus, 20q13, is found amplified in a variety of cancers, including breast (Chin *et al.*, 2006), and this amplification is linked to poor prognosis (Bergamaschi *et al.*, 2006). Gene copy number analysis across breast tumor subtypes suggested that the genomic amplification of *B-Myb* was not enriched in the basal-like subtype, but instead was enriched in the luminal B subtype (Bergamaschi *et al.*, 2006). Therefore, within luminal tumors, *B-Myb* amplification and expression is an event that appears to be selected for and portends a poor prognosis (Figures 2.2B, C). In basal-like tumors the high expression of *B-Myb* may be due to other regulatory mechanisms, possibly by virtue of their inherently high proliferation rates, by amplification of transcription factor(s) targeting *B-Myb*, or by selectively enhanced promoter activity. This dual behavior of basal-like tumors, a subtype typical of chemo-responsiveness but with poor patient outcomes, has been described before and termed the “basal-like tumor paradox” (Carey *et al.*, 2006), and the data presented here suggest that B-Myb may be a key regulator of this complex phenotype.

Recently, Schwab *et al*. 2007, linked the nonsynonymous *B-Myb* germline variant (S427G, rs2070235) to a decrease in overall cancer risk when combining neuroblastomas, chronic myelogenous leukemia, and colon cancers into a single dataset and comparing this grouping of cases to non-cancer bearing controls. While both studies are technically accurate (i.e. similar allele frequencies in the control populations; CBCS: homozygous major allele 72.6%, heterozygous 24%, homozygous minor allele 3.4%; Schwab *et al.*, 2007: 72.1%, 27.3%, and 0.4%, respectively), the association of *B-Myb* rs2070235 genotype and basal-like breast cancer susceptibility differs from the previous report and has not been described before. Here, in a population based case-control study, we found that the rs2070235 minor allele was associated with increased risk of basal-like breast cancer but not other subtypes in the CBCS (Table 2.2). The discrepancy between the Schwab data and what is reported here may be due to differences between breast cancer and the cancers examined in the Schwab paper. Even amongst breast cancers there is significant heterogeneity across the subtypes in terms of the etiologic role of the *B-Myb* allele, and an analysis of breast cancer without subdivision would have missed the association of rs2070235 with the basal-like breast cancer subtype.

Presently, the function of the S427G variant is unknown. Our *in vitro* studies in breast epithelial cell lines showed neither a phenotypic difference between the overexpression of *B-Myb* or its variant on chemosensitivity relative to each other, nor did we observe a difference in their baseline gene expression patterns. It is tempting to speculate that rs2070235 has functional consequences and somehow contributes to the etiology of basal-like tumors. The *B-Myb* variant was found to be nearly ten-fold more frequent in African-Americans (7%) versus non-African-Americans (0.8%) in the CBCS. This is relevant in

light of recent data showing that premenopausal African-Americans are approximately twice as likely to develop basal-like tumors compared with premenopausal Caucasians (Carey *et al.*, 2006; Millikan *et al.*, 2008). Of note, Schwab *et al.*, 2007 demonstrated that the B-Myb S427G protein was more stable than the wild-type protein. This increased stability may correspond to elevated B-Myb protein levels, possibly increasing transcriptional activity of G2/M cell cycle genes and leading to higher inherent proliferation rates. This alteration of cell cycle control may contribute to B-Myb's influence on poor outcome breast cancers. Additional assays are required to more fully investigate the role of rs2070235 and other variants in *B-Myb* that may also lie in linkage disequilibrium with the G allele. This is an important area of investigation, since it is possible that one or more variants in *B-Myb* could contribute to the higher frequency of basal-like breast cancer among African-American breast cancer patients and may contribute to the pathophysiology of basal-like breast cancer.

Since chemotherapy is currently the only option for basal-like patients, we explored if increased *B-Myb* expression had any effect on chemosensitivity *in vitro*. We observed a statistically significant increase in sensitivity to two TOP2A inhibitors, doxorubicin and etoposide, in the HME-CC and ME16C cells overexpressing *B-Myb* or its variant, but this phenotype was not observed in the basal-like tumor derived cell lines (Figure 2.5). Also, there was no significant difference in chemosensitivity to the other tested drugs in any of the cell lines, which included camptothecin, a DNA topoisomerase I (TOP1) inhibitor. TOP2A, a nuclear enzyme that relaxes both negative and positive DNA supercoils by creating double-stranded DNA breaks, is of particular importance for proper DNA duplication during S-phase of the cell cycle (Smith *et al.*, 1994). TOP2A inhibitors cause the enzyme to become trapped on double-strand DNA breakpoints, thereby causing G2 cell cycle checkpoint arrest. Since

B-Myb is a G2/M regulating gene, increased *B-Myb* expression may be promoting the TOP2A-inhibitor treated cells through the G2 checkpoint via induction of downstream B-Myb target genes. By facilitating cells through G2, with less regard for DNA damage, the *B-Myb* overexpressing cells may attempt to cycle again, eventually leading to the increased sensitivity to TOP2A phenotype observed; see hypothetical model (Figure 2.9). TOP2A itself was on the B-Myb induced gene list (Supplementary Table 2) and thus more of the target of doxorubicin and etoposide was present, adding to the observed sensitivity phenotype in *B-Myb* overexpressing cells.

In support of this hypothesis, the gene list identified by microarray analysis as being highly expressed in doxorubicin-treated, *B-Myb* overexpressing cells (Appendix II) was significantly enriched in genes required for G2/M progression. For example, this list included the newly described B-Myb interacting protein LIN-9, a protein required along with B-Myb for the transcription of G2/M genes (Osterloh *et al.*, 2007). In addition, our B-Myb target gene list contained many previously identified B-Myb targets including *CDC2*, *CyclinB1*, and *BIRC5* (Osterloh *et al.*, 2007; Zhu *et al.*, 2004). This gene list represents putative mammary B-Myb target genes and again suggests that B-Myb has its greatest effect on G2/M genes. In agreement with the gene list enrichment results, when *B-Myb* overexpressing cells were treated with low doses of doxorubicin or etoposide, more cells accumulated in G1 versus controls, suggesting that the control cell line was appropriately inhibiting cell cycle progression at G2, whereas *B-Myb* overexpressing lines were bypassing this checkpoint (Figures 2.7 and 2.8).

Figure 2.9

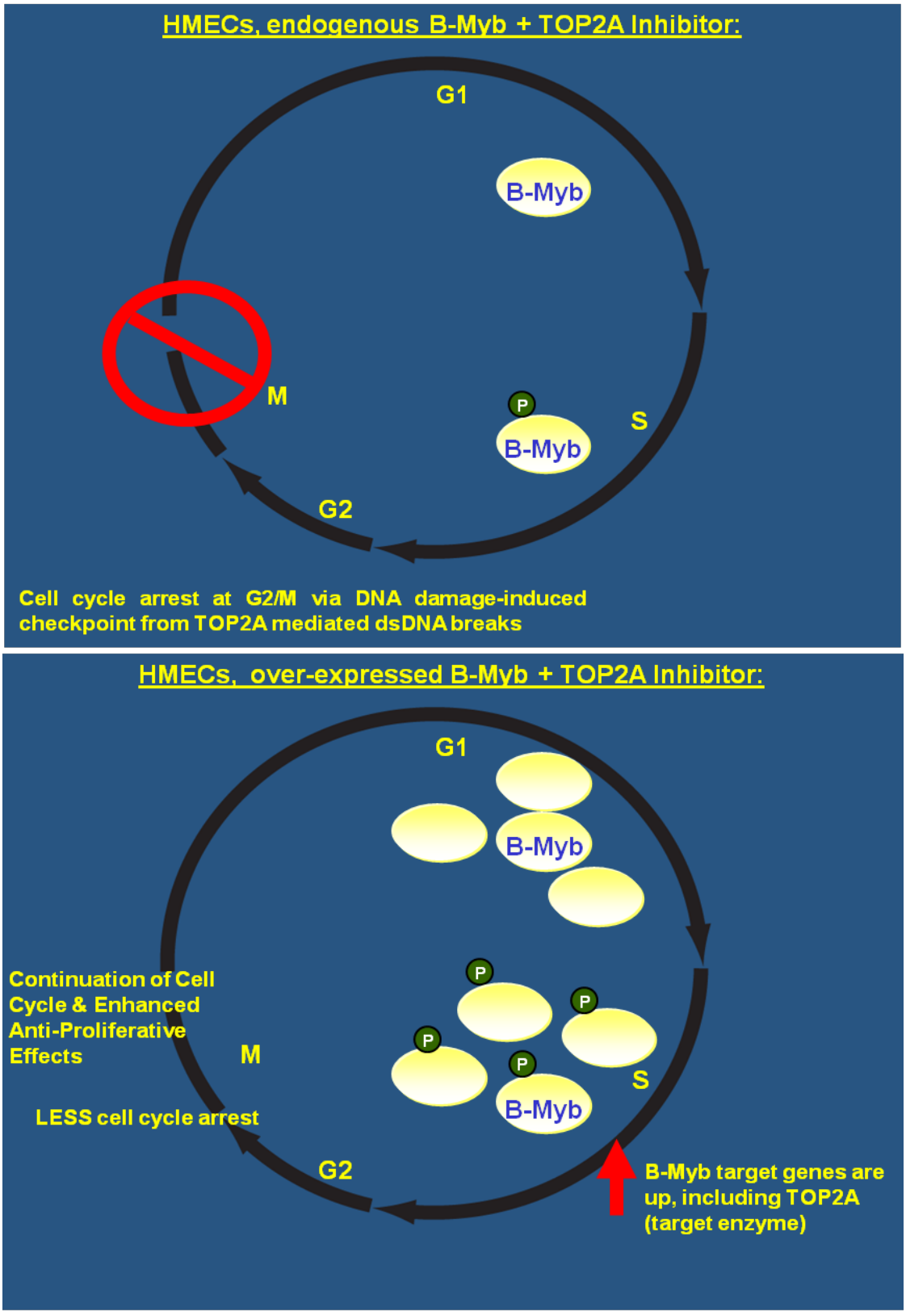

HMECs, endogenous B-Myb + TOP2A Inhibitor:
G1
B-Myb
M
P
B-Myb
S
G2
Cell cycle arrest at G2/M via DNA damage-induced checkpoint from TOP2A mediated dsDNA breaks
HMECs, over-expressed B-Myb + TOP2A Inhibitor:
G1
B-Myb
Continuation of Cell Cycle & Enhanced Anti-Proliferative Effects
M
P
P
P
P
B-Myb
S
LESS cell cycle arrest
G2
B-Myb target genes are up, including TOP2A (target enzyme)

Figure 2.9. *B-Myb* expression may increase anti-proliferative effects of TOP2A inhibitors. Hypothetical model depicting HME-CC cells treated with TOP2A inhibitor (A) expressing endogenous *B-Myb* levels or (B) *B-Myb* overexpression.

The results of this study enhance our understanding of the role of B-Myb in breast cancer by identifying new B-Myb target genes, by showing that this gene is highly expressed in basal-like breast cancers, and by showing it is of prognostic value for survival and predictive value for pathological complete response. Also, we have described a significant correlation between a *B-Myb* variant and an increased risk of basal-like breast cancer. These findings point to B-Myb as a biomarker that is of potential clinical importance for determining disease risk and for guiding treatment. In addition to its role in basal-like cancers, *B-Myb* may also be of great importance in luminal tumors, a breast cancer subtype with relatively good prognosis, since *B-Myb* expression was capable of stratifying the poor from the good actors within this group, and it is within this group where *B-Myb* is occasionally amplified on the DNA level. The link between *B-Myb* high expression and increased chemotherapy sensitivity *in vitro* is mirrored by similar findings *in vivo* where we and others have shown that basal-like tumors are, on average, the most sensitive to multi-agent chemotherapy regimens that contain an anthracycline (Carey *et al.*, 2007; Rouzier *et al.*, 2005). It is unlikely that *B-Myb* expression alone is responsible for the chemotherapy sensitivity of basal-like tumors, but when coupled with the loss of TP53 function, which is known to occur in basal-like tumors (Sorlie *et al.*, 2001), and the loss of RB function (Derenzini *et al.*, 2008; Herschkowitz *et al.*, 2008), multiple important checkpoints are deficient in basal-like tumors and it is this lack of control that may ultimately prove to be their Achilles' heel.

ACKNOWLEDGMENTS

We thank Drs. Melissa A. Troester and William K. Kaufmann for comments and discussion. This work was supported by funds from the NCI Breast SPORE program (P50-CA58223-09A1), by RO1-CA-101227-01, by the Breast Cancer Research Foundation and the V Foundation for Cancer Research.

REFERENCES

Bergamaschi A, Kim YH, Wang P, Sørlie T, Hernandez-Boussard T, Lonning PE *et al* (2006). Distinct patterns of DNA copy number alteration are associated with different clinicopathological features and gene-expression subtypes of breast cancer. *Genes, Chromosomes and Cancer* **45:** 1033-1040.

Carey LA, Dees EC, Sawyer L, Gatti L, Moore DT, Collichio F *et al* (2007). The Triple Negative Paradox: Primary Tumor Chemosensitivity of Breast Cancer Subtypes. *Clin Cancer Res* **13:** 2329-2334.

Carey LA, Perou CM, Livasy CA, Dressler LG, Cowan D, Conway K *et al* (2006). Race, Breast Cancer Subtypes, and Survival in the Carolina Breast Cancer Study. *JAMA* **295:** 2492-2502.

Chin K, DeVries S, Fridlyand J, Spellman PT, Roydasgupta R, Kuo W-L *et al* (2006). Genomic and transcriptional aberrations linked to breast cancer pathophysiologies. *Cancer Cell* **10:** 529-541.

Davidson CJ, Tirouvanziam R, Herzenberg LA, Lipsick JS (2005). Functional Evolution of the Vertebrate Myb Gene Family: B-Myb, but Neither A-Myb nor c-Myb, Complements Drosophila Myb in Hemocytes. *Genetics* **169:** 215-229.

Derenzini M, Donati G, Mazzini G, Montanaro L, Vici M, Ceccarelli C *et al* (2008). Loss of Retinoblastoma Tumor Suppressor Protein Makes Human Breast Cancer Cells More Sensitive to Antimetabolite Exposure. *Clin Cancer Res* **14:** 2199-2209.

Herschkowitz JI, He X, Fan C, Perou CM (2008). The functional loss of the retinoblastoma tumor suppressor is a common event in Basal-like and Luminal B breast carcinomas. *Breast Cancer Reserach*.

Hess KR, Anderson K, Symmans WF, Valero V, Ibrahim N, Mejia JA *et al* (2006). Pharmacogenomic Predictor of Sensitivity to Preoperative Chemotherapy With Paclitaxel and Fluorouracil, Doxorubicin, and Cyclophosphamide in Breast Cancer. *J Clin Oncol* **24:** 4236-4244.

Hoadley K, Weigman V, Fan C, Sawyer L, He X, Troester M *et al* (2007). EGFR associated expression profiles vary with breast tumor subtype. *BMC Genomics* **8:** 258.

Hosack DA, Jr GD, Sherman BT, Lane HC, Lempicki RA (2003). Identifying biological themes within lists of genes with EASE. *Genome Biology* **4**.

Hu Z, Fan C, Oh D, Marron J, He X, Qaqish B *et al* (2006). The molecular portraits of breast tumors are conserved across microarray platforms. *BMC Genomics* **7:** 96.

Hu Z, Troester MA, Perou CM (2005). High reproducibility using sodium hydroxide-stripped long oligonucleotide DNA microarrays. *Biotechniques* **38:** 121-4.

Ito M, Araki S, Matsunaga S, Itoh T, Nishihama R, Machida Y *et al* (2001). G2/M-Phase-Specific Transcription during the Plant Cell Cycle Is Mediated by c-Myb-Like Transcription Factors. *Plant Cell* **13:** 1891-1905.

Klempnauer K-H, Gonda TJ, Bishop MJ (1982). Nucleotide sequence of the retroviral leukemia gene v-myb and its cellular progenitor c-myb: The architecture of a transduced oncogene. *Cell* **31:** 453-463.

Leprince D, Gegonne A, Coll J, de Taisne C, Schneeberger A, Lagrou C *et al* (1983). A putative second cell-derived oncogene of the avian leukaemia retrovirus E26. *Nature* **306:** 395.

Malaterre J, Carpinelli M, Ernst M, Alexander W, Cooke M, Sutton S *et al* (2007). c-Myb is required for progenitor cell homeostasis in colonic crypts. *PNAS* **104:** 3829-3834.

Miller LD, Smeds J, George J, Vega VB, Vergara L, Ploner A *et al* (2005). An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival. *Proceedings of the National Academy of Sciences* **102:** 13550-13555.

Millikan R, Eaton A, Worley K, Biscocho L, Hodgson E, Huang W-Y *et al* (2003). HER2 Codon 655 Polymorphism and Risk of Breast Cancer in African Americans and Whites. *Breast Cancer Research and Treatment* **79:** 355.

Millikan RC, Newman B, Tse C-K, Patricia G. Moorman, Conway K, Smith LV *et al* (2008). Epidemiology of basal-like breast cancer. *Breast Cancer Research and Treatment* **109:** 123-139.

Mucenski ML, McLain K, Kier AB, Swerdlow SH, Schreiner CM, Miller TA *et al* (1991). A functional c-myb gene is required for normal murine fetal hepatic hematopoiesis. *Cell* **65:** 677.

Ness SA (2003). Myb protein specificity: evidence of a context-specific transcription factor code. *Blood Cells, Molecules, and Diseases* **31:** 192.

Newman B, Moorman P, Millikan R, Qaqish B, Geradts J, Aldrich T *et al* (1995). The Carolina Breast Cancer Study: integrating population-based epidemiology and molecular biology. *Breast Cancer Research and Treatment* **35:** 51-60.

Nomura N, Takahashi M, Matsui M, Ishii S, Date T, Sasamoto S *et al* (1988). Isolation of human cDNA clones of myb-related genes, A-myb and B-myb. *Nucl. Acids Res.* **16:** 11075-11089.

Osterloh L, vonEyss B, Schmit F, Rein L, Hübner D, Samans B *et al* (2007). The human synMuv-like protein LIN-9 is required for transcription of G2/M genes and for entry into mitosis. *The EMBO Journal* **26:** 144-157.

Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA *et al* (2000). Molecular portraits of human breast tumours. *Nature* **406:** 747.

Pomeroy SL, Tamayo P, Gaasenbeek M, Sturla LM, Angelo M, McLaughlin ME *et al* (2002). Prediction of central nervous system embryonal tumour outcome based on gene expression. *Nature* **415:** 436.

Rosinski JA, Atchley WR (1998). Molecular Evolution of the Myb Family of Transcription Factors: Evidence for Polyphyletic Origin. *Journal of Molecular Evolution* **46**.

Rouzier R, Perou CM, Symmans WF, Ibrahim N, Cristofanilli M, Anderson K *et al* (2005). Breast Cancer Molecular Subtypes Respond Differently to Preoperative Chemotherapy. *Clin Cancer Res* **11:** 5678-5685.

Sala A (2005). B-MYB, a transcription factor implicated in regulating cell cycle, apoptosis and cancer. *European Journal of Cancer* **41:** 2479-2484.

Schwab R, Bussolari R, Corvetta D, Chayka O, Santilli G, Kwok JMM *et al* (2007). Isolation and functional assessment of common, polymorphic variants of the B-MYB proto-oncogene associated with a reduced cancer risk. *Oncogene*.

Smith P, Soues S, Gottlieb T, Falk S, Watson J, Osborne R *et al* (1994). Etoposide-induced cell cycle delay and arrest-dependent modulation of DNA toposisomerase II in small-cell lung cancer cells. *British Journal of Cancer* **70:** 914-921.

Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H *et al* (2001). Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *PNAS* **98:** 10869-10874.

Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A *et al* (2003). Repeated observation of breast tumor subtypes in independent gene expression data sets. *PNAS* **100:** 8418-8423.

Tanaka Y, Patestos NP, Maekawa T, Ishii S (1999). B-myb is Required for Inner Cell Mass Formation at an Early Stage of Development. *J. Biol. Chem.* **274:** 28067-28070.

Toscani A, Mettus RV, Coupland R, Simpkins H, Litvin J, Orth J *et al* (1997). Arrest of spermatogenesis and defective breast development in mice lacking A-myb. *Nature* **386:** 713.

Trauth K, Mutschler B, Jenkins N, Gilbert D, Copeland N, Klempnauer K (1994). Mouse A-myb encodes a trans-activator and is expressed in mitotically active cells of the developing central nervous system, adult testis and B lymphocytes. *The EMBO Journal* **13:** 5994–6005.

Troester MA, Hoadley KA, Sorlie T, Herbert B-S, Borresen-Dale A-L, Lonning PE *et al* (2004). Cell-Type-Specific Responses to Chemotherapeutics in Breast Cancer. *Cancer Res* **64:** 4218-4226.

Tusher VG, Tibshirani R, Chu G (2001). Significance analysis of microarrays applied to the ionizing radiation response. *PNAS* **98:** 5116-5121.

van de Vijver MJ, He YD, van 't Veer LJ, Dai H, Hart AAM, Voskuil DW *et al* (2002). A Gene-Expression Signature as a Predictor of Survival in Breast Cancer. *N Engl J Med* **347:** 1999-2009.

Vanewijk PH, Hoekstra JA (1993). Calculation of the EC50 and Its Confidence Interval When Subtoxic Stimulus Is Present. *Ecotoxicology and Environmental Safety* **25:** 25.

Wang Y, Klijn JGM, Zhang Y, Sieuwerts AM, Look MP, Yang F *et al* (2005). Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *The Lancet* **365:** 671.

Whitfield ML, George LK, Grant GD, Perou CM (2006). Common markers of proliferation. *Nat Rev Cancer* **6:** 99-106.

Whitfield ML, Sherlock G, Saldanha AJ, Murray JI, Ball CA, Alexander KE *et al* (2002). Identification of Genes Periodically Expressed in the Human Cell Cycle and Their Expression in Tumors. *Mol. Biol. Cell* **13:** 1977-2000.

Zhu W, Giangrande PH, Nevins JR (2004). E2Fs link the control of G1/S and G2/M transcription. *The EMBO Journal* **23:** 4615-26.

CHAPTER III

TUMOR SUPPRESSOR ROLE FOR THE C-MYB ONCOGENE IN LUMINAL BREAST CANCER

PREFACE

This chapter represents the results described in two manuscripts. The first manuscript, describing the role of *c-Myb* in luminal breast cancer, is in the submission process. My role in this project included initiating, organizing, and designing the study. I performed the microarray data analysis, *c-Myb* knock-down in mammary cell lines, chemotherapy treatment, and both *in vitro* and *in vivo* tumorigenicity experiments. I initiated all aspects of manuscript preparation, including writing, figure preparation, and revisions. Katherine A. Hoadley aided in study design and contributed writing of the manuscript. Joel S. Parker assisted with microarray statistical analysis. Charles M. Perou was the principal investigator, conceived the project, and contributed to writing the manuscript.

Thorner AR, Parker J, Hoadley KA, Perou CM. Tumor suppressor role for the c-Myb oncogene in luminal breast cancer. (Manuscript in preparation)

The second manuscript, examining the role of c-Myb in p53 stabilization, is currently under review at *Molecular and Cellular Biology*. Chad Deisenroth and principal investigator Yanping Zhang carried out experiments describing the interaction between Hep27 and p53. I performed all microarray analysis and *c-Myb* knock-down experiments.

Deisenroth C, **Thorner AR**, Enomoto T, Perou CM, and Zhang, Y. Mitochondrial Hep27 is a c-Myb target gene that inhibits Mdm2 and stabilizes p53. (Submitted to MCB, 2009)

ABSTRACT

The transcription factor *c-Myb* has been well characterized as an oncogene in several human tumor types, and its expression in the hematopoietic stem/progenitor cell population is essential for proper hematopoiesis. However, the role of *c-Myb* in mammopoeisis and breast tumorigenesis is poorly understood, despite its high expression in the majority of breast cancer cases (60-80%). We find that *c-Myb* high expression in human breast tumors correlates with the luminal/ER+ phenotype and a good prognosis. RNAi knock-down of endogenous *c-Myb* levels in the MCF7 luminal breast tumor cell line increases tumorigenesis both *in vitro* and *in vivo*, suggesting a tumor suppressor role in luminal breast cancer. We created a mammary-derived *c-Myb* expression signature and found it highly correlated with a published mature luminal mammary cell signature and least correlated with a mammary stem/progenitor lineage gene signature. These data describe, for the first time, a tumor suppressor role for the *c-Myb* proto-oncogene in breast cancer that has implications for understanding luminal tumorigenesis and for guiding treatment.

INTRODUCTION

Breast cancer is a heterogeneous disease and numerous studies have defined at least five molecular subtypes of breast tumors using an "intrinsic" gene set (Hu *et al.*, 2006; Parker *et al.*, 2009; Perou *et al.*, 2000). The luminal/estrogen receptor alpha positive (ER+) subtypes are the most commonly diagnosed breast cancer (60-80 percent), with patients being classified as either good outcome Luminal A, or poor outcome Luminal B. Patients with Luminal A tumors have good overall survival, in part, because these tumors are slow growing, often respond to selective estrogen receptor modulators, and have infrequent *TP53* mutations (Riggs and Hartmann, 2003; Sorlie *et al.*, 2001). The luminal subtypes are defined by high expression of approximately 80 genes within the intrinsic gene list, including *ERα*, *GATA3*, *FOXA1*, and *c-Myb*, the latter of which is a known proto-oncogene frequently observed as amplified in a variety of tumor types (Oh and Reddy, 1999).

Nearly three decades ago the *c-Myb* transcription factor was identified as the mammalian homolog of *v-myb*, a transforming retroviral oncogene linked to avian leukemia (Klempnauer *et al.*, 1982; Leprince *et al.*, 1983). Since that time, *c-Myb* high expression has been associated with oncogenic activity and poor prognosis in several human cancers including T-cell leukemia, acute myelogenous leukemia, colorectal tumors, and most recently in adenoid cystic carcinomas (Biroccio *et al.*, 2001; Oh and Reddy, 1999; Persson *et al.*, 2009). In addition, *c-Myb* has been implicated in progenitor cell maintenance and is required for proper cellular differentiation in the hematopoietic system, neuronal cells, skin cells, and colonic crypts (Kopecki Z, 2007; Malaterre *et al.*, 2007; Malaterre *et al.*, 2008; Mucenski *et al.*, 1991). *c-Myb* high expression is frequently associated with a variety of immature cell lineages, and expression levels decrease as cells differentiate (Ess, 1999).

However, there has been very little discovered about the role of *c-Myb* in normal mammopoeisis and breast tumorigenesis, despite its high expression in virtually all ER+ tumors as well as in 29% of hereditary and typically ER-negative BRCA1 breast cancers (Guérin M, 1990; Kauraniemi *et al.*, 2000).

To gain insight into *c-Myb* and its involvement in breast cancer, we analyzed the expression of *c-Myb* in the context of breast tumor subtypes, and examined its association with patient outcomes. We also manipulated the expression of *c-Myb* via RNA interference in a luminal/ER+ mammary cell line, observed alterations in growth properties both *in vitro* and *in vivo,* and identified a mammary-specific *c-Myb* gene signature.

MATERIALS AND METHODS

Cell Culture. *c-Myb knock-down.* MCF7 cells (a gift from F. Tamanoi, University of California-Los Angeles, Los Angeles, CA) were maintained in RPMI-1640 plus 10% FBS at 37C and 5% CO_2. Stable knock-down of c-Myb in MCF7 cells was accomplished using a short hairpin RNA against c-Myb (shMYB; CGTTGGTCTGTTATTGCCAAGCACTTAAA) and compared to a knock-down control (shGFP) cloned into the pRS vector, purchased from OriGene (Catalog No. TR311329). Retrovial transduction was performed as described (Thorner *et al.*, 2009). Stable populations were selected by culturing in 2 ug/mL puromycin for two weeks. Cells were then plated at clonal density and >20 colonies screened by western blotting as described (Thorner *et al.*, 2009) for c-Myb (Abcam, ab45150), and β-tubulin (Santa Cruz, sc-9104). The clones with greatest knock-down were expanded for further analyses.

Addition of ZsGreen1. PT67 cells stably expressing retrovirus containing pLNCX2_ZsGreen1 (a generous gift from Dr. Kathryn B. Horwitz at the University of Colorado Health Sciences Center, Aurora, Colorado) were propagated as described (Harrell *et al.*, 2006). MCF7 cells stably expressing shMYB or shGFP were transduced, as described above, with pLNCX2_ZsGreen1-containing retrovirus and kept under constant selection using 400 ug/mL G418.

***In Vitro* Analyses.** *Doubling Time Assay*. MCF7 cells stably expressing shMYB or shGFP were seeded, in duplicate, into 10 centimeter dishes at 50,000 cells per plate. Cells were allowed 48 hours of growth prior to the first counting (t=0), followed by counting at 48, 72, and 124 hours (Beckman Z1 Coulter Particle Counter). Doubling times were estimated by linear regression.

Colony Formation Assay. Soft agar assays were performed in triplicate in six-well ultra-low attachment plates (Corning). Briefly, a medium-agar mix was prepared by combining 2x RPMI-1640 (Invitrogen, 23400-021), 5.6 mL 1x RPMI (Invitrogen), 2.4 mL FBS (Sigma), and 8 mL 1.8% Noble agar (Sigma, A5431-250G). A volume of 2.3 mL of the medium-agar mix was added to each well, to create a bottom layer, and allowed to solidify. MCF7 cells (shMYB or shGFP) were washed with PBS, trypsinized, counted, and 8,000 cells were combined with 3 mL of medium-agar mix to create the top agar layer in each well. Once the top agar layer solidified, 0.5 mL of selective media (RPMI-1640, 10% FBS, 2 ug/mL puromycin) was added to each well and changed with fresh media every three days. Cells were grown for 15-20 days until colonies were visible. Colonies were visualized by removing liquid media, adding 200 microliters of MTT dye (Cell-Titer 96, Promega

#G4100), incubating for one hour at 37 C, followed by scanning the plates and manual counting of colonies. Statistical significance was calculated using a two-tailed t-test.

MTT Assay. To estimate the IC50 of tamoxifen (Sigma-Aldrich #T9262) on cell lines, a modified MTT assay was performed as previously described (Thorner *et al.*, 2009).

***In Vivo* Tumor Analysis.** MCF7 cells stably expressing shMYB+ZsGreen1 or shGFP+ZsGreen1 were collected, counted, and 550,000 cells, embedded in Matrigel, were injected into each fourth mammary gland of Nude mice (Harlan Laboratories, Hsd:Athymic Nude-Foxn1^{nu}). Tumors were allowed to grow for 15 weeks then measured by caliper. Each tumor area was calculated and statistical significance between the means of experimental versus control determined using a one-tailed independent t-test.

Microarray Analysis. Poly-A(+) RNA was collected (Micro-FastTrack2.0 mRNA Isolation Kit, Invitrogen) from six replicates of MCF7 cells stably expressing shGFP and five replicates expressing shMYB, reverse transcribed and labeled using the Agilent Low RNA Input Linear Amplification Kit (Agilent Technologies, Santa Clara, CA, USA), and hybridized to Agilent Human 4x44K Custom Oligo microarrays, adapted from manufacturer's protocol. An untreated MCF7 cell line reference was co-hybridized to each array. Microarrays were scanned on an Agilent Technologies DNA Microarray Scanner with Surescan High-Resolution Technology (Part no. G2565CA) and the image analyzed using Agilent Feature Extraction Software. Data was normalized using Lowess normalization on the Cy3 and Cy5 channels. Microarray data is available at the UNC Microarray Database [http://genome.unc.edu] and will be available at the Gene Expression Omnibus.

Microarray Statistical Analyses. Supervised microarray analysis was performed on the MCF7-c-MYB RNAi vs. MCF7-vector control by selecting genes with an absolute signal intensity of at least 10 units in both dye channels and data present in at least 70% of experimental samples. A two-class, unpaired Significance Analysis of Microarrays (SAM) was performed to identify significant genes associated with *c-Myb* knock-down with a false discovery rate (FDR) of less than 5% (Tusher *et al.*, 2001).

Breast tumor microarray data on tumors without adjuvant chemotherapy (local-only Tx; n=145) from the Netherlands Cancer Institute (van de Vijver *et al.*, 2002) was used to analyze both *c-Myb*'s relation to survival and its expression across breast cancer subtypes. Gene expression levels of *c-Myb* were rank ordered, split into halves (high-to-low) and relation to survival tested using the chi-square test and visualized by Kaplan-Meier survival plots (WinSTAT v.2007.1). Testing the association of *c-Myb* expression versus subtypes was performed by ANOVA using the R system for statistical computing (R Development Core Team, 2006 http://www.R-project.org).

Other Statistical Analyses. To test the relation of the *c-Myb*, *GATA3*, and *ER* gene signatures to mammary cell lineage gene signatures, cell line models were used to characterize the different signatures. In each experiment, microarray data provides a signature of the activity for further evaluation. The gene signatures used in this experiment include *c-Myb* signature described above, a *GATA3*-signature from Usary *et al.* (Usary *et al.*, 2004), and an *ER*-signature from Oh *et al.* (Oh *et al.*, 2006); each of these three signatures were compared to the mammary cell lineage gene signatures previously defined (Lim *et al.*,

2009). Signatures were defined by the SAM statistics corresponding to differentially expressed genes (FDR< 5%) from the model comparison, and these SAM statistics were used to weight each gene. Tumor (UNC-324 breast tumor dataset; GSE10893) and mammary cell lineage samples were evaluated by calculating the inner product of the signature gene weights and the sample. The resulting value is a relative activity measure of the signature. Boxplots were generated to relate the activity measure of each signature across tumor subtypes and sorted mammary cell lines.

RESULTS

***c-Myb* High Expression Correlates with Good Prognosis.** To study the role of *c-Myb* in breast tumors, we first examined whether associations between *c-Myb* mRNA expression and patient outcomes exists. The Netherlands Cancer Institute locally-treated (no adjuvant chemotherapy) breast tumor microarray dataset (NKI-295, local-only Tx: n=145; (van de Vijver *et al.*, 2002) was classified into the intrinsic subtypes (Luminal A, Luminal B, HER2-enriched, Basal-like, and Normal-like) using the PAM50 classifier, as described (Parker *et al.*, 2009). An ANOVA was performed to determine statistical significance of *c-Myb* expression across the breast cancer subtypes (Figure 3.1A). Expression of *c-Myb* differed significantly across the subtypes with highest expression observed in the ER+, Luminal A and B subtypes and lowest expression in the Basal-like/ER- tumors. High *c-Myb* expression levels were also significantly correlated with smaller tumor size and lower grade (Figures 3.1B and C). Similar results were observed using two other published microarray datasets (data not shown) (Herschkowitz *et al.*, 2007; Miller *et al.*, 2005).

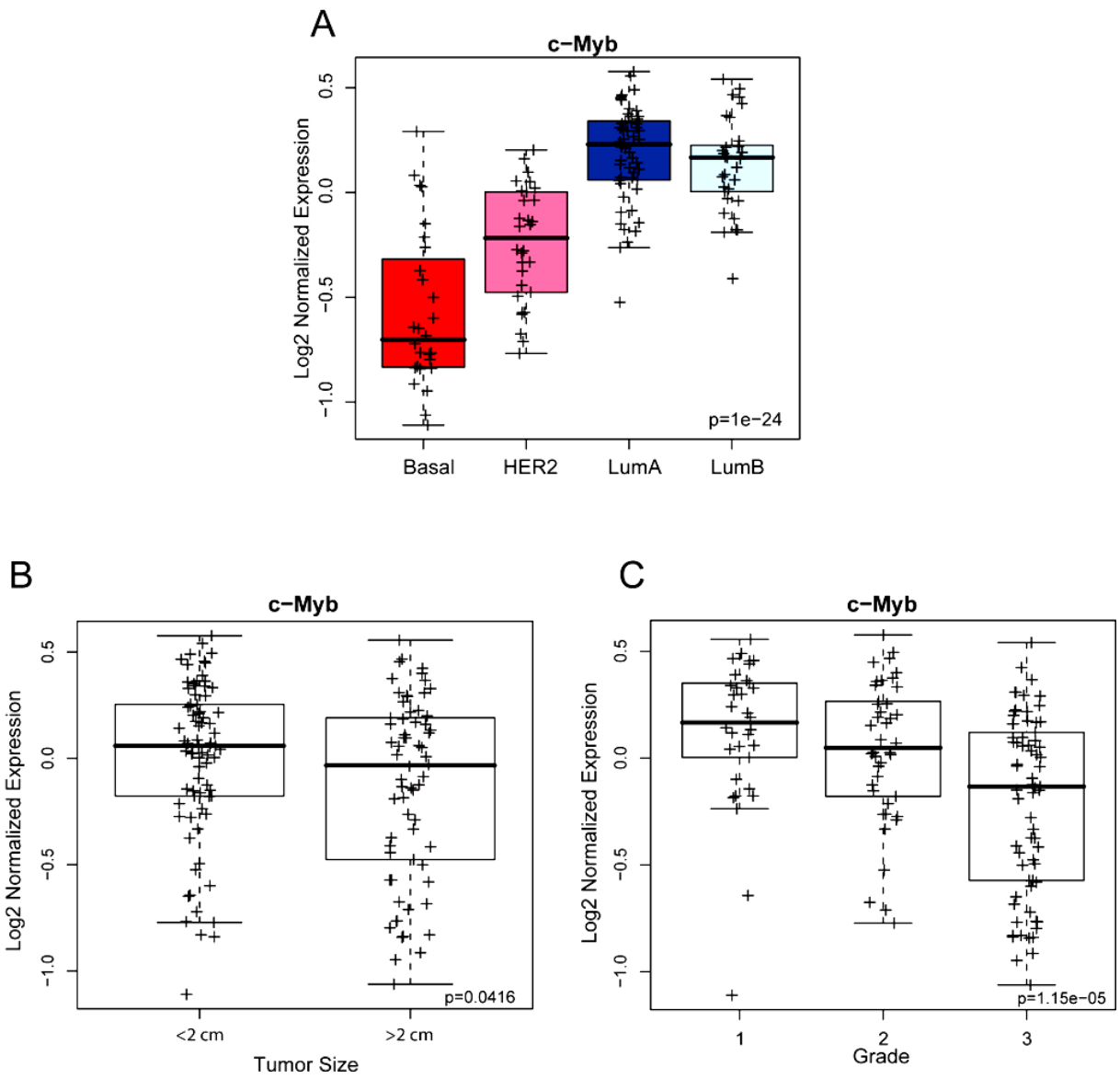


Figure 3.1. ***c-Myb*** **high expression correlates with luminal breast cancer, small tumor size and low tumor grade.** NKI local-only treated tumor dataset (n=145) was tested for (A) *c-Myb* expression across breast tumor subtypes. Statistical significance calculated by ANOVA. This same dataset was tested for the relationship of *c-Myb* gene expression to (B) tumor size or (C) tumor grade. Statistical significance determined by t-test or ANOVA.

Using this same NKI=145 cases dataset, we tested if *c-Myb* expression levels correlated with patient outcomes. Tumors were rank ordered (low-to-high) based on *c-Myb* mRNA expression levels and were analyzed for overall survival (OS) using Kaplan-Meier analysis in all tumors, as well as within each subtype. High *c-Myb* expression levels significantly correlated with better survival across all subtypes (Figure 3.2A; n=145), as well as in the combined Luminal A+B subtypes (Figure 3.2B; n=87), but not in the HER2-enriched or Basal-like subtypes (Figures 3.2C and D).

We also investigated if *c-Myb* expression levels correlate with achieving a pathological complete response (pCR). Using published data (Hess *et al.*, 2006) consisting of microarrays on breast tumors of patients receiving neoadjuvant paclitaxel, and subsequent 5FU-Adriamycin-Cyclophosphamide (T/FAC; n=133), we observed that high *c-Myb* levels significantly correlated with a poor pCR (Table 3.1); this finding is consistent with previous findings that high ER levels predict poor pCR rates (Hess *et al.*, 2006).

Table 3.1. Pathologic complete response, pCR, data of Hess *et al*., (2006) rank ordered, split into halves (high/low) based on *c-Myb* expression values, and analyzed by chi-square.

c-Myb:	High	Low
No pCR	55	44
pCR	11	23

χ^2=4.6, df=1, p=0.03

***c-Myb* Knock-down in MCF7 Cells Increases Tumorigenesis *in vitro* and *in vivo*.** Based on our observations that higher levels of *c-Myb* are predictive of good outcomes in all breast tumors, as well as in luminal tumors, we utilized RNA interference (short hairpin RNA:

shRNA) to knock-down endogenous c-Myb protein in the luminal tumor-derived cell line, MCF7 (Figure 3.3A). Microarray analysis revealed that transcript levels of *c-Myb* were decreased 2.5-fold relative to the control (data not shown), while western blot analysis showed little to no c-Myb protein expression. A cell proliferation assay was performed to compare the doubling time of MCF7 cells with the stable *c-Myb* knock-down (shMYB) versus control (shGFP) (Figure 3.3A). Cells with *c-Myb* knock-down grew faster *in vitro* (Figure 3.3B).

Furthermore, both cell lines were tested for anchorage-independent growth by means of colony formation in soft agar with MCF7 cells having depleted c-Myb protein able to form nearly 14-fold more colonies than controls (Figures 3.3C and D). We assayed both cell lines for their relative sensitivities to tamoxifen and found that MCF7 cells with reduced c-Myb levels were more resistant to tamoxifen based on IC50 estimates (Table 3.2).

Table 3.2. IC50 doses (72h) of tamoxifen on c-Myb knock-down (shMYB) versus control (shGFP).

Tamoxifen	IC50	Std Dev
shGFP	6.5	**6.2-6.9**
shMYB	7.9	**7.0-8.9**

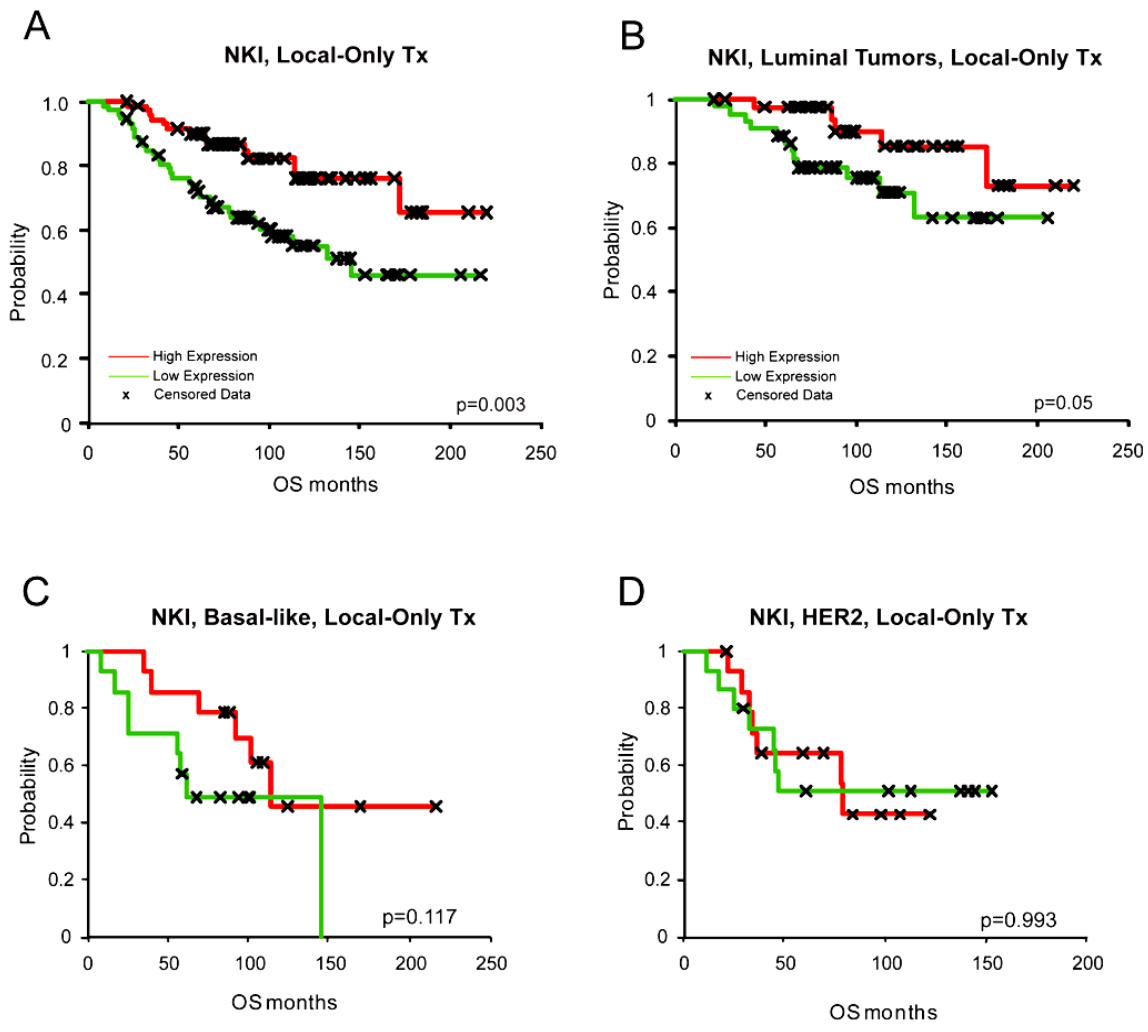


Figure 3.2. *c-Myb* high expression correlates with good prognosis in all breast subtypes combined and luminal tumors. Kaplan-Meier overall survival (OS) analysis based on *c-Myb* expression values rank ordered into halves (high-to-low): (A) All subtypes combined (n=145, local-only treatment (Tx)), (B) Luminal A and B subtypes only (n=87), (C) Basal-like (n=28) and (D) HER2+ (n=30).

Figure 3.3.

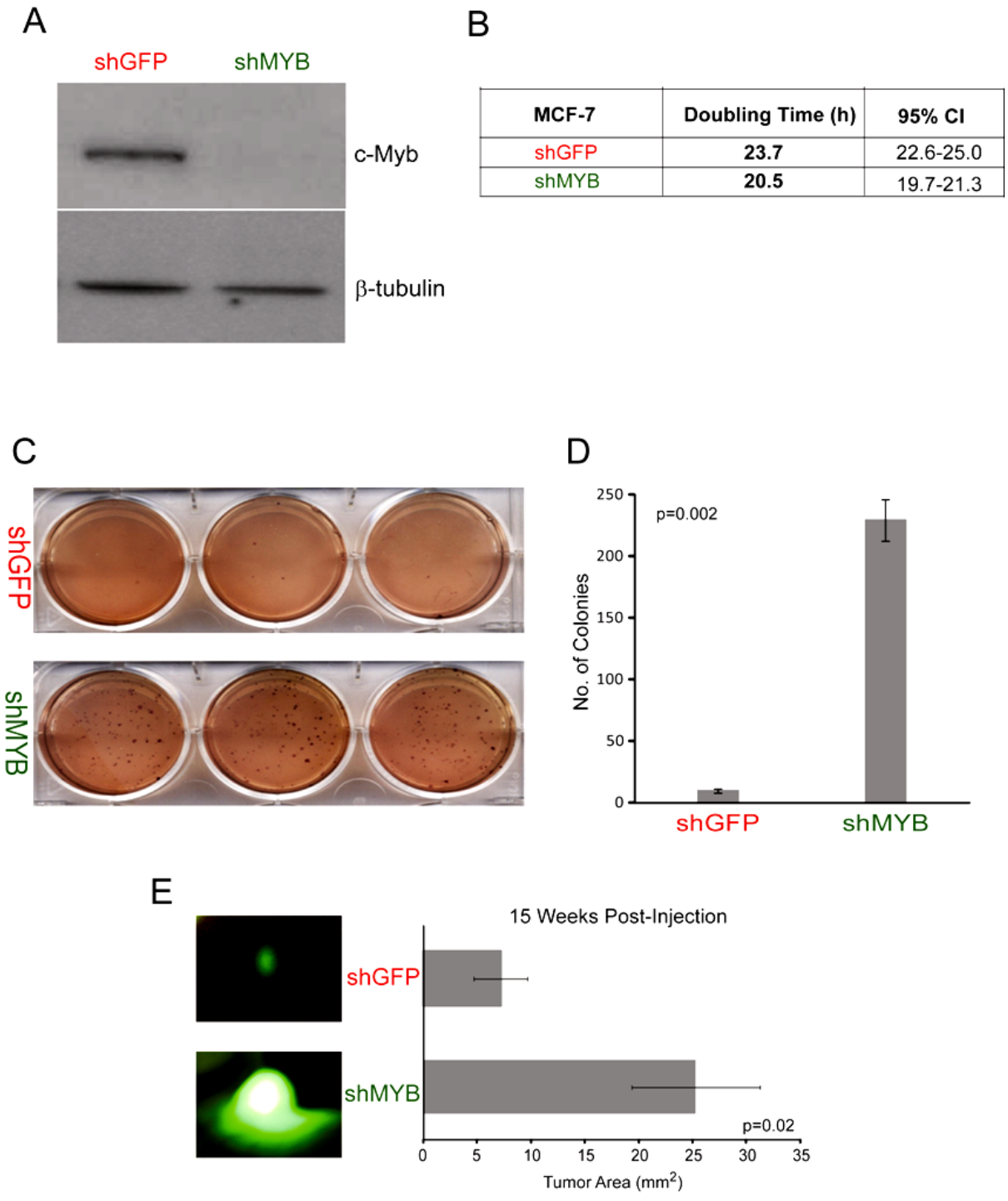

A
shGFP
shMYB
c-Myb
β-tubulin
B
MCF-7 | Doubling Time (h) | 95% CI
shGFP | 23.7 | 22.6-25.0
shMYB | 20.5 | 19.7-21.3
C
shGFP
shMYB
D
p=0.002
No. of Colonies
250
200
150
100
50
0
shGFP
shMYB
E
15 Weeks Post-Injection
shGFP
shMYB
p=0.02
0
5
10
15
20
25
30
35
Tumor Area (mm2)

Figure 3.3. c-Myb knock-down increases luminal tumor growth both *in vitro* and *in vivo*. (A) Western blot analysis of c-Myb knock-down (shMYB) versus control (shGFP) in MCF7 cells. (B) Doubling times calculated by linear regression. CI: confidence interval (C) Representative dishes displaying the effects of c-Myb knock-down in a soft agar colony formation assay. (D) Quantification of colony formation assay. Statistical significance determined by two-tailed independent t-test. (E) Tumor area of nude mice fifteen weeks after injection of MCF7 shGFP or shMYB; representative images were taken of each tumor type in live mice. Error bars represent standard error and p-value calculated by a two-tailed independent t-test.

To examine if the *in vitro* data were relevant in *in vivo* xenografts, we stably expressed ZsGreen1, a reef coral fluorescent protein, in both of the MCF7 cell lines (shMYB and shGFP) to allow for ease of *in vivo* tumor visualization. Nude mice (without estrogen pellets) were injected with $5.5x10^5$ cells that were embedded in Matrigel into each fourth mammary gland (shGFP, n=10; shMYB, n=9). Fifteen weeks post-injection representative images were taken of tumors in live mice, and tumor area calculated. Tumors containing the *c-Myb* knock-down line were significantly larger than controls (Figure 3.3E, p=0.02).

***c-Myb* Expression Signature Identifies Many Luminal/ER+ Subtype Defining Genes.** In order to identify the transcriptional targets of c-Myb in breast cells, Agilent microarrays were used to assess gene expression differences between MCF7 shMYB (n=5) versus MCF7 shGFP (n=6). In a two-class Significance Analysis of Microarrays (SAM) analysis (Tusher *et al.*, 2001), 1,211 significantly differentially expressed genes were identified using a false discovery rate (FDR) of less than 5% (data not shown), henceforth termed the "c-Myb gene signature". Several previously identified c-Myb target genes were on this list including *KIT*, *DHRS2* (*Hep27*), and *EMP2* (Hogg *et al.*, 1997; Lei *et al.*, 2004; Rushton *et al.*, 2003).

c-Myb is an estrogen receptor target gene (Gudas *et al.*, 1995), and is repeatedly observed as being highly expressed in the luminal "intrinsic" gene set (Oh *et al.*, 2006; Usary *et al.*, 2004). To determine if genes within the *c-Myb* gene signature overlap with the luminal intrinsic gene set, we used the 232 breast tumor microarray data of Herschkowitz *et al.*, 2007, clustered the tumors using the ~2000 intrinsic gene list described by Parker *et al.*, 2009, and defined the luminal cluster as genes highly correlated (0.65 node correlation; 79 genes total) with *ESR1*, a central gene in the luminal cluster (Figure 3.4 and Table 3.3).

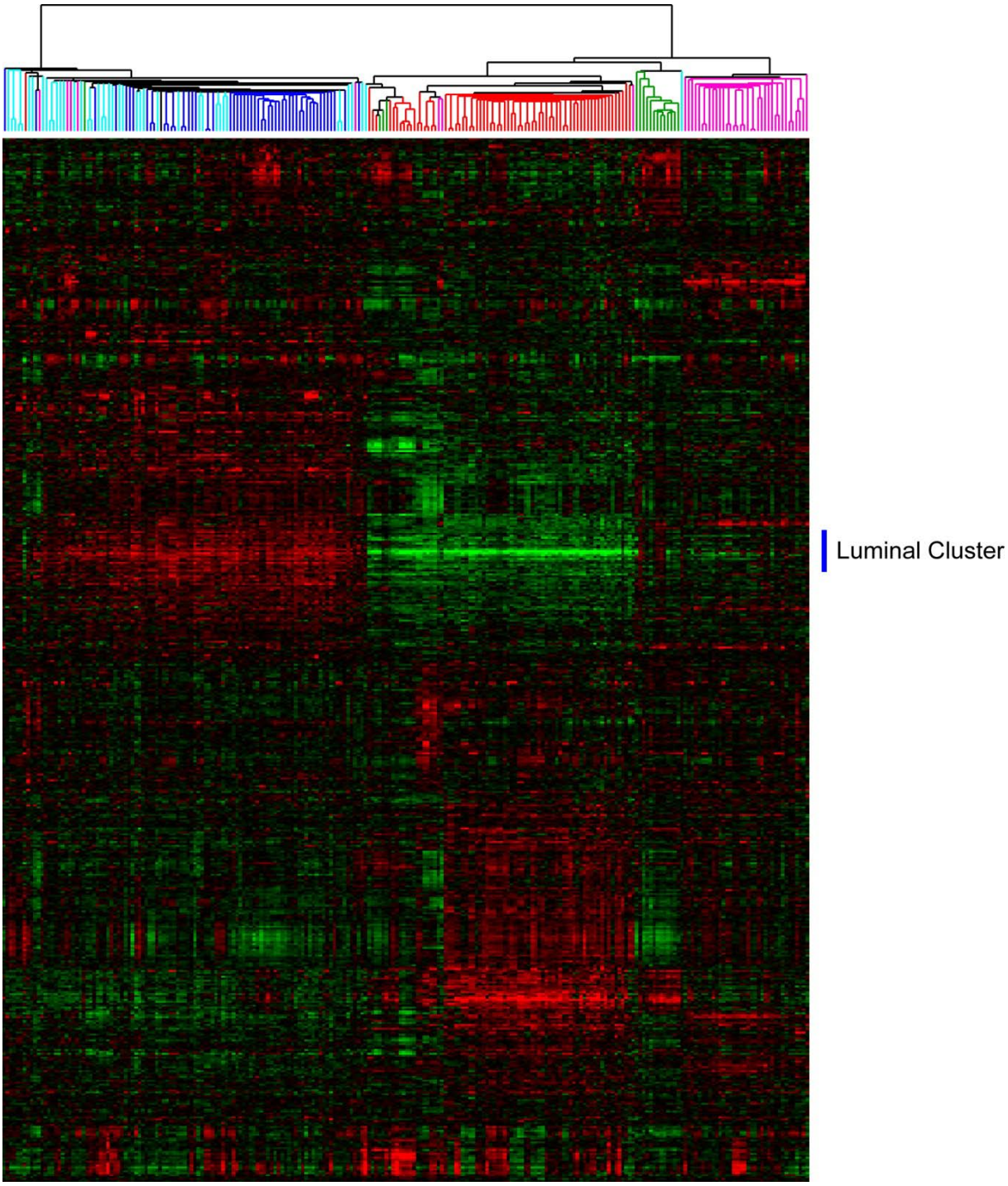


Figure 3.4. UNC breast tumor microarray dataset (n=232; Herschkowitz *et al.*, 2007) clustered using an intrinsic gene set (Parker *et al.*, 2009). The luminal gene cluster, identified as genes highly correlated with the ESR1 gene node (0.65 node correlation; 79 genes) is displayed on the right. Dendrogram branches are colored by subtype: luminal A: dark blue, luminal B: light blue, basal-like: red, normal-like: green, HER-2 enriched: pink.

Table 3.3. c-Myb, ER, and GATA3 gene signature define many luminal cluster genes. Gene names and GenBank accession numbers of the 79 genes highly correlated with the ESR1 gene node (0.65 node correlation). Overlapping genes with *c-Myb*-, *GATA3*- and *ER*-gene signatures are listed.

GenBank Accession Number	Luminal Cluster-Defining Genes	Signature Overlap
NM_030627	CPEB4	c-Myb
NM_003479	PTP4A2	ER
NM_017423	GALNT7	
NM_006379	SEMA3C	
NM_004694	SLC16A6	
NM_004374	COX6C	c-Myb
NM_175605	TTC10	
NM_006460	HIS1	
NM_015541	LRIG1	c-Myb
NM_000633	BCL2	
NM_019000	FLJ20152	c-Myb
NM_016121	KCTD3	c-Myb
NM_006996	SLC19A2	
NM_033211	LOC90355	c-Myb
NM_006113	VAV3	
NM_145252	LOC124220	GATA3
NM_130898	CREB3L4	
NM_020775	KIAA1324	c-Myb
AK000820	ZNF587	
NM_024762	ZNF552	
NM_144611	MGC32124	
NM_173075	APBB2	
NM_016613	DKFZp434L142	c-Myb
NM_006726	LRBA	
NM_206927	SYTL2	
NM_152400	FLJ39370	
BC071555	IL6ST	
NM_024863	TCEAL4	
NM_001012979		
NM_001006640	TCEAL1	
NM_014399	TM4SF13	GATA3
NM_014373	GPR160	GATA3
NM_080759	DACH1	
NM_017786	FLJ20366	c-Myb
NM_012319	SLC39A6	c-Myb
NM_207446	LOC400451	GATA3
BC023981	CA12	c-Myb+ER+GATA3

GenBank Accession Number	Luminal Cluster-Defining Genes	Signature Overlap
NM_003226	TFF3	GATA3
NM_005080	XBP1	c-Myb+ER+GATA3
NM_000044	AR	c-Myb
NM_004496	FOXA1	GATA3
NM_006408	AGR2	c-Myb+GATA3
NM_176813	BCMP11	
NM_021800	DNAJC12	
NM_001002295	GATA3	GATA3
NM_000125	ESR1	
NM_024817	FLJ13710	
NM_014599	MAGED2	
NM_000507	FBP1	c-Myb
NM_001002837	PIB5PA	
NM_020686	ABAT	
NM_014246	CELSR1	
NM_005375	MYB	c-Myb
NM_003462	DNALI1	c-Myb
NM_016569	TBX3	
NM_020974	SCUBE2	c-Myb
NM_032309	CHCHD5	
NM_174921	LOC201895	
NM_024573	C6orf211	
NM_001609	ACADSB	
NM_001001664	LOC339745	
NM_000015	NAT2	c-Myb
NM_000662	NAT1	c-Myb
NM_019600	FLJ10980	
NM_032780	TMEM25	c-Myb
NM_003866	INPP4B	
NM_018379	FLJ11280	
NM_018000	FLJ10116	ER
BC028374	C1orf34	GATA3
NM_007013	WWP1	
NM_003489	NRIP1	ER
A_23_P14432		
NM_014048	MKL2	
NM_032918	RERG	c-Myb+ER
NM_002513	NME3	
NM_016463	CXXC5	
NM_005264	GFRA1	c-Myb
NM_021814	ELOVL5	c-Myb+ER
NM_014912	CPEB3	

We also analyzed previously published *ER* and *GATA3* gene signatures to determine their luminal cluster contributions (Oh *et al.*, 2006; Usary *et al.*, 2004). The *c-Myb* signature had the largest number of genes overlapping with the luminal cluster (24%), followed by *GATA3* (10%) and *ER* (4%) signature genes, and unique combinations of the signatures accounting for 6% (Table 3.3). These data suggest that the Luminal/ER+ cluster is a combination of the effects of multiple transcription factors, with c-Myb being a major contributor.

***c-Myb* Gene Signature Correlates with Mature Luminal Mammary Cell Lineage.** Recently Lim et al. used florescence activated cell sorting (FACS) of normal, human mammary tissue to isolate four discrete cell populations within the mammary gland hierarchy (i.e. stromal, mammary stem cell, luminal progenitor, and mature luminal populations) (Lim *et al.*, 2009). Unique gene expression signatures for each of the mammary gland lineages were revealed by microarrays, and these signatures compared to our cell line-derived *ER*, *GATA3*, and *c-Myb* gene signatures (Figure 3.5, right graphs). The *ER* gene signature was significantly correlated with both the luminal progenitor and mature luminal populations, whereas the *GATA3* gene signature was highly correlated with the mature luminal population, and to a lesser extent, the luminal progenitor lineage (Figures 3.5A and 3.5B, respectively). The differentiated, mature luminal lineage displayed high c-Myb regulated activity that was not observed in the other cell populations, including the mammary stem cell population (Figure 3.5C).

Tumor gene expression profiles were processed in the same fashion as the sorted cell lines and tumor profiles were compared to the *ER*, *GATA3*, and *c-Myb* gene signatures to demonstrate which breast tumor subtypes were enriched for these signatures (Figure 3.5, left

Figure 3.5.

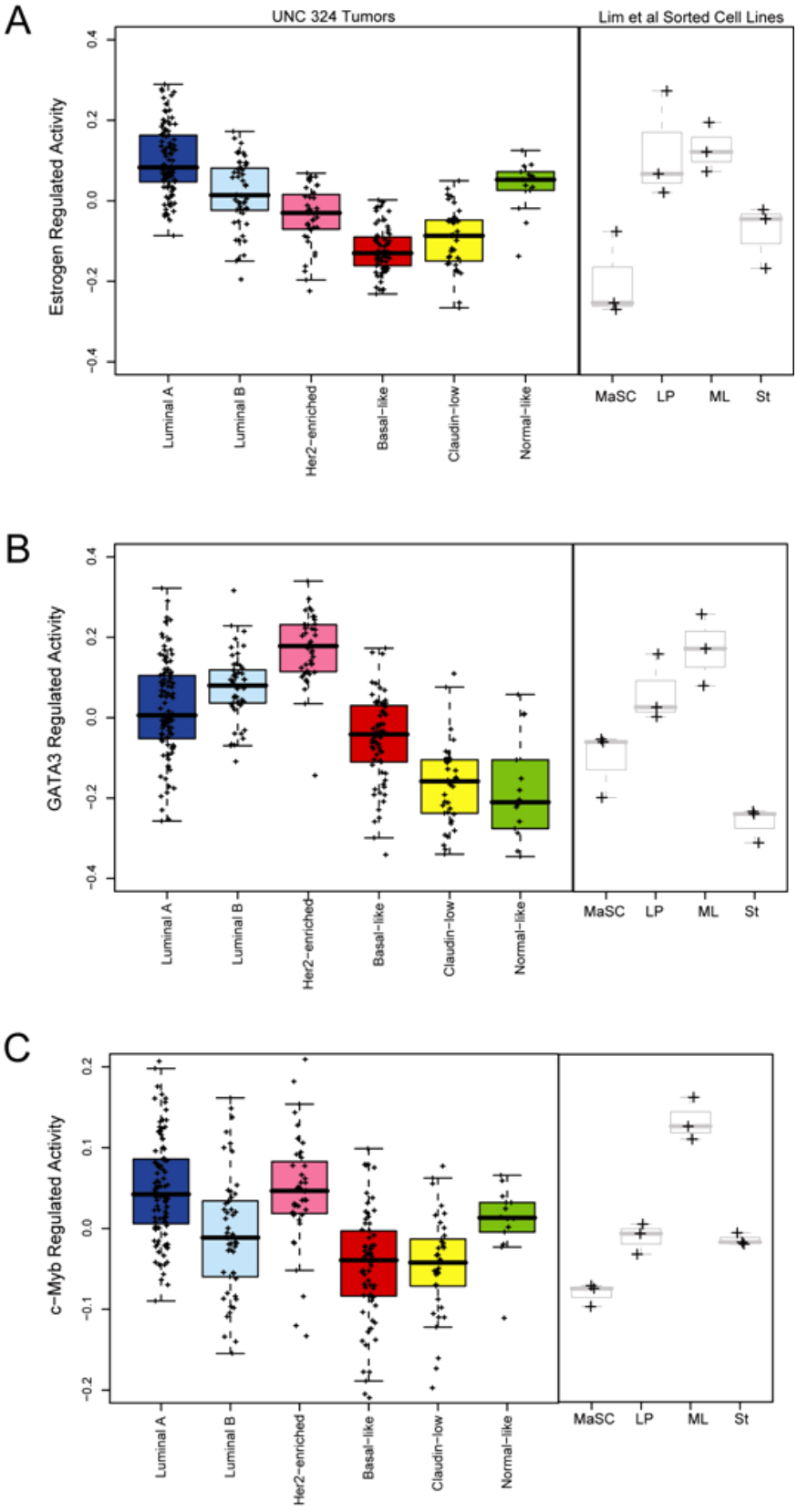

A
UNC 324 Tumors
Lim et al Sorted Cell Lines
Estrogen Regulated Activity
0.4
0.2
0.0
−0.2
−0.4
Luminal A
Luminal B
Her2−enriched
Basal−like
Claudin−low
Normal−like
MaSC
LP
ML
St
B
GATA3 Regulated Activity
0.4
0.2
0.0
−0.2
−0.4
Luminal A
Luminal B
Her2−enriched
Basal−like
Claudin−low
Normal−like
MaSC
LP
ML
St
C
c−Myb Regulated Activity
0.2
0.1
0.0
−0.1
−0.2
Luminal A
Luminal B
Her2−enriched
Basal−like
Claudin−low
Normal−like
MaSC
LP
ML
St

Figure 3.5. *c-Myb* gene signature correlates with mature (differentiated) luminal mammary cells. (A) ER-, (B) GATA3-, and (C) c-Myb- regulated activities across UNC breast tumors (n=324; GSE10893) or mammary gland cell lineage gene signatures (Lim *et al.*, 2009). Tumor gene expression was weighted according to *ER*, *GATA3*, or *c-Myb* gene signatures by calculating the inner product of each signature and the tumor expression profile. Lim *et al*. lineage signatures were processed in the same fashion as tumors to demonstrate the developmental context of these signatures.
MaSC: mammary stem cell-enriched; LP: luminal progenitor; ML: mature luminal; St: Stromal.

graphs). The *ER* gene signature was highest in the luminal A subtype and lowest in the ER-negative, basal-like tumors (Figure 3.5A). The *GATA3* gene signature was most strongly correlated with the HER2-enriched subtype (Figure 3.5B). The *c-Myb* gene signature was strongly correlated with the Luminal A and HER2-enriched subtypes, and was least correlated with basal-like tumors.

DISCUSSION

The essential role of the *c-Myb* oncogene in leukemogenesis has been appreciated for several decades. Its expression is required for maintenance of both acute and chronic myeloid leukemia cells (Anfossi, 1989; Calabretta, 1991). *c-Myb* is also necessary for normal hematopoiesis; knock-out mice succumb to embryonic lethality (e15) due to unsuccessful blood cell lineage formation (Mucenski *et al.*, 1991), and lineage-specific knockouts have revealed that *c-Myb* is required for proper T-cell differentiation (Bender *et al.*, 2004). Studies in colorectal carcinoma have found high *c-Myb* expression to correlate with poor prognosis (Biroccio *et al.*, 2001), and more recently the fusion of *c-Myb* and the transcription factor *NFIB* has been reported as a potential "hallmark" of adenoid cystic carcinomas (Persson *et al.*, 2009). Based on these data it has been hypothesized that *c-Myb* is necessary for progenitor cell maintenance, and decreasing c-*Myb* expression is crucial for cellular differentiation (Ramsay and Gonda, 2008).

Our data in breast tumors, however, suggest a tumor suppressor role for c-MYB where high expression portents a good outcome and high differentiation status of tumors. It is known that *c-Myb* is a direct target of the estrogen receptor alpha (ER), and c-Myb protein levels are increased in ER+ tumors (Guérin M, 1990). Here, our findings confirm that *c-Myb*

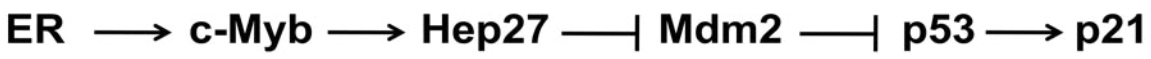
high expression correlates with the ER+, Luminal A and B subtypes of breast cancer (Figure 3.1A). Unlike the other tumor types discussed above, breast tumors having high *c-Myb* expression levels correspond with a good prognosis, even within luminal breast tumors alone (Figures 3.1A and B). We purposefully used a microarray dataset where patients received no adjuvant chemotherapy so as not to confound the survival data with the use of chemotherapeutics, particularly estrogen antagonists.

To analyze the role of c-Myb in breast tumorigenesis, we chose to knock-down endogenous c-Myb levels in MCF7 cells, which is a luminal breast tumor cell line (Neve *et al.*, 2006). When c-Myb protein levels were significantly decreased relative to the parental cell line, we observed faster cell growth, decreased tamoxifen sensitivity, and increased colony formation *in vitro* (Figures 3.3B-D). These observations were preserved *in vivo* where MCF7 cells, with c-Myb knocked-down, grew significantly faster in nude mice than MCF7 controls (Figure 3.3E). Many of the control tumors did not grow or disappeared by 15 weeks post-injection. This suggests that the *c-Myb* knock-down cell line is less estrogen-dependent than the parental line given that no exogenous hormone was implanted in these mice. This result also correlated with the observed tamoxifen resistance in the c-Myb knock-down cells (Table 3.1).

A recent study found that the estrogen receptor can directly relieve transcriptional attenuation in the first intron of *c-Myb*, thus giving greater insight into how ER regulates *c-Myb* expression (Drabsch *et al.*, 2007). The authors proposed that c-Myb is required for the proliferation of MCF7 cells because they observed decreased proliferation in c-Myb knock-down MCF7 cells versus controls when treated with β-estradiol. These results are contradictory to ours, where we observed increased growth and tumorigenesis in the c-Myb

knock-down line. This may be in part due to our use of a more estrogen sensitive MCF7 cell line; unlike the cell line used by Drabsch *et al.*, our MCF7 isolate requires estrogen to grow and in the absence of hormone (i.e. phenol-red free and charcoal stripped FBS media) they do not proliferate. We were also able to show that the c-Myb knock-down cell line forms more colonies *in vitro* and grows faster *in vivo*, data not shown by these authors.

Our results from both human breast tumor microarray data and c-Myb knock-down in MCF7 cells suggest a novel tumor suppressor role for c-Myb in breast cancer. In collaboration with Chad Deisenroth and Yanping Zhang (University of North Carolina), we identified a c-Myb target gene, *Hep27* (*DHRS2*) (Rushton *et al.*, 2003), as a novel regulator of the p53 pathway (Deisenroth, Thorner *et al.*, submitted). The Hep27 protein inhibits Mdm2, a well-known inhibitor of p53, thereby stabilizing the p53 protein (see pathway diagram below).



In MCF7 cells, when estrogen was added to the media, we observed a corresponding increase in both c-Myb and Hep27 (Figure 3.6A). When *c-Myb* was exogenously expressed at increasing levels, there was a corresponding increase in Hep27 and p53 protein levels (Figure 3.6B). Conversely, when c-Myb levels were depleted via shRNA, there were corresponding decreases in both Hep27 and p53 protein (Figure 3.6C). Taken together, these results suggest the existence of this c-Myb-Hep27-p53 pathway in MCF7 cells.

The SAM analysis of our MCF7 c-Myb knock-down line versus parental line showed that *Hep27*'s expression is significantly reduced in the knock-down line (data not shown).

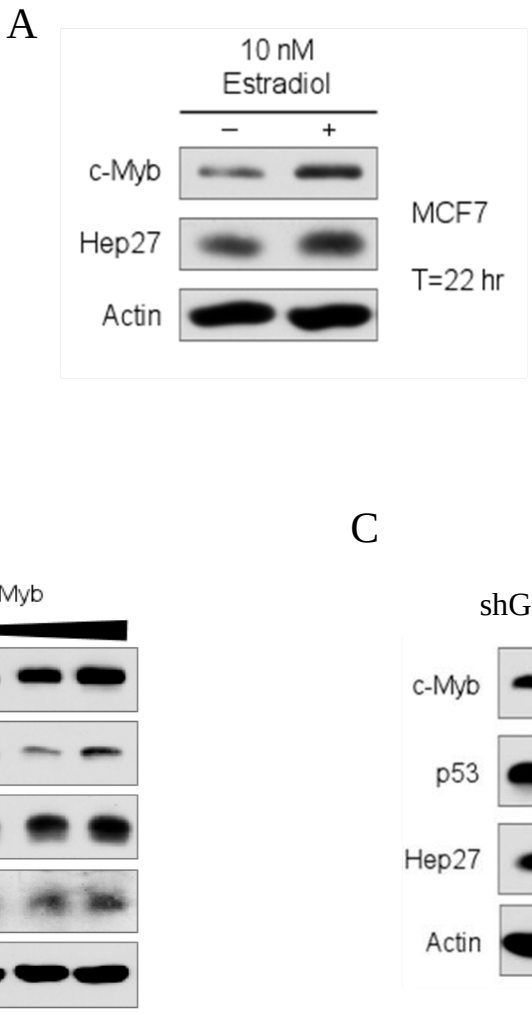


Figure 3.6. c-Myb induces p53 stabilization and activation in a Hep27 dependent manner. (A) 10 nM estradiol was added to media for 22 hours and cell lysate collected for western blot analysis. (B) Expression construct expressing c-Myb was transfected into MCF7 cells for 24 hours and indicated proteins were detected by western blot. (C) Western blot for c-Myb, p53, and Hep27 protein levels in MCF7 cells with c-Myb knock-down (shMYB) or control (shGFP).

Figure 3.7.

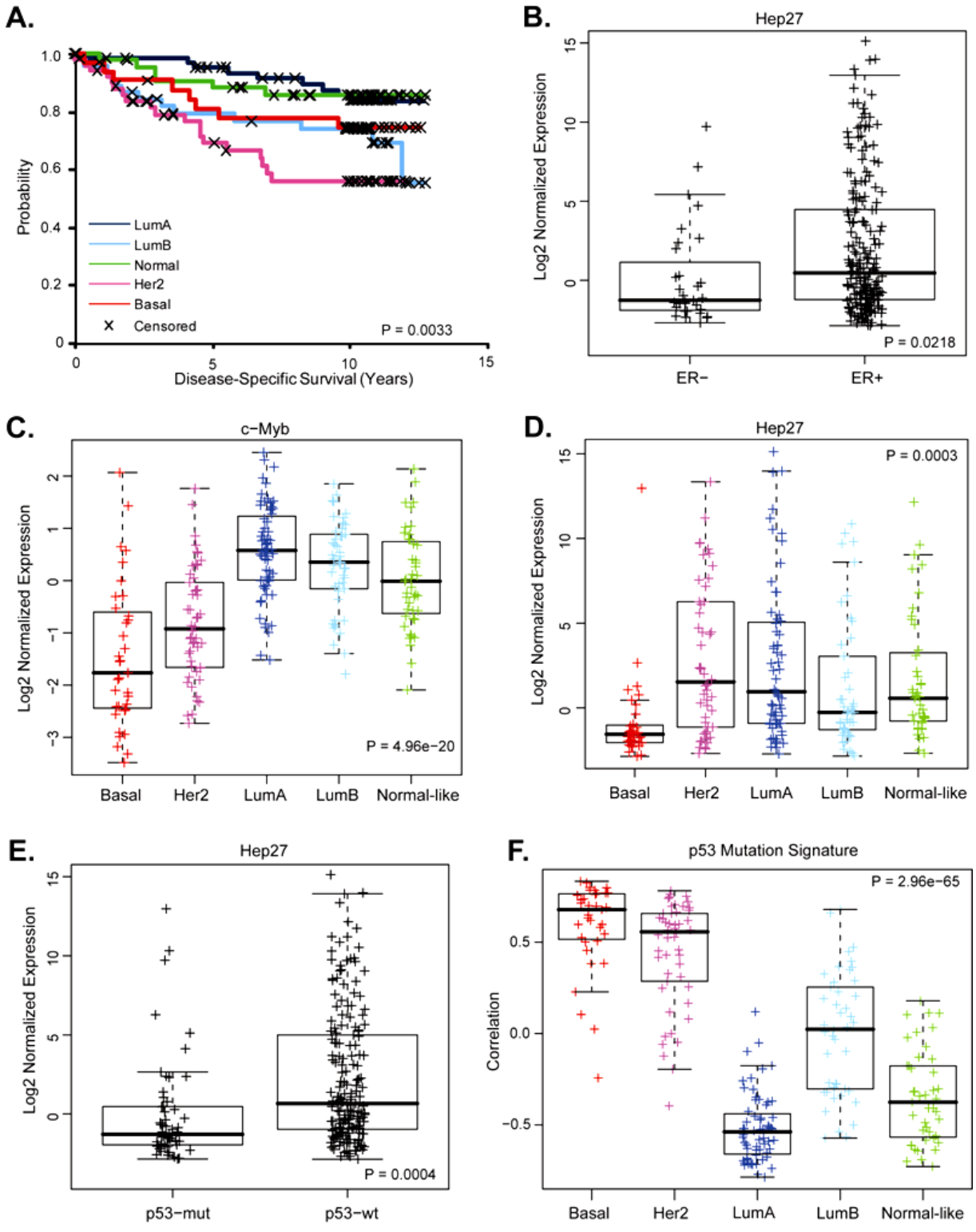

A.
Probability
1.0
0.8
0.6
0.4
0.2
0
LumA
LumB
Normal
Her2
Basal
Censored
P = 0.0033
0
5
10
15
Disease-Specific Survival (Years)
B.
Hep27
Log2 Normalized Expression
15
10
5
0
P = 0.0218
ER−
ER+
C.
c−Myb
Log2 Normalized Expression
2
1
0
−1
−2
−3
P = 4.96e−20
Basal
Her2
LumA
LumB
Normal-like
D.
Hep27
Log2 Normalized Expression
15
10
5
0
P = 0.0003
Basal
Her2
LumA
LumB
Normal-like
E.
Hep27
Log2 Normalized Expression
15
10
5
0
P = 0.0004
p53−mut
p53−wt
F.
p53 Mutation Signature
Correlation
0.5
0.0
−0.5
P = 2.96e−65
Basal
Her2
LumA
LumB
Normal-like

Figure 3.7. Potential c-Myb-Hep27-Mdm2-p53 pathway in luminal breast cancer. The Miller *et al*., 2005 dataset (n=236) was classified into the breast tumor intrinsic subtypes (Basal-like, HER2-enriched, Luminal A, Luminal B, and Normal-like) using the PAM50 predictor (Parker *et al*., 2009). (A) Kaplan-Meier survival analysis of disease-specific survival stratified by subtype. p-value determined by log-rank test, testing the null hypothesis that the survival curves are identical across the subtypes. (B) Hep27 mRNA expression in ER- and ER+ tumors (n=232). p-values calculated by t-test showing different expression values across ER status or biologically defined breast tumor subtypes. (C) c-Myb and (D) Hep27 expression varies by intrinsic subtype. (E) Hep27 expression in p53-mutant and p53-wild-type tumors. (F) A p53-mutation signature (Troester *et al*., 2006) was applied to this dataset and correlated with breast tumor subtype. For figures C-F, p-values were calculated by ANOVA, testing the null hypothesis that all group means are equal.

Again, this suggests that in luminal tumors the c-Myb-Hep27-p53 pathway may be intact and, therefore, tumors with higher levels of c-Myb will correspondingly have higher levels of stabilized, wild-type tumor suppressor p53. To further investigate this, we analyzed the breast tumor microarray data of Miller *et al.* (Miller *et al.*, 2005). The tumors in this dataset are representative of all breast tumor subtypes and their prognosis (overall survival) is similar to patients in other breast tumor microarray datasets (Figure 3.7A). We observed higher *c-Myb* and *Hep27* levels in ER+, p53 wild-type tumors, both common features of the Luminal breast tumor subtype (Figures 3.7B-D). We also found that p53 wild-type tumors expressed higher levels of *Hep27* (Figure 3.7E), and luminal tumors are least likely to be p53 mutant (Figure 3.7F). Luminal tumors have been repeatedly observed as chemotherapy resistant (Parker *et al.*, 2009; Rouzier *et al.*, 2005), but until now the potential mechanism was unknown. Here, our results showing *c-Myb* high-expressing tumors have poor pathological complete response to chemotherapy (Table 3.1) may be due to an intact c-Myb-Hep27-p53 pathway in the luminal subtypes.

The identification of a *c-Myb* signature was also informative from a possible developmental perspective. As compared with the luminal tumor defining gene signature, we observed that more genes from the *c-Myb* gene list overlapped this cluster than other luminal defining gene lists (*GATA3* and *ER*). This is suggestive that the c-Myb transcription factor is potentially regulating many genes in luminal tumors.

During hematopoiesis the expression of *c-Myb* is highest in progenitor cell lineages and is down-regulated during differentiation (Ramsay and Gonda, 2008). We used a recent study that isolated four cell lineages in the normal mammary gland (Lim *et al.*, 2009): stromal, mammary stem cell, luminal progenitor, and mature luminal populations.

Comparisons of our *c-Myb* gene signature, as well as the previously published *GATA3* and *ER* gene signatures, to the Lim *et al*. data showed the opposite result for c-Myb in the mammary lineage when compared to the hematopoietic linage. Namely, the *c-Myb* gene signature significantly correlated with the mature luminal cell population, and was least expressed within the mammary stem cell enriched population (Figure 3.5). In addition, the estrogen-regulated activity was highest in both luminal progenitor and mature luminal cells, while the GATA3-regulated activity was also highest in the mature luminal population, but was also increased in both the luminal progenitor and mammary stem cell populations (Figure 3.5C).

Directly targeting c-Myb as a form of cancer therapy has been suggested and implemented in several tumor types (Calabretta, 1991; Ratajczak, 1992). Based on positive results, an antisense oligonucleotide targeting c-Myb has been developed as a targeted therapeutic and a Phase I clinical trial begun for patients with advanced hematologic malignancies (National Clinical Trials Identifier: NCT00780052). It has been suggested that this form of treatment could be of value in patients with other cancers expressing high *c-Myb*, including breast tumors (Ramsay and Gonda, 2008). However, our findings indicate c-Myb is not behaving as an oncogene in ER+, luminal breast tumors, which is the most common form of human breast cancer. Rather, *c-Myb* in the mammary gland is being expressed in the mature luminal cell population and acting in a pathway to stabilize the tumor suppressor, p53. Therefore, high *c-Myb* expression is beneficial in luminal breast cancer and reducing c-Myb protein levels via antisense therapy could be detrimental. In total, our current findings have yielded unique insights into the role of c-Myb in luminal breast cancer and suggest that it is behaving as a tumor suppressor in this disease.

ACKNOWLEDGEMENTS

We thank Drs. Jerry Usary and J. Chuck Harrell for technical assistance. This work was supported by funds from the NCI Breast SPORE program to UNC-CH (P50-CA58223-09A1), by the NCI (RO1-CA-101227), the Breast Cancer Research Foundation and by the V Foundation for Cancer Research.

REFERENCES

Anfossi G, Gewirtz A.M., and Calabretta, B (1989). An oligomer complementary to c-myb-encoded mRNA inhibits proliferation of human myeloid leukemia cell lines. *Proc Natl Acad Sci* **86:** 3379-3383.

Bender TP, Kremer CS, Kraus M, Buch T, Rajewsky K (2004). Critical functions for c-Myb at three checkpoints during thymocyte development. *Nat Immunol* **5:** 721-9.

Biroccio A, Benassi B, D'Agnano I, D'Angelo C, Buglioni S, Mottolese M *et al* (2001). c-Myb and Bcl-x Overexpression Predicts Poor Prognosis in Colorectal Cancer : Clinical and Experimental Findings. *Am J Pathol* **158:** 1289-1299.

Calabretta B, RB Sims, M Valtieri, D Caracciolo, C Szczylik, D Venturelli, M Ratajczak, M Beran, and AM Gewirtz (1991). Normal and leukemic hematopoietic cells manifest differential sensitivity to inhibitory effects of c-myb antisense oligodeoxynucleotides: an in vitro study relevant to bone marrow purging. *Proc Natl Acad Sci* **88:** 2351-2355.

Drabsch Y, Hugo H, Zhang R, Dowhan DH, Miao YR, Gewirtz AM *et al* (2007). Mechanism of and requirement for estrogen-regulated MYB expression in estrogen-receptor-positive breast cancer cells. *Proceedings of the National Academy of Sciences* **104:** 13762-13767.

Ess KC, David P Witte, Charles P Bascomb and Bruce J Aronow (1999). Diverse developing mouse lineages exhibit high-level c-Myb expression in immature cells and loss of expression upon differentiation. *Oncogene* **18:** 1103-1111.

Gudas JM, Klein RC, Oka M, Cowan KH (1995). Posttranscriptional regulation of the c-myb proto-oncogene in estrogen receptor-positive breast cancer cells. *Clin Cancer Res* **1:** 235-243.

Guérin M SZ, Andrieu N, Riou G. (1990). Strong association between c-myb and oestrogen-receptor expression in human breast cancer. *Oncogene* **5:** 131-5.

Harrell JC, Dye WW, Allred DC, Jedlicka P, Spoelstra NS, Sartorius CA *et al* (2006). Estrogen Receptor Positive Breast Cancer Metastasis: Altered Hormonal Sensitivity and Tumor Aggressiveness in Lymphatic Vessels and Lymph Nodes. *Cancer Res* **66:** 9308-9315.

Herschkowitz J, Simin K, Weigman V, Mikaelian I, Usary J, Hu Z *et al* (2007). Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. *Genome Biology* **8:** R76.

Hess KR, Anderson K, Symmans WF, Valero V, Ibrahim N, Mejia JA *et al* (2006). Pharmacogenomic Predictor of Sensitivity to Preoperative Chemotherapy With Paclitaxel and Fluorouracil, Doxorubicin, and Cyclophosphamide in Breast Cancer. *J Clin Oncol* **24:** 4236-4244.

Hogg A, Schirm S, Nakagoshi H, Bartley P, Ishii S, Bishop JM *et al* (1997). Inactivation of a c-Myb/estrogen receptor fusion protein in transformed primary cells leads to granulocyte/macrophage differentiation and down regulation of c-*kit* but not c-*myc* or *cdc2*. **15:** 2885-2898.

Hu Z, Fan C, Oh D, Marron J, He X, Qaqish B *et al* (2006). The molecular portraits of breast tumors are conserved across microarray platforms. *BMC Genomics* **7:** 96.

Kauraniemi P, Hedenfalk I, Persson K, Duggan DJ, Tanner M, Johannsson O *et al* (2000). MYB Oncogene Amplification in Hereditary BRCA1 Breast Cancer. *Cancer Res* **60:** 5323-5328.

Klempnauer K-H, Gonda TJ, Bishop MJ (1982). Nucleotide sequence of the retroviral leukemia gene v-myb and its cellular progenitor c-myb: The architecture of a transduced oncogene. *Cell* **31:** 453-463.

Kopecki Z MML, D. H. Adams, X. Strudwick, T. Mantamadiotis, A. Stoppacciaro, A. Gabrielli, R. G. Ramsay, A. J. Cowin (2007). Collagen loss and impaired wound healing is associated with c-Myb deficiency. *The Journal of Pathology* **211:** 351-361.

Lei W, Rushton JJ, Davis LM, Liu F, Ness SA (2004). Positive and Negative Determinants of Target Gene Specificity in Myb Transcription Factors. *J. Biol. Chem.* **279:** 29519-29527.

Leprince D, Gegonne A, Coll J, de Taisne C, Schneeberger A, Lagrou C *et al* (1983). A putative second cell-derived oncogene of the avian leukaemia retrovirus E26. *Nature* **306:** 395.

Lim E, Vaillant F, Wu D, Forrest NC, Pal B, Hart AH *et al* (2009). Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. *Nat Med* **15:** 907-913.

Malaterre J, Carpinelli M, Ernst M, Alexander W, Cooke M, Sutton S *et al* (2007). c-Myb is required for progenitor cell homeostasis in colonic crypts. *PNAS* **104:** 3829-3834.

Malaterre J, Mantamadiotis T, Dworkin S, Lightowler S, Yang Q, Ransome MI *et al* (2008). c-Myb Is Required for Neural Progenitor Cell Proliferation and Maintenance of the Neural Stem Cell Niche in Adult Brain. *Stem Cells* **26:** 173-181.

Miller LD, Smeds J, George J, Vega VB, Vergara L, Ploner A *et al* (2005). An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival. *Proceedings of the National Academy of Sciences* **102:** 13550-13555.

Mucenski ML, McLain K, Kier AB, Swerdlow SH, Schreiner CM, Miller TA *et al* (1991). A functional c-myb gene is required for normal murine fetal hepatic hematopoiesis. *Cell* **65:** 677.

Neve RM, Chin K, Fridlyand J, Yeh J, Baehner FL, Fevr T *et al* (2006). A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. *Cancer Cell* **10:** 515-527.

Oh D, Troester M, Usary J, Hu Z, He X, Fan C *et al* (2006). Estrogen-regulated genes predict survival in hormone receptor-positive breast cancers. *J Clin Oncol* **24:** 1656 - 1664.

Oh IH, Reddy EP (1999). The *myb* gene family in cell growth, differentiation and apoptosis. *Oncogene* **18:** 3017-33.

Parker JS, Mullins M, Cheang MCU, Leung S, Voduc D, Vickery T *et al* (2009). Supervised Risk Predictor of Breast Cancer Based on Intrinsic Subtypes. *J Clin Oncol* **27:** 1160-1167.

Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA *et al* (2000). Molecular portraits of human breast tumours. *Nature* **406:** 747.

Persson M, Andren Y, Mark J, Horlings HM, Persson F, Stenman G (2009). Recurrent fusion of MYB and NFIB transcription factor genes in carcinomas of the breast and head and neck. *Proceedings of the National Academy of Sciences***:** -.

Ramsay RG, Gonda TJ (2008). MYB function in normal and cancer cells. *Nat Rev Cancer* **8:** 523.

Ratajczak M, JA Kant, SM Luger, N Hijiya, J Zhang, G Zon, and AM Gewirtz (1992). In vivo treatment of human leukemia in a scid mouse model with c-myb antisense oligodeoxynucleotides. *PNAS* **89:** 11823-11827.

Riggs BL, Hartmann LC (2003). Selective Estrogen-Receptor Modulators -- Mechanisms of Action and Application to Clinical Practice. *N Engl J Med* **348:** 618-629.

Rouzier R, Perou CM, Symmans WF, Ibrahim N, Cristofanilli M, Anderson K *et al* (2005). Breast Cancer Molecular Subtypes Respond Differently to Preoperative Chemotherapy. *Clin Cancer Res* **11:** 5678-5685.

Rushton JJ, Davis LM, Lei W, Mo X, Leutz A, Ness SA (2003). Distinct changes in gene expression induced by A-Myb, B-Myb and c-Myb proteins. *Oncogene* **22:** 308.

Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H *et al* (2001). Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *PNAS* **98:** 10869-10874.

Thorner AR, Hoadley KA, Parker JS, Winkel S, Millikan RC, Perou CM (2009). In vitro and in vivo analysis of B-Myb in basal-like breast cancer. *Oncogene* **28:** 742-51.

Tusher VG, Tibshirani R, Chu G (2001). Significance analysis of microarrays applied to the ionizing radiation response. *PNAS* **98:** 5116-5121.

Usary J, Llaca V, Karaca G, Presswala S, Karaca M, He X *et al* (2004). Mutation of GATA3 in human breast tumors. *Oncogene* **23:** 7669-7678.

van de Vijver MJ, He YD, van 't Veer LJ, Dai H, Hart AAM, Voskuil DW *et al* (2002). A Gene-Expression Signature as a Predictor of Survival in Breast Cancer. *N Engl J Med* **347:** 1999-2009.

CHAPTER IV

CONCLUSION

Over the past decade our understanding of breast cancer at the molecular level has significantly improved due, in part, to the identification of the breast tumor intrinsic subtypes by gene expression microarrays and other advances in genetics and cell biology. The intrinsic subtypes of breast cancer have been recapitulated in many labs using differing RNA collection techniques and varying microarray platforms (Herschkowitz *et al.*, 2007; Hu *et al.*, 2006; Parker *et al.*, 2009; Perou *et al.*, 2000; Sorlie *et al.*, 2001; Sorlie *et al.*, 2003). To date, at least six breast cancer subtypes have been identified including luminal A, luminal B, normal breast-like, HER2-enriched, basal-like, and claudin-low. The subtypes not only have biological relevance but also clinical relevance, as each tumor type has unique clinical outcomes and response to treatment (Carey *et al.*, 2007; Fan *et al.*, 2006). Several dominant genes drive the classification of the luminal and HER2-enriched tumor types, namely the expression of estrogen receptor/ESR1 (luminal), and amplification and high expression of HER2/ERBB2. Other subtypes are defined by their lack of gene expression, such as in the claudin-low subtype where low expression of claudin genes 3, 4, and 7 is observed. The basal-like subtype lacks expression of both the estrogen and progesterone receptors, as well as HER2.

The molecular subtypes are defined by an intrinsic gene set (Hu *et al.*, 2006; Perou *et al.*, 2000). The expression of these genes varies across the subtypes, but does not vary within replicates of the same patient sample. Of the hundreds of genes that are used to classify each subtype, relatively few have been well defined in terms of their contribution to breast tumorigenesis. Therefore, my research focused on the MYB family of transcription factors, which is a family comprised of three genes all found within the intrinsic gene set (Hu *et al.*, 2006). Specifically, my goal was to evaluate the *in vitro* and *in vivo* function of *B-Myb* and *c-Myb* (two MYB family members with implications in tumorigenesis in other tissue types) in the mammary gland.

B-Myb Gene in Breast Cancer Biology

High expression of *B-Myb* was observed in the proliferation gene signature, which is an easily identifiable and dominant gene signature that is a marker of tumor cell proliferation rates (Whitfield *et al.*, 2002). The proliferation genes are highly expressed in the both the luminal B and basal-like subtypes, and the fast proliferation of these tumor types may contribute to their poor prognoses. Basal-like tumors have no known molecular therapeutic targets and at this time can only be treated with cytotoxic chemotherapy. Therefore, I chose to focus on how *B-Myb* expression may be contributing to this relatively ill-defined subtype.

Using several large breast tumor microarray datasets we found *B-Myb* to be significantly highly expressed in the basal-like subtype followed by next highest expression in the luminal B and HER2-enriched subtypes, which are all tumor subtypes with poor outcomes. We also determined that *B-Myb* expression levels alone were able to predict overall survival across all breast cancer subtypes as well as within the luminal A, luminal B,

and HER2-enriched tumors. *B-Myb* expression levels were not capable of predicting survival within the basal-like tumors alone, however, patients with this disease have very poor outcomes and virtually all express high levels of this protein.

These results suggest that *B-Myb* expression is not only relevant in basal-like disease, but levels of *B-Myb* may play an important role in all subtypes. For example, the luminal A tumors have the lowest expression of *B-Myb* relative to the other subtypes, yet the overall survival of patients with this subtype could be predicted using univariate analysis by *B-Myb* gene expression, suggesting that higher levels of *B-Myb* portends worse outcome.

High expression levels of *B-Myb* in basal-like disease has not been linked to genomic amplification (Bergamaschi *et al.*, 2006). This group identified amplification of the *B-Myb* locus (20q13) in luminal B tumors, but not in other breast cancer subtypes. This indicates that the high expression of *B-Myb* in luminal tumors may be selected for and leads to poor outcomes, whereas in basal-like tumors its high expression is the result of other mechanisms, such as amplification of transcription factor(s) targeting *B-Myb*, enhanced promoter activity, or because of the inherent high proliferation rate of this tumor type.

B-Myb expression levels were also predictive of pathologic complete response. We observed that patients with higher levels of *B-Myb* expression had a better response to chemotherapy. At first glance this may appear counterintuitive, whereby tumors having the poorest outcomes have the best response to chemotherapy treatment. This unusual finding has been observed before and been dubbed the "triple negative (basal-like) tumor paradox" (Carey *et al.*, 2007). The explanation for this paradox is that many highly proliferative basal-like tumors (that show high B-Myb and proliferation-signature expression) are in fact chemotherapy sensitive, and these patients respond and have a good outcome; however, the

majority of basal-like patients (whose tumors are also highly proliferative) do not completely respond and have residual disease, and they have a high likelihood of recurrence, and in fact, they do recur at a high frequency. Thus, some patients respond and do well, while most do poorly, thus the paradox.

Recently, a *B-Myb* germline variant causing a non-synonymous serine to glycine amino acid change at residue 427 (S427G, rs2070235) was identified as decreasing overall cancer risk when combining a variety of tumor types and comparing to non-cancer bearing controls (Schwab *et al.*, 2007). However, using a population-based case-control study from the Carolina Breast Cancer Study, we identified women carrying two copies of this variant and determined that they were at a two-fold increased risk of basal-like breast cancer, but not other breast tumor subtypes. The disagreement between these two findings may be due to the different tumor types studied, or because the other study combined a variety of tumors whereas we studied breast tumors only. We did not observe any increase or decreased risk in breast cancer as a whole and it wasn't until we separated tumors into subtypes that we were able to observe the increased risk in only the basal-like subtype.

Our findings mark the first published basal-like breast cancer risk variant. It should also be noted that the frequency of the *B-Myb* variant is approximately 10-fold higher in African-Americans compared with Caucasians. This is especially relevant in light of recent data showing that premenopausal African-Americans are approximately twice as likely to develop basal-like tumors compared with premenopausal Caucasians (Carey *et al.*, 2006; Millikan *et al.*, 2008). It may be of potential benefit for African-Americans to be genotyped for this variant to identify carriers who would be ideal candidates for early and frequent

breast cancer screenings with the goal of earlier detection and diagnosis of basal-like breast cancer.

To evaluate the function of *B-Myb*'s role in chemotherapy response we used *in vitro* models of normal and tumor-derived mammary cell lines. Normal mammary tissue derived lines and basal-like tumor derived lines expressing endogenous levels of *B-Myb* were compared to those overexpressing *B-Myb*, or the susceptibility variant. Normal derived lines overexpressing *B-Myb*, or variant, were more sensitive to topoisomerase 2α (TOP2A)-targeting chemotherapeutics, but not to the other drugs tested. TOP2A inhibitors caused the enzyme to become trapped on double-strand DNA breakpoints, thereby causing G2 cell cycle checkpoint arrest. *B-Myb* has been identified as a G2/M regulating gene and increased *B-Myb* expression may be promoting the TOP2A-inhibitor-treated cells to continue to cycle via the G2 checkpoint through induction of downstream B-Myb target genes. By facilitating cells through G2, with less regard for DNA damage, the B-Myb-overexpressing cells may attempt to cycle again, eventually leading to the observed increased sensitivity to TOP2A phenotype. Using microarray analysis, TOP2A expression was increased in the TOP2A-inhibitor treated cell lines overexpressing *B-Myb*. Therefore, more of the target protein was present in these cells potentially leading to the observed increased sensitivity phenotype. In support of this hypothesis, many known and potential B-Myb target genes were identified by cell cycle analysis and enrichment for G2/M cell cycle genes was observed. This is in agreement with what has been observed in other tissue types, but had not been documented in the mammary epithelial cells.

The results of this study enhance our understanding of the role of B-Myb in breast cancer by identifying new B-Myb target genes, by showing that this gene is highly expressed

in basal-like breast cancers and that it is of prognostic value for survival and predictive value for pathological complete response. B-Myb appears to not only be important in basal-like tumors, where highest expression is observed, but also in luminal tumors, where it was capable of predicting outcomes in these diseases that have relatively good prognosis.

Currently the function of the B-Myb S427G variant is unknown, and we did not observe any gene expression differences between the wild-type and variant B-Myb, nor were differences in chemosensitivity identified. Recent evidence has shown that the variant is more resistant to degradation than wild-type B-Myb (Schwab *et al.*, 2007). It is tempting to speculate that if the half-life of the variant is longer, carriers of this variant have more B-Myb protein in the cell leading to increased cell cycling with less regard to DNA damage. A woman carrying two copies of the variant may, over her lifetime, accumulate a number of genetic mutations in her mammary cells eventually giving rise to basal-like disease. One experiment to test this hypothesis would be to study mouse models of cancer crossed with *B-Myb* variant knock-in mice and determine the effects this gene has on tumorigenesis.

Clearly *B-Myb*, a gene found in the intrinsic gene set, plays a significant role in basal-like disease and potentially in the aggressiveness of other breast tumor subtypes. Its overexpression, combined with the frequent loss of p53 and Rb observed in basal-like disease (Derenzini *et al.*, 2008; Herschkowitz *et al.*, 2008; Sorlie *et al.*, 2001), may begin to explain the basal-like paradox, and the *B-Myb* variant could be of future clinical importance for determining disease risk, especially in African-Americans.

c-Myb Gene in Breast Cancer Biology

The *c-Myb* oncogene has been linked to cancer since its discovery as the mammalian homolog of *v-Myb*, a transforming viral oncogene causing avian leukemia. We observed significant high *c-Myb* expression in the ER-positive, luminal breast tumor subtypes. This is in concordance with previous studies correlating high c-Myb protein levels in ER-positive breast cancer (Guérin M, 1990), and with *c-Myb* being a direct ER target gene (Drabsch *et al.*, 2007). However, the role of c-Myb in ER-positive breast cancer remains virtually unknown.

High *c-Myb* expression has been linked to poor overall survival in a variety of leukemias and colorectal cancer (Biroccio *et al.*, 2001; Ramsay and Gonda, 2008). These results are consistent with the fact that *c-Myb* was identified as an oncogene and we hypothesized that we would observe similar results in breast cancer; however, the results were markedly different. Using multiple different breast tumor microarray datasets with patient survival information, we found that higher levels of *c-Myb* expression conferred better outcomes. In fact, even within the luminal A subtype, a subtype with patients virtually all expressing high levels of *c-Myb*, we were able to stratify outcome based on *c-Myb* expression levels alone, suggesting a dosage effect (i.e. tumors with very high *c-Myb* expression levels portend better patient outcomes than those with high *c-Myb* expression levels). High *c-Myb* expression also correlated with smaller tumor size and lower tumor grade. This data indicated that the *c-Myb* oncogene may be playing a very different role, potentially as a tumor suppressor, in breast cancer compared to the previously studied tumor types.

We chose to stably knock-down endogenous c-Myb protein levels in a luminal tumor cell line (MCF7) to assess the role of c-Myb in luminal breast tumorigenesis. We found that when c-Myb levels were reduced the tumor grew faster, was more anchorage independent, and was able to form tumors in nude mice. In general, when no exogenous estrogen (i.e. estrogen pellet) is added to nude mice, MCF7 cells grow very slowly or do not form tumors at all. However, in the MCF7 cells were c-Myb levels were reduced, no estrogen was required for tumor growth; these results suggest that with lower levels of c-Myb, the tumor is less dependent on estrogen for growth. In agreement with our patient survival data, these results demonstrated that having *c-Myb* expression in the tumor bestows a protective effect, precisely what one would observe with a tumor suppressor gene. However, this begged the question how is this transcription factor behaving as a tumor suppressor in breast cancer?

In collaboration with Dr. Yanping Zhang's lab at UNC, we were able to experimentally determine that c-Myb was working in concert with a mitochondrial protein, Hep27, to regulate p53 stabilization in the luminal tumors. Deisenroth *et al*. (submitted) described an interaction between Hep27 and Mdm2, a p53 inhibitor, whereby Hep27 inhibits Mdm2 function thereby stabilizing p53. Therefore, as Hep27 protein levels increase, there is a corresponding increase in stable p53. We, and others, have described Hep27 as a transcriptional target of c-Myb (Lei *et al.*, 2004). We showed that gradually increasing exogenous amounts of *c-Myb* to MCF7 cells resulted in a corresponding increase in Hep27, stable p53, and the p53 target gene, p21. This suggests that a c-Myb-Hep27-p53 pathway exists in this luminal tumor cell line. Furthermore, using a large human breast tumor microarray dataset we found evidence of the existence of this *in vivo*.

It has been established that luminal/ER-positive tumors respond poorly to chemotherapy, likely due to the fact that the overwhelming majority of these tumors are p53 wild-type (Troester *et al.*, 2006). Furthermore, luminal tumors treated with chemotherapy also show increased p21 protein level post-treatment (Troester *et al.*, 2004). Our results demonstrate that the stability of p53 in this tumor type is actually dependent on ER and c-Myb. When c-Myb levels are increased, p53 is stabilized, and we see a corresponding upsurge in p21 levels.

To further identify phenotypic differences between the c-Myb knock-down line and its parental control, we hybridized labeled mRNA to Agilent microarrays and performed significance analysis of microarrays (SAM). Many gene expression differences were observed with over 1,200 transcripts ("c-Myb signature") differentially expressed with a low false discovery rate. These genes were capable of stratifying breast tumors into groups recapitulating their intrinsic subtypes, indicative of their unique expression across the tumor subtypes.

A recent publication described the use of cell surface markers, EpCAM and CD49f, to isolate four discrete cell populations from the normal, human mammary gland: mammary stem cell enriched (MaSC), luminal progenitor, mature luminal, and stromal (Lim *et al.*, 2009). This group used *in vitro* assays to show they had isolated the described cell types, and microarray gene expression analysis was also carried out on each cell type to identify unique gene expression profiles. We used the gene expression data of Lim *et al.*, 2009 to further analyze our c-Myb signature, as well as previous published gene signatures for ER and GATA3 (a transcription factor that is essential for mammary gland formation) (Oh *et al.*, 2006; Usary *et al.*, 2004).

Our c-Myb gene list strongly correlated with the mature luminal gene profile, and was least correlated with the MaSC signature. *c-Myb* expression levels alone displayed this same correlation. This is in stark contrast to the observation in hematopoiesis, colonogenesis, and neurogenesis where *c-Myb* expression is highest in the stem cell populations and decreases and cells differentiate (Malaterre *et al.*, 2007; Malaterre *et al.*, 2008; Ramsay and Gonda, 2008). We observed the exact opposite expression pattern, lowest *c-Myb* expression levels in the MaSC population, increased levels in the luminal progenitor cells, and highest levels in the differentiated luminal cells. Again, this highlights how differently *c-Myb* is behaving in the mammary gland versus other tissue types in which it has been studied. Why this is the case is currently unknown, but we speculate that in the breast, *c-Myb* is under the strict control of the estrogen receptor, a hormone-responsive transcription factor that is not expressed in the hematopoietic system or the colon. Also, differing transcriptional co-factors that are mammary tissue-specific may be interacting with c-Myb to promote transcriptional activation of different target genes in the mammary gland.

The estrogen-regulated genes, and ESR1 expression alone, were also highest in the mature luminal population followed by having second highest expression in the luminal progenitor cells, and lowest expression levels in the MaSC population. GATA3, a known target of ER and a gene essential in mammary development (Kouros-Mehr *et al.*, 2006), was observed as having its highest expression in the luminal progenitor population; however, the GATA3 gene signature activity was observed highest in the mature luminal population. These results confirm previous observations that ER is not expressed in the MaSC population (Sleeman *et al.*, 2007), and its expression begins in the luminal progenitor cells. The ER targets GATA3 at this stage to promote proper mammary gland differentiation. In the more

differentiated populations, the ER targets the transcription factor *c-Myb*, whose protein targets the activation of genes important for maintenance within the mature luminal cells, such as p53 via Hep27.

The gene signature data presented above may have implications in hypothesizing the cell-of-origin for the luminal tumor subtypes. Recent evidence proposes that the ER-positive, luminal tumors arise from the mature luminal cell population in the mammary gland, whereas ER-negative tumors are derived from an earlier mammary cell lineage (Prat and Perou, 2009). This is consistent with our findings that the ER-regulated activity and ESR1 expression was lowest in the MaSC population, but highest in the differentiated luminal cells. Also, luminal tumors rarely mutated for p53, and in fact, the p53 pathway is most likely intact in this breast cancer subtype (Troester *et al.*, 2006). Here, we described that p53 stabilization in luminal tumors was likely due to an intact ER-c-Myb-p53 pathway, and our data showed that *c-Myb* expression and its downstream transcriptional-regulated activity was highest in the mature luminal population. Further adding to this was our evidence that *c-Myb*'s expression, as well as its regulated activity, is highest in the luminal tumors and lowest in the ER-negative tumor subtypes.

The c-Myb data presented here have very important clinical implications, especially in light of an ongoing clinical trial for patients with advanced hematologic malignancies (National Clinical Trials Identifier: NCT00780052). The drug in this trial is a c-Myb antisense oligonuclotide given to patients with the intent of reducing the amount of c-Myb protein in the tumor. It has been suggested that this form of treatment could be of value in patients with other cancers expressing high *c-Myb*, including breast tumors (Ramsay and Gonda, 2008). However, c-Myb appears to be playing a tumor suppressor role in luminal

breast cancer, and treating patients with an antisense oligonucleotide against this gene could prove detrimental by leading to increased tumorigenesis.

In summary, the MYB family of transcription factors including *B-Myb* and *c-Myb*, both identified as members of the intrinsic gene signature, are undoubtedly playing important roles in the etiology of breast cancer. Our findings point to both genes as important markers of both breast cancer outcome and risk, particularly the *B-Myb* germline variant shown to cause increased risk of basal-like disease. *B-Myb* is therefore a biomarker that is of potential clinical importance for determining disease risk and for guiding treatment. The discovery of a tumor suppressor role for the *c-Myb* oncogene will greatly impact the future of luminal tumor treatment, notably the decision to expand a current clinical trial treating leukemia patients with antisense olionucleotides against c-Myb to breast cancer patients. The work presented here begins to describe important genes from the intrinsic gene set, their involvement in breast tumorigenesis, and lays the groundwork for many future analyses.

REFERENCES

Bergamaschi A, Kim YH, Wang P, Sørlie T, Hernandez-Boussard T, Lonning PE *et al* (2006). Distinct patterns of DNA copy number alteration are associated with different clinicopathological features and gene-expression subtypes of breast cancer. *Genes, Chromosomes and Cancer* **45:** 1033-1040.

Biroccio A, Benassi B, D'Agnano I, D'Angelo C, Buglioni S, Mottolese M *et al* (2001). c-Myb and Bcl-x Overexpression Predicts Poor Prognosis in Colorectal Cancer : Clinical and Experimental Findings. *Am J Pathol* **158:** 1289-1299.

Carey LA, Dees EC, Sawyer L, Gatti L, Moore DT, Collichio F *et al* (2007). The Triple Negative Paradox: Primary Tumor Chemosensitivity of Breast Cancer Subtypes. *Clin Cancer Res* **13:** 2329-2334.

Carey LA, Perou CM, Livasy CA, Dressler LG, Cowan D, Conway K *et al* (2006). Race, Breast Cancer Subtypes, and Survival in the Carolina Breast Cancer Study. *JAMA* **295:** 2492-2502.

Derenzini M, Donati G, Mazzini G, Montanaro L, Vici M, Ceccarelli C *et al* (2008). Loss of Retinoblastoma Tumor Suppressor Protein Makes Human Breast Cancer Cells More Sensitive to Antimetabolite Exposure. *Clin Cancer Res* **14:** 2199-2209.

Drabsch Y, Hugo H, Zhang R, Dowhan DH, Miao YR, Gewirtz AM *et al* (2007). Mechanism of and requirement for estrogen-regulated MYB expression in estrogen-receptor-positive breast cancer cells. *Proceedings of the National Academy of Sciences* **104:** 13762-13767.

Fan C, Oh DS, Wessels L, Weigelt B, Nuyten DSA, Nobel AB *et al* (2006). Concordance among Gene-Expression-Based Predictors for Breast Cancer. *N Engl J Med* **355:** 560-569.

Guérin M SZ, Andrieu N, Riou G. (1990). Strong association between c-myb and oestrogen-receptor expression in human breast cancer. *Oncogene* **5:** 131-5.

Herschkowitz J, Simin K, Weigman V, Mikaelian I, Usary J, Hu Z *et al* (2007). Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. *Genome Biology* **8:** R76.

Herschkowitz JI, He X, Fan C, Perou CM (2008). The functional loss of the retinoblastoma tumor suppressor is a common event in Basal-like and Luminal B breast carcinomas. *Breast Cancer Reserach*.

Hu Z, Fan C, Oh D, Marron J, He X, Qaqish B *et al* (2006). The molecular portraits of breast tumors are conserved across microarray platforms. *BMC Genomics* **7:** 96.

Kouros-Mehr H, Slorach EM, Sternlicht MD, Werb Z (2006). GATA-3 Maintains the Differentiation of the Luminal Cell Fate in the Mammary Gland. *Cell* **127:** 1041.

Lei W, Rushton JJ, Davis LM, Liu F, Ness SA (2004). Positive and Negative Determinants of Target Gene Specificity in Myb Transcription Factors. *J. Biol. Chem.* **279:** 29519-29527.

Lim E, Vaillant F, Wu D, Forrest NC, Pal B, Hart AH *et al* (2009). Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. *Nat Med* **15:** 907-913.

Malaterre J, Carpinelli M, Ernst M, Alexander W, Cooke M, Sutton S *et al* (2007). c-Myb is required for progenitor cell homeostasis in colonic crypts. *PNAS* **104:** 3829-3834.

Malaterre J, Mantamadiotis T, Dworkin S, Lightowler S, Yang Q, Ransome MI *et al* (2008). c-Myb Is Required for Neural Progenitor Cell Proliferation and Maintenance of the Neural Stem Cell Niche in Adult Brain. *Stem Cells* **26:** 173-181.

Millikan RC, Newman B, Tse C-K, Patricia G. Moorman, Conway K, Smith LV *et al* (2008). Epidemiology of basal-like breast cancer. *Breast Cancer Research and Treatment* **109:** 123-139.

Oh D, Troester M, Usary J, Hu Z, He X, Fan C *et al* (2006). Estrogen-regulated genes predict survival in hormone receptor-positive breast cancers. *J Clin Oncol* **24:** 1656 - 1664.

Parker JS, Mullins M, Cheang MCU, Leung S, Voduc D, Vickery T *et al* (2009). Supervised Risk Predictor of Breast Cancer Based on Intrinsic Subtypes. *J Clin Oncol* **27:** 1160-1167.

Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA *et al* (2000). Molecular portraits of human breast tumours. *Nature* **406:** 747.

Prat A, Perou CM (2009). Mammary development meets cancer genomics. *Nat Med* **15:** 842-844.

Ramsay RG, Gonda TJ (2008). MYB function in normal and cancer cells. *Nat Rev Cancer* **8:** 523.

Schwab R, Bussolari R, Corvetta D, Chayka O, Santilli G, Kwok JMM *et al* (2007). Isolation and functional assessment of common, polymorphic variants of the B-MYB proto-oncogene associated with a reduced cancer risk. *Oncogene*.

Sleeman KE, Kendrick H, Robertson D, Isacke CM, Ashworth A, Smalley MJ (2007). Dissociation of estrogen receptor expression and in vivo stem cell activity in the mammary gland
10.1083/jcb.200604065. *J. Cell Biol.* **176:** 19-26.

Sorlie T, Perou C, Tibshirani R, Aas T, Geisler S, Johnsen H *et al* (2001). Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci USA* **98:** 10869 - 10874.

Sorlie T, Tibshirani R, Parker J, Hastie T, Marron JS, Nobel A *et al* (2003). Repeated observation of breast tumor subtypes in independent gene expression data sets. *PNAS* **100:** 8418-8423.

Troester M, Herschkowitz J, Oh D, He X, Hoadley K, Barbier C *et al* (2006). Gene expression patterns associated with p53 status in breast cancer. *BMC Cancer* **6:** 276.

Troester MA, Hoadley KA, Sorlie T, Herbert B-S, Borresen-Dale A-L, Lonning PE *et al* (2004). Cell-Type-Specific Responses to Chemotherapeutics in Breast Cancer. *Cancer Res* **64:** 4218-4226.

Usary J, Llaca V, Karaca G, Presswala S, Karaca M, He X *et al* (2004). Mutation of GATA3 in human breast tumors. *Oncogene* **23:** 7669-7678.

Whitfield ML, Sherlock G, Saldanha AJ, Murray JI, Ball CA, Alexander KE *et al* (2002). Identification of Genes Periodically Expressed in the Human Cell Cycle and Their Expression in Tumors. *Mol. Biol. Cell* **13:** 1977-2000.

APPENDIX II

Genes expressed significantly higher, as determined by Significance Analysis of Microarrays (SAM), in doxorubicin-treated *B-Myb* overexpressing HME-CC cells versus doxorubicin-treated vector control cells. A list of 217 genes with an FDR<3% was obtained and ranked based on fold-change.

Genbank Accession #	Fold Change	Gene Name
NM_002466	9.50	V-myb myeloblastosis viral oncogene homolog (avian)-like 2
NM_017786	4.10	Hypothetical protein FLJ20366
NM_018136	3.28	Asp (abnormal spindle)-like, microcephaly associated (Drosophila)
NM_016343	3.21	Centromere protein F, 350/400ka (mitosin)
NM_006528	3.08	Tissue factor pathway inhibitor 2
NM_018131	3.07	Chromosome 10 open reading frame 3
NM_001067	3.01	Topoisomerase (DNA) II alpha 170kDa
NM_005733	2.99	Kinesin family member 20A
AF108138	2.96	Chromosome 15 open reading frame 20
NM_001786	2.88	Cell division cycle 2, G1 to S and G2 to M
NM_014750	2.88	Discs, large homolog 7 (Drosophila)
NM_006101	2.72	Kinetochore associated 2
NM_001813	2.71	Centromere protein E, 312kDa
NM_181803	2.71	Ubiquitin-conjugating enzyme E2C
NM_004701	2.64	Cyclin B2
NM_007174	2.63	Citron (rho-interacting, serine/threonine kinase 21)
AK098670	2.62	Cell division cycle associated 2
NM_022346	2.62	Chromosome condensation protein G
NM_003258	2.58	Thymidine kinase 1, soluble
ENST00000330382	2.56	

Genbank Accession #	Fold Change	Gene Name
NM_018101	2.56	Cell division cycle associated 8
NM_019013	2.55	Hypothetical protein FLJ10156
NM_030919	2.52	Chromosome 20 open reading frame 129
NM_001389	2.51	Down syndrome cell adhesion molecule
NM_007280	2.51	Opa-interacting protein 5
NM_006845	2.49	Kinesin family member 2C
NM_006461	2.47	Sperm associated antigen 5
NM_003318	2.47	TTK protein kinase
NM_080668	2.47	Cell division cycle associated 5
NM_005192	2.46	Cyclin-dependent kinase inhibitor 3 (CDK2-associated dual specificity phosphatase)
NM_031966	2.45	Cyclin B1
NM_012484	2.45	Hyaluronan-mediated motility receptor (RHAMM)
K03200	2.45	
NM_004336	2.43	BUB1 budding uninhibited by benzimidazoles 1 homolog (yeast)
NM_018685	2.41	Anillin, actin binding protein (scraps homolog, Drosophila)
NM_002358	2.35	MAD2 mitotic arrest deficient-like 1 (yeast)
NM_005480	2.33	Trophinin associated protein (tastin)
NM_024734	2.27	Calmin (calponin-like, transmembrane)
BC000784	2.25	Baculoviral IAP repeat-containing 5 (survivin)
NM_002875	2.23	RAD51 homolog (RecA homolog, E. coli) (S. cerevisiae)
BC032409	2.22	CDNA clone IMAGE:5217021, with apparent retained intron
NM_005563	2.21	Stathmin 1/oncoprotein 18
NM_001827	2.19	CDC28 protein kinase regulatory subunit 2
AL832427	2.16	Fanconi anemia, complementation group D2

Genbank Accession #	Fold Change	Gene Name
BC030154	2.15	Ribonucleotide reductase M2 polypeptide
AK024476	2.15	CTF18, chromosome transmission fidelity factor 18 homolog (S. cerevisiae)
NM_177978	2.11	Chordin
NM_025245	2.10	Pre-B-cell leukemia transcription factor 4
NM_018451	2.09	Centromere protein J
NM_006596	2.07	Polymerase (DNA directed), theta
NM_032117	2.06	GAJ protein
AB058697	2.06	Hypothetical protein FLJ10719
NM_015341	2.04	Barren homolog (Drosophila)
NM_138555	2.04	Kinesin family member 23
NM_022092	2.03	CTF18, chromosome transmission fidelity factor 18 homolog (S. cerevisiae)
NM_031217	2.02	Kinesin family member 18A
BC046178	2.02	CDNA clone IMAGE:4452583, partial cds
THC1572841	2.01	
NM_012112	2.00	TPX2, microtubule-associated protein homolog (Xenopus laevis)
NM_005879	1.98	TRAF interacting protein
NM_032997	1.97	ZW10 interactor
NM_033316	1.97	Antigen p97 (melanoma associated) identified by monoclonal antibodies 133.2 and 96.5
NM_017915	1.97	Hypothetical protein FLJ20641
NM_014791	1.97	Maternal embryonic leucine zipper kinase
AK094154	1.96	CDNA FLJ36835 fis, clone ASTRO2010996
AK097948	1.94	Hypothetical protein FLJ40629
NM_020242	1.94	Kinesin-like 7
NM_002129	1.93	High-mobility group box 2

Genbank Accession #	Fold Change	Gene Name
BC044933	1.92	Hypothetical protein LOC146909
NM_015426	1.92	DKFZP434C245 protein
NM_173466	1.91	Hypothetical protein DKFZp434P055
AL137347	1.90	Hypothetical protein DKFZP761M1511
NM_003600	1.90	Serine/threonine kinase 6
NM_013277	1.90	Rac GTPase activating protein 1
NM_014321	1.89	Origin recognition complex, subunit 6 homolog-like (yeast)
NM_006479	1.88	RAD51 associated protein 1
ENST00000317847	1.87	
NM_005030	1.87	Polo-like kinase 1 (Drosophila)
NM_004636	1.86	Sema domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3B
NM_001826	1.86	CDC28 protein kinase regulatory subunit 1B
BC041395	1.85	Homo sapiens, Similar to diaphanous homolog 3 (Drosophila), clone IMAGE:5277415, mRNA
NM_052969	1.85	Ribosomal protein L39-like
NM_003981	1.84	Protein regulator of cytokinesis 1
NM_001992	1.84	Coagulation factor II (thrombin) receptor
BC046178	1.84	CDNA clone IMAGE:4452583, partial cds
I_3537291	1.84	
AK001581	1.83	Hypothetical protein FLJ10719
NM_013239	1.82	Protein phosphatase 2A 48 kDa regulatory subunit
NM_013277	1.82	Rac GTPase activating protein 1
NM_002266	1.81	Karyopherin alpha 2 (RAG cohort 1, importin alpha 1)
NM_004848	1.81	Chromosome 1 open reading frame 38

Genbank Accession #	Fold Change	Gene Name
NM_016448	1.80	RA-regulated nuclear matrix-associated protein
NM_000057	1.80	Bloom syndrome
NM_005496	1.80	SMC4 structural maintenance of chromosomes 4-like 1 (yeast)
NM_182687	1.80	Membrane-associated tyrosine- and threonine-specific cdc2-inhibitory kinase
NM_003878	1.80	Gamma-glutamyl hydrolase (conjugase, folylpolygammaglutamyl hydrolase)
NM_018492	1.79	T-LAK cell-originated protein kinase
AI791206	1.79	MRNA (fetal brain cDNA g6_1g)
NM_152259	1.79	Leucine-rich repeat kinase 1
BC035696	1.79	Similar to RIKEN cDNA 2700049P18 gene
NM_145061	1.78	Chromosome 13 open reading frame 3
NM_021077	1.78	Neuromedin B
AB006624	1.78	KIAA0286 protein
X85137	1.78	Kinesin family member 11
NM_000227	1.78	Laminin, alpha 3
ENST00000320402	1.75	
NM_018454	1.75	Nucleolar and spindle associated protein 1
NM_145018	1.75	Hypothetical protein FLJ25416
NM_024629	1.74	MLF1 interacting protein
XM_293746	1.73	
NM_006187	1.73	2'-5'-oligoadenylate synthetase 3, 100kDa
THC1410715	1.73	
BC005970	1.73	Vaccinia related kinase 1
NM_003579	1.73	RAD54-like (S. cerevisiae)
NM_007294	1.72	Breast cancer 1, early onset

Genbank Accession #	Fold Change	Gene Name
NM_021874	1.71	Cell division cycle 25B
ENST00000325544	1.71	
NM_018098	1.70	Epithelial cell transforming sequence 2 oncogene
NM_015161	1.70	ADP-ribosylation factor-like 6 interacting protein
NM_152521	1.70	
NM_004260	1.69	RecQ protein-like 4
NM_018098	1.69	Epithelial cell transforming sequence 2 oncogene
NM_014109	1.69	ATPase family, AAA domain containing 2
NM_000356	1.68	Treacher Collins-Franceschetti syndrome 1
NM_012291	1.68	Extra spindle poles like 1 (S. cerevisiae)
NM_003920	1.67	Timeless homolog (Drosophila)
NM_024918	1.67	Chromosome 20 open reading frame 172
NM_018186	1.67	Hypothetical protein FLJ10706
NM_003686	1.66	Exonuclease 1
AW082201	1.66	
NM_024094	1.65	Defective in sister chromatid cohesion homolog 1 (S. cerevisiae)
NM_002106	1.65	H2A histone family, member Z
NM_138419	1.65	Family with sequence similarity 54, member A
AK056691	1.65	Haspin
NM_004843	1.64	Interleukin 27 receptor, alpha
NM_016195	1.64	M-phase phosphoprotein 1
NM_002452	1.63	Nudix (nucleoside diphosphate linked moiety X)-type motif 1
NM_032883	1.63	Chromosome 20 open reading frame 100
NM_018369	1.62	DEP domain containing 1B

Genbank Accession #	Fold Change	Gene Name
AB046816	1.61	KIAA1596
NM_031942	1.61	Cell division cycle associated 7
NM_002106	1.61	H2A histone family, member Z
NM_001255	1.61	CDC20 cell division cycle 20 homolog (S. cerevisiae)
NM_025108	1.60	Hypothetical protein FLJ13909
AA470111	1.59	Similar to protein phosphatase 2A 48 kDa regulatory subunit isoform 1; serine/threonine protein phosphatase 2A, 48kDa regulatory subunit; PP2A, subunit B, PR48 isoform; PP2A B subunit PR48; NY-REN-8 a
NM_030928	1.59	DNA replication factor
NM_002916	1.59	Replication factor C (activator 1) 4, 37kDa
AK027267	1.59	RAS-like, family 11, member B
NM_001254	1.59	CDC6 cell division cycle 6 homolog (S. cerevisiae)
I_3512484	1.59	
NM_012177	1.58	F-box protein 5
AK075134	1.58	Cell division cycle associated 7
NM_018132	1.58	Chromosome 6 open reading frame 139
ENST00000329801	1.58	
BC010176	1.57	Similar to RIKEN cDNA 2410004L22 gene (M. musculus)
NM_004523	1.57	Kinesin family member 11
XM_302460	1.56	
NM_024808	1.56	FLJ22624 protein
ENST00000332692	1.56	
U79240	1.56	PAS domain containing serine/threonine kinase
NM_017760	1.56	More than blood homolog
NM_002388	1.55	MCM3 minichromosome maintenance deficient 3 (S. cerevisiae)

Genbank Accession #	Fold Change	Gene Name
NM_005483	1.55	Chromatin assembly factor 1, subunit A (p150)
NM_004111	1.55	Flap structure-specific endonuclease 1
NM_014176	1.54	HSPC150 protein similar to ubiquitin-conjugating enzyme
ENST00000325202	1.54	
NM_014708	1.54	Kinetochore associated 1
NM_014865	1.54	Chromosome condensation-related SMC-associated protein 1
NM_003914	1.54	Cyclin A1
AB028070	1.53	Activator of S phase kinase
I_2001857	1.52	
AL117629	1.52	DKFZP434C245 protein
NM_005441	1.52	Chromatin assembly factor 1, subunit B (p60)
NM_138639	1.52	BCL2-like 12 (proline rich)
BC045739	1.51	Chromosome 15 open reading frame 23
NM_024037	1.51	Hypothetical protein MGC2603
THC1562595	1.51	
NM_016095	1.51	DNA replication complex GINS protein PSF2
NM_007086	1.50	WD repeat and HMG-box DNA binding protein 1
NM_014220	1.50	Transmembrane 4 superfamily member 1
NM_000121	1.49	Erythropoietin receptor
NM_004526	1.49	MCM2 minichromosome maintenance deficient 2, mitotin (S. cerevisiae)
I_2011745	1.49	
NM_173083	1.48	Lin-9 homolog (C. elegans)
NM_004629	1.47	Fanconi anemia, complementation group G
NM_173553	1.47	Hypothetical protein FLJ25801

Genbank Accession #	Fold Change	Gene Name
NM_014256	1.47	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 3
NM_018087	1.46	Hypothetical protein FLJ10407
NM_003503	1.46	CDC7 cell division cycle 7 (S. cerevisiae)
NM_006437	1.45	Poly (ADP-ribose) polymerase family, member 4
NM_014505	1.45	Potassium large conductance calcium-activated channel, subfamily M, beta member 4
BC021694	1.45	Adducin 3 (gamma)
AK024475	1.45	DKFZP434I216 protein
NM_001197	1.44	BCL2-interacting killer (apoptosis-inducing)
NM_001237	1.44	Cyclin A2
NM_145249	1.43	Family with sequence similarity 14, member B
NM_016824	1.43	Adducin 3 (gamma)
NM_005914	1.43	MCM4 minichromosome maintenance deficient 4 (S. cerevisiae)
NM_005342	1.43	High-mobility group box 3
BC018086	1.42	Similar to RIKEN cDNA A430101B06 gene
NM_175923	1.42	Hypothetical protein MGC42630
AK000898	1.41	Zwilch
NM_002915	1.41	Replication factor C (activator 1) 3, 38kDa
AK055071	1.41	Phosphatidylinositol glycan, class K
NM_018455	1.40	Uncharacterized bone marrow protein BM039
NM_018140	1.40	Centrosomal protein 72 kDa
NM_000169	1.40	Galactosidase, alpha
THC1519373	1.39	
AF117229	1.39	Protein x 0004
THC1551878	1.38	

Genbank Accession #	Fold Change	Gene Name
NM_014467	1.38	Sushi-repeat-containing protein, X-linked 2
BC039021	1.37	CDNA clone IMAGE:6043059, partial cds
NM_007047	1.37	Butyrophilin, subfamily 3, member A2
I_932423	1.35	
AK074989	1.35	Hypothetical protein BC007899
AF258562	1.34	Deoxythymidylate kinase (thymidylate kinase)
THC1579998	1.34	