AMERICAN UNIVERSITY OF BEIRUT

RBM20 AND ARHGAP25 EXPRESSION ALTERATION

AND ROLES IN BREAST CANCER *IN VITRO* MODELS

by
RAHME MAHMOUD HILAL

Approved by:

_____
Dr. Diana Jaalouk, Associate Professor   Advisor
Department of Biology

_____
Dr. Rabih Talhouk, Professor   Member of Committee
Department of Biology

_____
Dr. Rihab Nasr, Associate Professor   Member of Committee
Department of Anatomy, Cell Biology and
Physiological Sciences

Date of thesis/dissertation defense: December 17, 2020

AMERICAN UNIVERSITY OF BEIRUT

RBM20 AND ARHGAP25 EXPRESSION ALTERATION AND ROLES IN BREAST CANCER IN VITRO MODELS

by
RAHME MAHMOUD HILAL

A thesis
submitted in partial fulfillment of the requirements
for the degree of Master of Science
to the Department of Biology
of the Faculty of Arts and Sciences
at the American University of Beirut

Beirut, Lebanon
December 2020

AMERICAN UNIVERSITY OF BEIRUT

RBM20 AND ARHGAP25 EXPRESSION ALTERATION AND ROLES IN BREAST CANCER IN VITRO MODELS

by
RAHME MAHMOUD HILAL

Approved by:

_____

Dr. Diana Jaalouk, Associate Professor   Advisor
Department of Biology

_____

Dr. Rabih Talhouk, Professor   Member of Committee
Department of Biology

_____

Dr. Rihab Nasr, Associate Professor   Member of Committee
Department of Anatomy, Cell Biology and
Physiological Sciences

Date of thesis/dissertation defense: December 17, 2020

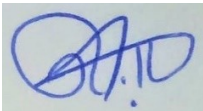# AMERICAN UNIVERSITY OF BEIRUT

PROJECT RELEASE FORM

Student Name: _____Hilal_____Rahme_____Mahmoud_____
           Last   First   Middle

☐
I authorize the American University of Beirut, to: (a) reproduce hard or electronic copies of my project; (b) include such copies in the archives and digital repositories of the University; and (c) make freely available such copies to third parties for research or educational purposes:

☐ As of the date of submission

☐ One year from the date of submission of my project.

☐ Two years from the date of submission of my project.

☒ Three years from the date of submission of my project.

_____January 18, 2021_____

Signature   Date

ACKNOWLEDGMENTS

ABSTRACT
OF THE THESIS OF

Rahme Mahmoud Hilal   for   <u>Master of Science</u>

                                                <u>Major</u>: Cell and Molecular Biology

Title: <u>RBM20 and ARHGAP25 Expression Alteration and Roles in Breast Cancer *in vitro* Models</u>

Breast cancer is the second leading type of cancer occurrence and mortality in women worldwide and studies are constantly attempting to find new treatment strategies, as well as better understand the mechanisms and pathways that control its initiation and progression. Alternative splicing is one of the major cellular processes that are deregulated in cancer, with more and more splicing factors being implicated in various steps of the tumorigenic process. RBM20 is a nuclear splicing factor mainly restricted in expression to the heart and skeletal muscle cells. Previous studies have implicated RBM20 in several heart defects and diseases that result from its altered splicing behavior due to mutations that are majorly loss of function. Several members of the RBM family have been implicated in various types of cancer, like RBM5, RBM6, RBM10 and others. However, RBM20 has not been shown to play any role in cancer. ARHGAP25 is a cytoplasmic RacGAP with known functions in cells of the hematopoietic lineage. Since it is part of the Rho/rac pathway, it has a major function in cell motility and migration. It also plays roles in immune system signaling pathways, as well as free radical production and phagocytosis in neutrophils. Importantly, it has been recently shown to have significant functions in each of rhabdomyosarcoma, colorectal cancer and lung cancer. In this study, we used western blotting, immunofluorescence staining, qRT-PCR, Co-IP and drug inhibition of AMPK to investigate the expression levels, subcellular localization and interactions between these two proteins, as well as potential pathways controlling the expression and activity of RBM20 in breast cancer. We showed that RBM20 expression, and more evidently, RBM20 isoform expression, is altered in two breast cancer *in vitro* models, MDA-MB-231 and MCF-7 cell lines, that are considered more and less aggressive models, respectively. Not only did we observe RBM20 isoform switching, but a novel and much smaller isoform was consistently seen exclusively in the MDA-MB-231 cells line which also had significantly higher levels of ARHGAP25 protein compared to MCF-7 cells. While ARHGAP25 subcellular localization was not altered in either cell line, that of RBM20 showed a significant shift towards the cytoplasm in the MCF-7 cell line. Phage display biopanning

assays have shown that the C-terminal of RBM20 can potentially directly interact with a 7-amino acid peptide sequence that matches a sequence in the ARHGAP25 protein. Upon that, we aimed to validate this interaction in the breast cancer cell lines. Since the PI3K/mTOR pathway is implicated in RBM20 expression, we inhibited AMPK using the drug Dorsomorphin Dihydrochloride and showed that AMPK does not directly affect RBM20 protein expression levels but may potentially play a role in its phosphorylation levels. The results indicate that RBM20 and ARHGAP25 are both implicated in breast cancer and can potentially interact in a direct protein-protein manner. The functions each is involved in raises the possibilities that they affect and regulate several cancer hallmarks to be investigated.

TABLE OF CONTENTS

4

ILLUSTRATIONS

TABLES

ABBREVIATIONS

%	Percent
˚C	Degrees Celcius
AML	Acute Myeloid Leukemia
AMPK	AMP-Activated Protein Kinase
APS	Ammonium Persufate
ARHGAP25	RhoGTPase-Activating Protein 25
ATCC	American Tissue Cell Culture
BCSC	Breast Cancer Stem Cell
CACNA1C	Calcium Channel Voltage-Gated L type Alpha 1C subunit
CAMK2δ	Ca2+/Calmodulin Dependent Protein Kinase II Delta
CCL28	C-C motif Ligand 28
cDNA	Circular DNA
circRNA	Circular RNA
CO2	Carbon Dioxide
C-RBM20	C-terminal RBM20
CRSL	Core Research Science Laboratory
CS-RBD	Consensus Sequence RNA-Binding Domain
DAPI	4',6-Diamidino-2-Phenylindole
DMSO	Dimethyl Sulfoxide
ECL	Electrogenerated Chemiluminescence
EDTA	Ethylene Diamine Tetraacetic Acid
EMT	Epithelial/Mesenchymal Transition
ER	Estrogen Receptor
ERK	Extracellular Signal-Regulated Protein Kinase
FBS	Fetal Bovine Serum
FHOD3	Formin Homology 2 Domain-Containing Protein 3
GAP	GTPase-Activating Protein
GDI	Guanine Nucleotide Dissociation Inhibitors
GIF	Guanine Nucleotide Exchange Factor
H2O	Water
HER2	Human Epidermal Growth Factor Receptor 2
hnRNPK	Heterogenous Nuclear Ribonucleoprotein K
HOXA4	Homeobox A4
HRP	HorseRadish Peroxidase
IgA	Immunoglobulin A
MaSC	Mammary Stem Cell

MDS	Myelodysplasia
MMP	Matrix Metalloprteinases
mRNA	Messenger RNA
mTOR	Mammalian Target of Rapamycin
NaOH	Sodium Hydroxide
PBS	Phosphate Buffered Saline
PDLIM5	PDZ and LIM Domain Protein 5
PFA	Paraformaldehyde
PH domain	Pleckstrin Homology Domain
PI3K	Phosphoinositol-3-Kinase
PKB	Protein Kinase B
PKM	Pyruvate Kinase Muscle
PR	Progesterone Receptor
PVDF	Polyvinylidene Fluoride
QKI	Quaking KH domain containing RNA binding protein
qRT-PCR	Quantitative Real-Time Polymerase Chain Reaction
RBM	RNA Binding Motif
RBP	RNA Binding Protein
RPMI	Roswell Park Memorial Institute Medium
RRM	RNA Recognition Motif
RYR2	Ryanodine Receptor 2
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
SEM	Standard Error of the Mean
SF3B	Splicing Factor 3B
snRNPs	Small Nuclear Ribonucleic Particles
SRSF	Serine/Arginine-Rich Splicing Factor 2
TDU	Terminal Ductular Lobuloalveolar Units
TEMED	Tetramethylethylenediamine
TNBC	Triple Negative Breast Cancer
TRA2β	Transformer 2β Homolog
ttn	Titin Gene
U2AF1	U2 Small Nuclear RNA Auxillary Factor 1
ZASP	Z-Band Alternatively Spliced PDZ-Motif Protein
ZnF	Zinc Finger
ZRSR2	Zinc Finger RNA Binding Motif & Serine/Arginine-Rich 2
μ	Micro

CHAPTER I

LITERATURE REVIEW

A. Breast, Mammary Glands & their Development

The mammary gland is a highly adaptive organ that continues its stages of development during puberty, pregnancy, lactation and involution (removal of the milk-producing cells and replacing them with adipocytes when the former become redundant at weaning) (Figure 1), all of which is controlled by endocrine and nutritional factors (Zhou et al. 2019, Capuco et al. 2013). Although the mammary gland does not develop into anything functional in the embryo, the main cell lineages are established by birth (Capuco et al. 2013). The branches of the mammary gland are composed of a bilayered epithelium that forms ducts during puberty and in the virgin adult and alveoli during pregnancy and lactation. The inner layer is comprised of luminal cells while the outer layer of myoepithelial or basal cells (Zhou et al. 2019).
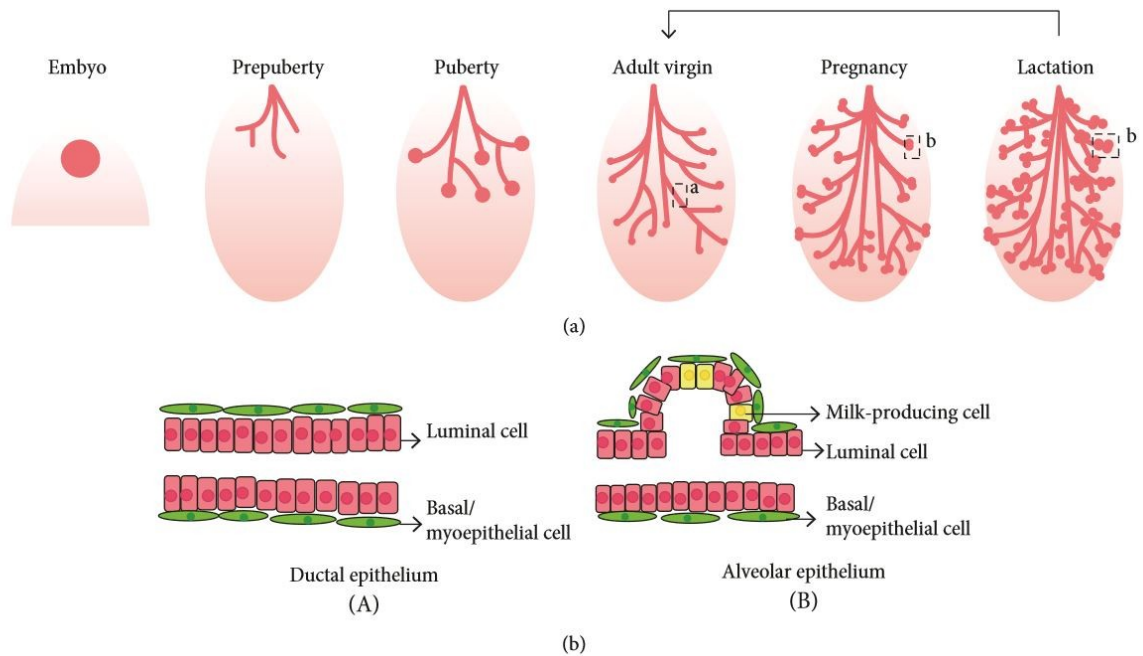
Figure 1 The mammary gland and its postnatal development (*Adapted from Zhou et al., 2019*). (a) Mammary gland development and cycling during the different stages and cycles. Growth of ducts and alveoli. (b) The ductal and alveolar mammary epithelium phenotypes.

Even during the prepubertal stage, mammary cells undergo rapid extension and proliferation into ducts. The distal ends (relative to the teats) form buds known in humans as terminal ductular lobuloalveolar units (TDUs). Hormonal changes during pregnancy, namely the predominance of progesterone (Capuco et al. 2013), induce the alveolar epithelium to quickly proliferate and differentiate until lactation which is when these luminal cells synthesize and secrete milk (Zhou et al. 2019). At this stage, the

myoepothelium contracts to pump and deliver the milk. During weaning, apoptosis, coupled with extracellular matrix remodeling, allows the expanded epithelium to retract (Zhou et al. 2019). While progesterone induces alveolar development, estrogen drives ductal proliferation (Capuco et al. 2013). Other regulators include some immune components, such as the STAT6 and IL4 axis that is active in T lymphocyte development. That is in addition to the interplay between mammary and immune cells, like mast cells and macrophages, during mammary gland development and lactation (Capuco et al. 2013). Prolactin is also one of the main modulators of mammary development. Although it plays a major role in that process, it is also speculated to contribute to breast cancer metastasis to bone and, since it also has an important function in bone development, once it is secreted by breast cancer cells residing in the bone, leads to faster osteoclast development and bone breakdown (Shemanko et al, 2016).

The amount of milk secreted during lactation depends on several factors, including the amount and activity of secretory epithelial cells, as well as the death and production of new such cells. The mammary growth that occurs during early lactation is key in determining the amount of milk produced later (Capuco et al. 2013). Secretory cells undergo some changes at the onset of lactation like the closure of tight junctions between cells, increased metabolism and increased differentiation into the secretory phenotype. Persistence of these cells also depends on several anti-apoptotic and pro-survival factors (Capuco et al. 2013). In addition to providing the neonate with nutrients, breastfeeding also helps establishing the microbial flora in the neonate's gastrointestinal tract and protecting against infections. This is done by attracting plasma cells to the mammary gland with

CCL28 where they will secrete IgA that will get transported across the mammary epithelial cells into the lumen with the milk (Niimi et al, 2018). It is hypothesized that the bacteria present in breast milk migrates from the maternal gut to the mammary gland to be incorporated into the milk. If that is the case, then these 2 processes also provide the mother with protection against infection of the mammary ducts (Niimi et al, 2018).

Involution comes as a result of the apoptotic death and phagocytic removal of secretory cells and largely depends on the pregnancy state of the female at the time of weaning. At this point, alveolar structures are lost; rapidly in non-pregnant females compared to pregnant ones (Capuco et al. 2013). This process occurs after the gradual declination of the number of secretory cells throughout the lactation period.

The mammary gland contains a group of adult stem cells, termed the mammary stem cells or MaSCs. These function in mammary gland homeostasis, repair and renewal (Zhou et al. 2019), making them targets of modulation, aiming at increasing milk production in domesticated mammals (Capuco et al. 2013). Mesenchymal stem cells are also present in the breast and contribute to the stroma (Capuco et al. 2013). The MaSC surface markers that have found consensus across all studies are CD24, β1 integrin and α6 integrin, in addition to others that appeared in some studies but not others (Zhou et al. 2019). The roles of these adult stem cells are especially pivotal since the mammary gland does most of its development postnatally. A striking role of MaSCs and one that supports the notion that they are long-lived within the mammary gland is the one played in alveolar epithelium remodeling during the pregnant cycle (Zhou et al. 2019). Human MaSCs are

localized in the mammary ducts while lobule progenitors that arise from the asymmetric division of the MaSCs reside in the TDUs (Capuco et al. 2013).

Importantly, some markers are shared by MaSCs and breast cancer stem cells (BCSC) such as CD29, CD61 and others. In some cases, BCSCs could originate from MaSCs since the fast and often self-renewal of the latter increases their susceptibility to oncogenic hits and mutations (Zhou et al. 2019).

B. Breast Cancer

Cancer seems to be the disease of the century due to its growing prevalence rate and challenges in treatment. Cancer is generally characterized by specific hallmarks that become evident in transformed cells (Hanahan & Weinberg, 2011). Some of the most important of these include metabolic switching, resistance to apoptosis and enhanced cancer cell proliferation. This is manifested in the inability of cancer cells to undergo their supposedly "programmed cell death" upon excessive and irreparable DNA damage and independence from proliferation induction by their microenvironment (Hanahan & Weinberg, 2011). In addition, cancer is not only known for its altered protein expression levels and pathway activation, but for novel protein-protein interactions within the cells that may give rise to new functions.

Breast cancer, in specific, is the most prevalent form of cancer in women with leading mortality rates, especially in women aged 35 to 54 years. This is still the case even after witnessing a significant decrease in mortality rates from breast cancer in the past 2

decades due to early screening and detection (Bonilla at al, 2017). This makes it a hot topic for research and new discoveries are constantly being made regarding the biology and physiology of breast tumors. After increasing steadily for several years, the rates of breast cancer occurrence have started to decrease or stabilize in many developed countries while it is still unstable in most developing ones (Bonilla at al, 2017). In addition, breast cancer occurrence, although sporadic in most cases, is increased by several risk factors that include the following: prolonged estrogen exposure, environmental pollutants, radiation, smoke and alcohol abuse, obesity, unhealthy diets, physical inactivity and age (Bonilla at al. 2017, WHO).

As the biological profiles of tumors are highly variant, the surface and molecular biomarkers on cancer cells are used to classify them into different subtypes, which differ in terms of grade and prognosis (Masood, 2016). Breast cancers are categorized into five subtypes mainly, but not solely, based on their surface expression of each of the estrogen receptor, progesterone receptor and human epidermal growth factor receptor 2 (HER2) (Bonilla at al, 2017). 50 % of breast tumors are of the Luminal A subtype which is characterized by being ER +ve and/or PR +ve and HER2 –ve. It usually has a low histological grade and the best prognosis among the 5 subtypes (Bonilla at al, 2017). Luminal A metastases usually target the bone. The Luminal B subtype accounts for 15 % of breast cancers and resembles Luminal A in terms of biomarker expression and image characteristics. However, it has a higher grade (of III or IV) and a much worse prognosis, especially as compared to the Luminal A subtype (Bonilla at al, 2017). The distinction between the Luminal A and Luminal B subtypes can be in-part attributed to the low levels

of Ki-67 in the former and its high levels in the latter, given that this protein "helps control

how fast cancer cells grow" (Raghunath at al, 2019). In addition to being ER positive, the

Luminal A and Luminal B subtypes present with a higher-than-normal expression of

GATA3. The therapy favored for these tumors is endocrine therapy (Masood, 2016). The

HER2+ subtype, which represents 15 % to 30 % of breast tumors, is characterized by being

positive for HER2, due to an Erb-B2 amplification or overexpression, and negative for both

ER and PR (Bonilla at al, 2017). It is usually associated with ductal carcinoma *in situ*

(Masood, 2016) and metastasizes to the bone, liver and brain (Bonilla at al, 2017). This

subtype used to have the worst prognosis before the use of trastuzumab or herseptin, the

targeted anti-HER2 therapy, the use of which gave HER2+ breast tumors the best

prognosis. However, this targeted therapy is not very efficient in treating metastatic HER2+

breast cancer (Bonilla at al, 2017). The triple negative breast cancer (TNBC) subtype is, as

the name implies, negative for all 3 receptors and represents only 12 % to 17 % of breast

tumors but accounts for the most deaths since it is highly aggressive. It is of high

histological grade and metastasizes to axillary nodes, in addition to the lungs and brain, the

sites of common relapse (Bonilla at al, 2017). TNBCs also express high levels of Ki-67,

reflecting their high proliferation rate (Masood, 2016). Finally, the basal like subtype is

described by being ER –ve with a high histological grade that sometimes resembles the

triple negative breast cancer subtype in terms of physiology and requires combination

chemotherapy as treatment (Bonilla at al, 2017). Both, the TNBC and basal-like subtypes,

can be seen as negative for all three growth factor hormonal receptors and expressing

higher levels of high molecular weight cytokeratins, but breast tumors that are specifically

and distinctly differentiated from TNBCs and classified as the basal-like subtype are

positive for CK5/6 and EGFR (epidermal growth factor receptor) (Masood, 2016). Distinguishing between these subtypes is crucial since their response to neoadjuvant therapy is variable (Dai et al, 2016). A less commonly used hormonal growth factor receptor for classification is the androgen receptor (Dai et al, 2016). Importantly, the discussed subtypes vary in their expression of EMT and stem cell markers, as well as interferon-regulated genes that allow the tumor to evade the immune response (Dai, 2016).

Technological developments in the last years, spanning over earlier detection methods, the use of genomics (prognosis, chances of relapse, etc.), less invasive surgical treatments and advanced/novel systemic therapies, have revolutionized the approach to breast cancer (Bonilla at al, 2017). Early detection, as well as the discovery of smaller lumps and carcinomas, by mammography has contributed to the decrease in breast cancer mortality rates. By 2015, the International Agency for Research on Caner (IRAC) finally verified that mammographic screening actually does lead to an about 40% reduction in breast cancer mortality rates (Bonilla at al, 2017). Other factors contributing to this reduction are the canonical clinical/self-breast examination and the advancements in systemic therapies and surgery. This is especially the case for systemic therapies, like adjuvant chemotherapy, targeted therapy and hormone therapy, that can target metastatic breast tumors. As for the surgical techniques currently implemented, these include breast conservative surgery followed by radiotherapy, made more possible with the introduction of oncoplastic surgical techniques (Bonilla at al. 2017, Raghunath at al, 2019).

As the main setback in breast cancer treatment is targeting metastases, especially to the brain, new methods are actively being explored. The 2 breast cancer subtypes with the

highest incidences of brain metastases are the TNBC and HER2+ subtypes (Raghunath at al, 2019). The metastasized tumors cause a problem mainly due to their genetic variability from their primary tumors. Medical advances have allowed the moving forward from the previously applied treatments that included surgery, whole brain radiation therapy and stereotactic radiotherapy (Raghunath at al, 2019). The first is now only limited to patients with certain numbers of intracranial lesions, the second is no longer a general recommendation after surgery and even systemic therapy is being reevaluated with the emergence of targeted therapies. The major track now seems to be radiotherapy in combination with targeted and immune therapy (Raghunath at al, 2019).

C. Alternative Splicing & Splicing Factors in Cancer

Splicing is a major post-transcriptional modification that is necessary for organism complexity. It is mainly categorized into either constitutive or alternative splicing (Weeland et al, 2015). Constitutive splicing is the excision of all introns (non-coding sequences) and splicing the exons (coding sequences) back together to form the mature mRNA from the pre-mRNA. Alternative splicing is removing different introns from different pre-mRNA transcripts of a gene, resulting in varying final transcripts with different exons included or excluded. This gives rise to different protein isoforms that may have various structures, intracellular localizations and functions (Weeland et al, 2015). This process is achieved by the spliceosome, a complex of 5 snRNPs (small nuclear ribonucleic particles), that recruits other proteins to the splice site. The splice sites need to be flanked by certain sequences on the pre-mRNA (cis-acting regulatory sequences) recognized by RBPs (RNA-binding

proteins) that allow the spliceosome to cut and stitch at the right locations depending on the context. These RBPs can act as either enhancers or silencers (Weeland et al, 2015).

Until recently, very few alterations in splicing regulators had been reported in cancer. Now, however, it is becoming more and more evident that splicing factor alterations are involved in every step of the tumorigenic process (Weeland et al, 2015). The splice variants they give rise to can be involved in increased tumor cell proliferation, resistance to apoptosis, ability to invade surrounding tissue and metastasize to distant sites, the metabolic switch, induction of angiogenesis, acquiring drug resistance and evading the immune response (Figure 2). Splicing factors in cancer can be altered in different manners (Weeland et al, 2015). In some cases, somatic mutations occur within the coding DNA sequences of the splicing factor genes themselves, giving rise to truncated or nonfunctional splicing factors. In other cases, these coding sequences would remain intact but the expression levels of splicing factors get altered due to some transcription factors that are usually overexpressed in the context of cancer, such as myc, or due to modifications at different levels of signaling pathways as growth factor receptors and their downstream cascades get deregulated in cancer cells (Weeland et al, 2015). This means that the oncoproteins and tumor suppressor proteins that induce the change in cellular characteristics from normal to transformed may be intact themselves but would behave irregularly due to different alternative splicing patterns that would give rise to different isoforms of these proteins with no or added functions.
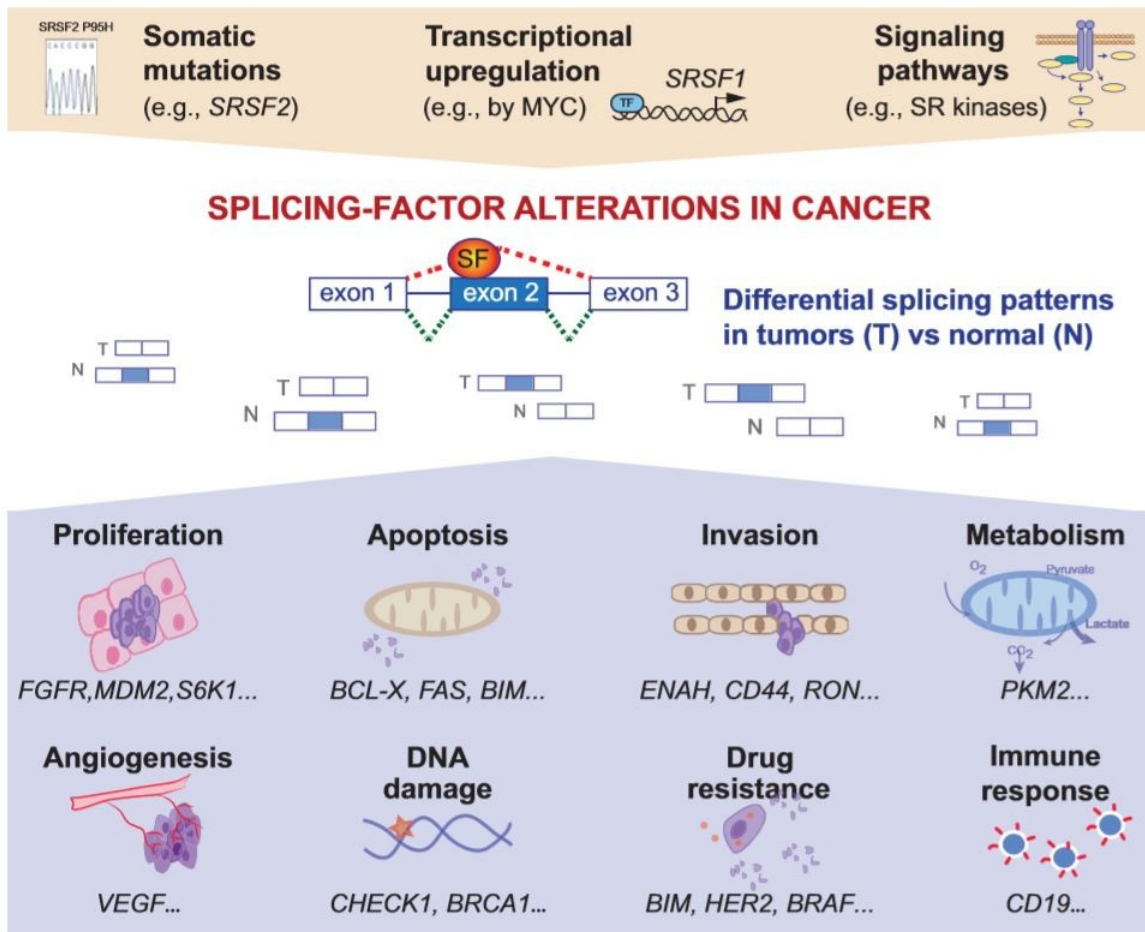
Figure 2 Splicing factor alterations in cancer (adapted from Anczuków et al. 2016). This drawing shows the type of splicing factor alterations in cancer and the scope of their effects.

Just like other tumor descriptions, there is a distinction between hematological malignancies and solid tumors when it comes to the splicing factors that can be altered (Anczuków & Krainer, 2016). Some splicing factors altered in hematological malignancies include SF3B1 (splicing factor3b subunit 1), SRSF2 (serine/arginine-rich splicing factor 2),

U2AF1 (U2 small nuclear RNA auxiliary factor 1) and ZRSR2 (zinc finger RNA binding motif and serine/arginine-rich 2). Those found in solid tumors include SF3B1 (especially common in ER +ve breast tumors), SRSF1, SRSF3, SRSF6, TRA2β (transformer 2β homolog) and hnRNPK (heterogeneous nuclear ribonucleoprotein K) (Anczuków & Krainer, 2016). In addition, some act as oncoproteins while others act as tumor suppressors. The former include SRSF1, SRSF3, SRSF6, HNRNPA2/B1 and HNRNPH while the latter include QKI, RBM5, RBM6 and RBM10 (Anczuków & Krainer, 2016).

Due to the evolving understanding of the role of RNA splicing in cancer, the therapeutic targeting of this process also quickly kicked in. For example, a class of compounds that target SF3b, a component of the U2 snRNP, are widely used in the treatment of cancers, such as myelodysplasia (MDS), and work by blocking branch point exposure to this splicing factor (Kim & Abdel-Wahab, 2017). Another class of compounds, the anticancer sulfonamides, are used to target RBM39 and RBM23, which are serine/arginine-rich RNA-binding proteins, and induce their proteasomal degradation through ubiquitilation (Ting at al, 2019) in MDS (Kim & Abdel-Wahab, 2017) and acute myeloid leukemia (AML) (Thomas & Majeti, 2019). By specifically blocking RBM39 in AML, the cancer cell would be led to synthetic lethality since this RNA-binding protein does not act as an oncoprotein itself and isn't mutated but is rather required by the AML cell to survive (Thomas & Majeti, 2019, Wang et al, 2019). However, the amount of compounds tested and used as targeted therapies for RNA splicing abnormalities in cancer is still limited and mainly restricted to sulfonamides and the SF3b-specific compounds. In

addition, even these compounds still face some problems as the safety of the first group and the efficacy of the second group need to be reevaluated (Kim & Abdel-Wahab, 2017).

1. *RBM Family*

The RNA Binding Motif (RBM) proteins are a family of RBPs with RNA recognition motifs. They are related in terms of structure and, to a lower extent, function (Sutherland et al, 2005). Different members of the RBM family have been shown to play various roles. These include: splice site selection, intron splicing, non-sense mediated RNA decay, gametogenesis, heart development (like RBM24 and RBM20) and apoptosis regulation (Sutherland et al, 2005; Blech-Hermoni & Ladd, 2013; Chai et al, 1997). In addition, some RBM proteins, including RBM24a and RBM24b, have been implicated in the development and differentiation of somites (Maragh et al, 2014). RBMX can also potentially contribute to the degenerative process of the damaged retinae (Dai et al, 2015).

Several RBM proteins have been found to be implicated in the process of apoptosis (Sutherland et al, 2005). The small variant of RBM10 is found to be co-expressed with caspase-3 in the context of breast cancer, potentially playing a role in the apoptotic pathway (Martin-Garabato et al, 2008). Interestingly, it seems that the role of RBM proteins in apoptosis is not solely restricted to the context of cancer. RBM5, in addition to being recognized as an RNA binding protein and tumor suppressor, also became known for being a modulator of apoptosis (Sutherland et al, 2005). After this discovery, RBM3 also showed apoptosis modulatory functions. The sequence and domain similarity between RBM5 and

each of RBM6 and RBM10 directed the attention towards these 2 RBM family proteins also playing a role in apoptosis (Sutherland et al, 2005). In addition, the X chromosome RBM proteins, namely RBMX, RBM3 and RBM10, are directly correlated with the Bax gene, the apoptosis promoter in the context of breast cancer (Martínez-Arribas et al, 2006). The structural similarity between members of the RBM family raises the possibility of other RBM proteins also somehow being involved in apoptosis.

D. RBM20

1. Overview

RBM20 (RNA-Binding Motif protein 20) is an RNA-binding protein that can recognize and bind some of the cis-acting regulatory elements on the pre-mRNA strand and carries out its function in the spliceosome. It is a major splicing factor mainly restricted to the heart and skeletal muscle (Weeland et al, 2015). Traces of RBM20 expression are also seen in the colon, small intestine and ovaries (Filippello et al, 2013). The gene is found on chromosome 10 q25.2. It is transcribed into four pre-mRNA transcripts, three of which are translated into functional proteins (Human RBM20 GeneCard). RBM20 is composed of 1,227 amino acids of 135 kDa with three functional domains: two zinc finger domains and one RNA Recognition Motif (RRM) - type RNA-binding domain (necessary for splicing regulation *in vivo*) (Filippello et al, 201324, Watanabe et al, 2018) that may allow RBM20 proteins to cluster on and repress exons (Murayama et al, 2018). It is an RNA-binding protein that is conserved within vertebrates, characterized by "a leucine (L)-rich region at the N-terminus, an arginine/serine(RS)-rich region just downstream from the RRM domain

29

and a glutamate (E) rich region between the RS-rich region and the ZnF2 domain"

(Filippello et al, 2013, Watanabe et al, 2018) (Figure 3). Other domains and motifs that

have been previously found to be shared among RBM proteins, such as the consensus

sequence RNA-binding domain (CS-RBD), ribonucleoprotein domain (RNP) and RNP

consensus sequence (RNP-CS) (Sutherland et al, 2005), remain to be specified within the

RBM20 structure. The RS-rich region, encoded by exon 9 of *RBM20* (Upadhyay &

Mackereth, 2020), contains a stretch of RSRSP that is responsible for the nuclear

localization of RBM20, crucial for its splicing control (Filippello et al, 2013, 4). Its

subnuclear localization follows a pattern similar to that of other SR proteins

(serine/arginine rich splicing factors); nuclear speckles excluded from the nucleolus

(Filippello et al, 2013). RBM20 regulates the splicing of many gene transcripts by

recognizing and binding to a UCUU sequence on the mRNA transcript, allowing it to

indirectly mediate several cellular pathways and functions (Watanabe et al, 2018). It is

required for several intron retention and exon skipping steps of alternative splicing. The

UCUU tetramer is enriched 50 nucleotides before and a 100 nucleotides after exons that are

regulated by RBM20 (Upadhyay & Mackereth, 2020). Mutations that have been identified

in the RRM domain (encoded by exons 6 through 8 of *RBM20*), such as V535I and I536T,

have been few, especially compared to mutations in the RS-rich region, and have been

shown to only partially inhibit RBM20 function. It has been recently uncovered that upon

RRM binding of the UCUU sequence on the mRNA transcript, a C-terminal α3 helix folds

to stabilize the binding after being unstructured before RRM binding to the last uracil

(Upadhyay & Mackereth, 2020). In addition, the C-terminal α3 helix plays a key role in the

recognition of the 3' uracil of the tetramer, although the former is distant from the bound

RNA and is found on the opposite side of the RRM domain (Upadhyay & Mackereth, 2020).
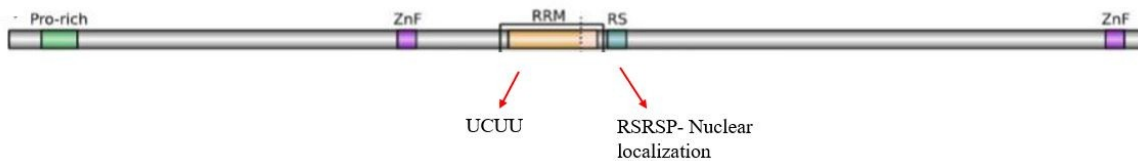
Figure 3 Functional domains of RBM20 (adapted and modified from Upadhyay & Mackereth, 2020). 2 zinc finger domains, an RNA recognition motif that specifically binds the UCUU tetramer on the pre-mRNA transcript and an arginine/serine-rich region that is responsible for the nuclear localization of RBM20 where it performs its function as part of the spliceosome.

2. *Role of RBM20 in Dilated Cardiomyopathies*

Missense mutations in the rbm20 gene have been shown to result in dilated cardiomyopathies (Brauch at al, 2009) with a 3 % frequency within patient populations (Refaat et al, 2012). The major and most studied protein regulated by splicing by RBM20 is titin, a giant sarcomeric protein (Weeland et al, 2015). In the absence of RBM20 and/or its function, the shorter titin isoform, N2B, is replaced by a larger isoform and an aberrantly giant one, N2BA and N2BA-G, respectively, characteristic of dilated cardiomyopathies (Murayama et al, 2018). Other proteins that act as splicing targets for RBM20 include calcium channel ryanodine receptor 2 (RYR2), Ca2+/calmodulin dependent protein kinase II delta (CAMK2δ), the calcium channel voltage-gated L type alpha 1C subunit

(CACNA1C/Cav1.2), formin homology 2 domain containing protein 3 (FHOD3), Z-band alternatively spliced PDZ-motif protein (ZASP/lDB3/CYPHER), and the PDZ and LIM domain protein 5 (PDLIM5/ENH) (Beqqali et al, 2016).

It has been shown that RBM20 regulates titin pre-mRNA splicing in response to external stimuli, such as insulin signaling (Zhu et al, 2017). Insulin treatment of serum-starved cells increases the percentage of the N2B titin isoform in the presence of normal RBM20 levels but not in its absence (Zhu et al, 2017). RBM20 has also been shown to play a very important role in the titin isoform switching regulated by the thyroid hormone, triiodothyronine (T3) (Zhu et al, 2015). Since insulin primarily acts on the PI3K/Akt/mTOR signaling pathway, studies have shown that blocking PI3K and mTOR activity results in mis-splicing of titin pre-mRNA. The expression levels of RBM20, being the master regulator of titin isoform switching, increase upon the activation of the PI3K/Akt/mTOR signaling by insulin or T3 (Zhu et al, 2017).

The Ttn gene transcripts regulated by RBM20 are not only the ones that get translated into the different Titin isoforms. Instead, RBM20 also regulates the production of circular RNA from the titin gene (Khan et al, 2016). Circular RNA molecules are usually produced by the spliceosome. The single-stranded RNA molecules that lack a poly-A tail are enclosed in a circle due to covalent bonding after splicing. The RNA sequence is excised from the pre-mRNA strand by exon back-splicing (Khan et al, 2016). These play important roles in the cell including contributions to gene expression. Interestingly, the hotspot for circRNAs in the pre-mRNA transcripts of titin are found within the I-band, the region most regulated by RBM20 splicing, with a few other ones close to the Z-disk and

away from RBM20 control (Khan et al, 2016). Specifically, the exons that give rise to the circRNAs are flanked by many RBM20 binding sites. Alterations in RBM20 expression and activity leading to variations in the formation of an entire family of circRNAs also contributes to heart disease, especially dilated cardiomyopathies (Khan et al, 2016).

Most missense mutations in splicing factors, and specifically RNA-binding proteins (RBPs), are found in their RRM domains. Interestingly, however, the missense mutation hotspot in RBM20 is in the RSRSP stretch within the RS domain (Murayama et al, 2018). The RRM and zinc finger domains were found to be dispensable and still repress exon inclusion in the *Ttn* transcript, favoring the N2B isoform. It was shown that phosphorylation of one or more serine residues in the RSRSP stretch was critical for splicing regulation of *Ttn* pre-mRNA by RBM20. RBM20 protein constitutively phosphorylated at this stretch localizes to the nucleus where it can carry out its role (Murayama et al, 2018). It is also speculated that this regulation on the phosphorylation-dephosphorylation level may play a role in the expression of longer titin isoforms in the embryonic heart. Additionally, phosphorylation of the Ser637 residue depends on that of the Ser639 residue and the simultaneous phosphorylation of both is critical for nuclear localization (Murayama et al, 2018). Other mutations in the sequence of RBM20, implicated in dilated cardiomyopathy and titin missplicing, have been identified in the glutamate (E)-rich region which is highly conserved evolutionarily. While a mutation in this region does not affect the subcellular localization of RBM20, it reduces the stability of the protein, even though transcript levels are normal (Beqqali et al, 2016).

3. *RBM Proteins in Cancer*

Several proteins of the same family as RBM20 have been shown to have altered expression in several tumors including breast cancer. Mutations in the gene for RBM10, mentioned above, has been shown to be directly related to the progression and development of lung adenocarcinoma and its poor prognosis. It achieves that by downregulating E-cadherin, motivating the cells to migrate and invade, and Fas, suppressing apoptosis (Yin et al, 2018). Mutations in RBM10 have been also observed in pancreatic cancer (Witkiewicz et al, 2015). RBM39 and RBM23 mutations, treated using aryl sulfonamides as mentioned earlier, are often mutated in AML (Hsiehchen et al, 2020, Ting at al, 2019). RBM38 has also been recognized as a tumor suppressor, mutated in several types of cancer including AML and breast cancer (Wampfler et al, 2016). When active, it inhibits the degradation of p53, the potent tumor suppressor that regulates cell cycle progression and advance. It also regulates p73 and/or p21^{CIP1} (the cell cycle inhibitor) stability (Wampfler et al, 2016). Its loss-of-function in cancer lifts part of the regulation and releases the breaks on the advance of the cell cycle. The loss of function of RBM38 could occur through mutations or downregulations. Interestingly, RBM38 has been recently shown to mediate miRNA-mRNA interactions, further contributing to its regulation of protein expression levels (Wampfler et al, 2016). In breast tumors, RBM38 can inhibit the proliferation of cancer cells by downregulating the expression of c-myc and upregulating that of PTEN, as well as inhibiting zonula occludins and EMT (epithelial-mesenchymal transition) (Li et al, 2017, Wu et al, 2017, Xue et al, 2014). Another member of the family involved in cancer, and

specifically breast cancer, is RBMS2 which also functions as a tumor suppressor by stabilizing p21 and, thereby, inhibiting breast cancer cell proliferation and even inducing cell cycle arrest when expressed and functional (Sun et al, 2018). As it also functions as a tumor suppressor, RBMS2 expression is downregulated in breast cancer. Moreover, RBMS3 was also found to be downregulated and RBMXL2 to be upregulated (Sun et al, 2018). Another RBM family protein, RBMY, expressed exclusively in mature male germ cells but completely absent in neoplastic male germ cells, was shown to be present in some cancer contexts (Schreiber et al, 2003). RBM3 is upregulated in some cancers like astrocytoma (Zhang et al, 2013, Schreiber et al, 2003) and RBM8A is upregulated in hepatocellular carcinoma, leading to higher levels of cancer cell migration, EMT and survival (Liang et al, 2017). However, RBM20 has never been implicated in any form of cancer to date.

4. RBM20 Regulation by the PI3K/mTOR/Akt Pathway

Just as every other protein in our cells, RBM20 expression and activation by phosphorylation are regulated by kinase cascades. RBM20 is regulated by the PI3K/mTOR/ Akt signaling pathway (Zhu et al, 2017). PI3K includes 3 different subclasses of enzymes that can all phosphorylate phosphatidylinositol. This signaling pathway can be activated by several cell surface receptors (Campa et al, 2015). Once active, the pathway goes on to stimulate several signaling effectors with domains that recognize phosphoinositides, including PH domains (Campa et al, 2015).

The control of RBM20 protein expression levels by the PI3K pathway is achieved through two downstream antagonistic substrates of mTOR that get phosphorylated upon the activation of the latter; 4E-BP1 promotes gene expression upon phosphorylation and increases RBM20 expression while p70S6K1 reduces gene expression upon phosphorylation and decreases RBM20 expression (Zhu et al, 2017).

AMP-activated protein kinase (AMPK) is a heterotrimeric protein that gets activated when the intracellular ratios of AMP/ATP and/or ADP/ATP are high (Ling et al, 2020). Exchange of ATP to AMP or ADP at the gamma subunit results in the phosphorylation of the alpha subunit in the kinase domain activation loop. The phosphorylation could be carried out by LKB1 or CaMKK2. However, phosphorylation of certain serine residues in the alpha subunit has the opposite effect and suppresses AMPK activity (Ling et al, 2020). mTOR and AMPK are both major nutrient sensors that are highly involved in the regulation of cell growth and metabolism. AMPK is activated upon energy stress from nutrient depletion. Low AMPK activity was found to occur simultaneously with high mTORC1 activity (Ling et al, 2020). Interestingly, inhibition of mTORC1, but not mTORC2, resulted in a reduction in serine phosphorylation in the alpha subunit of AMPK, allowing the latter to be activated. mTORC1 also inhibits the phosphorylation of tyrosine residues that activate AMPK (Ling et al, 2020). In short, mTOR activity increases inhibitory phosphorylation and decreases activating phosphorylation on AMPK. This implies that mTORC1 can directly inhibit AMPK activity (Ling et al, 2020). Conversely, AMPK can inhibit mTOR but indirectly by activating tuberous sclerosis complex 2 (TSC2) (Ling et al, 2020). As AMPK activation inhibits

cellular proliferation under conditions of nutrient stress, the ability of mTORC1 to inhibit AMPK and allow cell growth and proliferation seems as a plausible mechanism used in tumor microenvironments.

As RBM20 expression and activation are controlled by the PI3K/mTOR pathway (Zhu et al, 2017) and mTOR and AMPK are mutually regulated by one another, this suggests that AMPK is implicated in RBM20 expression control. As mentioned above, mTOR activation leads to the inhibition and/or inactivation of AMPK while the latter can also inhibit mTOR through this negative feedback loop (Ling et al, 2020). Therefore, it is possible that when AMPK is activated, it will regulate the PI3K signaling pathway by directly affecting mTOR which will, in turn, lead to a reduction in or inhibition of RBM20 expression.

5. *Phage Display Biopanning Assay for RBM20 Interactions*

Previously in our lab, phage display biopanning assays were performed to determine the proteins that potentially interact with RBM20 in a direct protein-protein interaction.
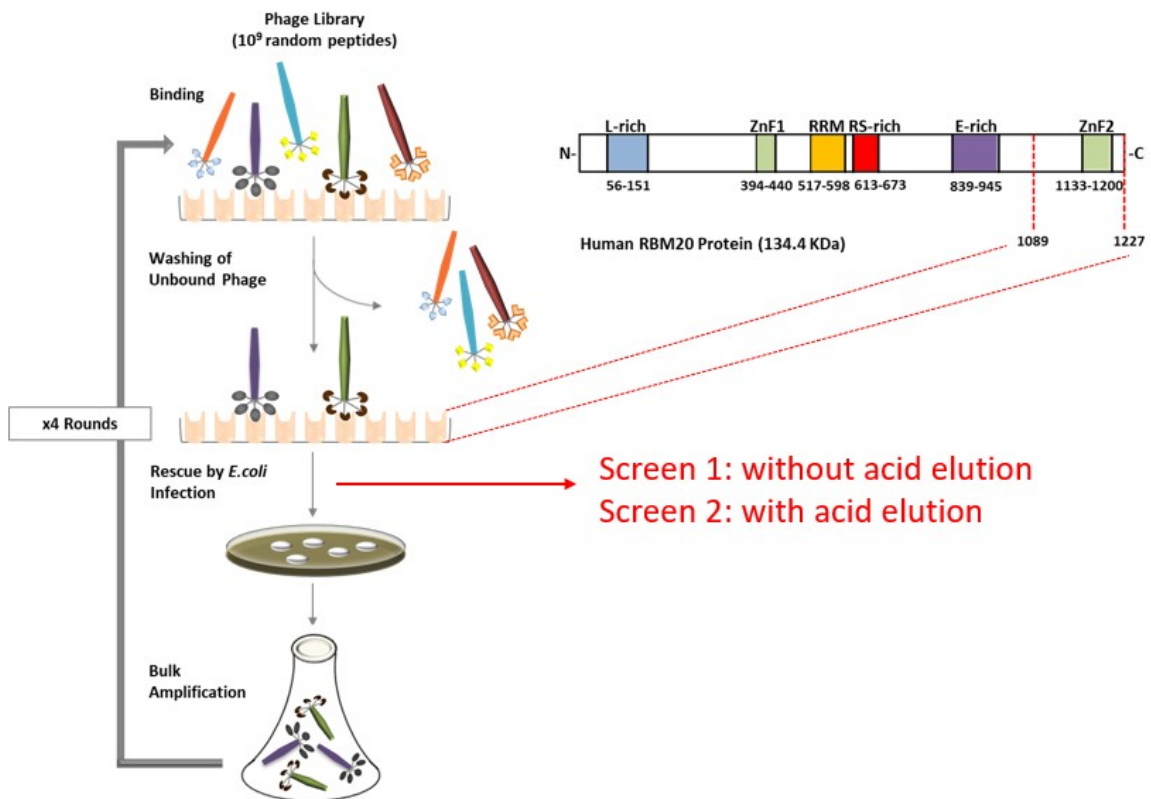
Figure 4 Phage Display Biopanning Assay procedure (Adapted From Zahr et al., 2019). Plated RBM20 C-terminus incubated with phage library. Performed either with or without acid elution. Bound phages isolated, amplified and sequenced.

For that purpose, the C-terminus of the RBM20 protein (C-RBM20) was plated on a petri dish, to which a phage library was added (Figure 4). The phage library was constituted of 10^9 phages, each of which carrying distinct random peptide sequences of 7 amino acids on its surface. After incubating the phage library with the plated RBM20 C-terminus, the plate was washed to remove the phages that did not bind the plated protein. To make sure

the binding was specific, this was performed over four rounds to be more certain of the specificity of the phages that bound. The procedure was performed either with acid elution to determine the high affinity binders or without acid elution to look for the low affinity binders. The phages that bound to the C-terminus of RBM20 during the four rounds of each of with or without acid elution were collected, amplified and had their genomes sequenced to determine the amino acid sequence of the peptides that succeeded in interacting with C-RBM20. After that, each of the specified 7 amino acid sequences was computationally blasted against the library of known proteins to look for homology between the isolated peptide sequences and motifs within proteins. One of the high affinity peptide binders matched a 7 amino acid motif within the RhoGTPase-activating protein 25 (ARHGAP25) (experiment and analysis done by Dr. Hind Zahr) (Table 1).

Screen/ Round	Peptide (% Frequency)	Matching motif	Protein Mimic (gene)
S1/RI	TVNEGMT (98)	651-INEGMT-656 36-VNEGM-40 228-VNEGM-232 490-TVTDGMT-496 647-VTEGMT-652 1190-VNAGMT-1195	Piwi-like protein 1: PIWIL1 (Piwil1) Neurogenic locus notch homolog protein 2: NOTCH2 (Notch2) Probable cation-transporting ATPase 13A5: ATP13A5 (Atp13a5) Protein sidekick-1: SDK1 (Sdk1) Cullin-9: CUL9 (Cul9) Dedicator of cytokinesis protein 9: DOCK9 (Dock9)
S1/RII	MVRDHSV (94)	6834-MVRDH-6838 346-MIRDHEV-352 1650-MVVDHSV-1656 173-VREHSV-178 472-MVRQHS-477 622-VREHSV-627	Nesprin-1: SYNE-1 (Syne-1) Rho GTPase-activating protein 25: ARHGAP25 (Arhgap25) Protein furry homolog: FRY (Fry) Kinesin-like protein: KIF13B (Kif13b) Zinc finger protein ZXDC: ZXDC (Zxdc) Protein FAM83G: FAM83G (Fam83g)
S1/RIII	MVRDHSV (88)	6834-MVRDH-6838 346-MIRDHEV-352 1650-MVVDHSV-1656 173-VREHSV-178 472-MVRQHS-477 622-VREHSV-627	Nesprin-1: SYNE-1 (Syne-1) Rho GTPase-activating protein 25: ARHGAP25 (Arhgap25) Protein furry homolog: FRY (Fry) Kinesin-like protein: KIF13B (Kif13b) Zinc finger protein ZXDC: ZXDC (Zxdc) Protein FAM83G: FAM83G (Fam83g)
S1/RIV	MVRDHSV (100)	6834-MVRDH-6838 346-MIRDHEV-352 1650-MVVDHSV-1656 173-VREHSV-178 472-MVRQHS-477 622-VREHSV-627	Nesprin-1: SYNE-1 (Syne-1) Rho GTPase-activating protein 25: ARHGAP25 (Arhgap25) Protein furry homolog: FRY (Fry) Kinesin-like protein: KIF13B (Kif13b) Zinc finger protein ZXDC: ZXDC (Zxdc) Protein FAM83G: FAM83G (Fam83g)

Table 1 Phage display biopanning assay results (modified from Zahr, H et al in press) showing the matching 7 amino acid motifs with the corresponding proteins.

E. ARHGAP25

1. *RhoGTPases & Their Regulation*

RhoGTPases belong to the superfamily of Ras small GTPases. They can be activated by a vast range of cell surface receptors and their second messengers play important roles in cell survival by promoting cell growth and inhibiting apoptosis, in addition to regulating gene expression and modulating actin cytoskeletal dynamics (Campa et al, 2015). These small GTPases, such as Rho, Rac and Cdc42, are regulated by 3 types of proteins. First, guanine nucleotide exchange factors (GIFs) allow the dissociation of GDP from and the binding of GTP to the GTPases in order to activate the latter (Csépányi-Kömi et al, 2012). Second, guanine nucleotide dissociation inhibitors (GDIs) keep the small G proteins in their inactive state (GDP-bound). Finally, the family of GTPase activating proteins (GAPs) act by accelerating hydrolysis of GTP bound to the RhoGTPases by the inherent GTPase activity of the latter in order to reduce their activity (Csépányi-Kömi et al, 2012). Although Rac is inactive in its GDP-bound state, it can still indirectly activate some downstream substrates (Campa et al, 2015).

2. *RhoGTPase-Activating Protein 25 (ARHGAP25)*

ARHGAP25 (RhoGTPase-activating protein 25) is a Rho GTPase activator, expressed preferentially in the cytoplasm and plasma membrane of hematopoietic cells and some cells of the spleen, that acts specifically on Rac (Csépányi-Kömi et al, 2012). Studies have also shown that it can act on both Rac1 and Rac2, the latter being specific to hematopoietic cells similar to ARHGAP25 (7, Campa et al, 2015). The gene encoding ARHGAP25 is located on chromosome 2p13 (Csépányi-Kömi et al, 2012) near a recombination hotspot (Chami et al, 2014) and encodes a protein that is 639 amino acids long with the RacGAP domain between amino acids 151 and 340 (Csépányi-Kömi et al, 2012). Its N-terminus is a PH domain, present in many signaling proteins and usually responsible for binding other proteins (Simple Module Architecture Research Tool), and the C-terminus consists of a coiled coil sequence (Csépányi-Kömi et al, 2012). In addition, ARHGAP25 possesses a 200 amino acid ID between the GAP and CC domains. This domain does not have a known function although the majority of the phosphorylation events occur at that location (Wang et al, 2016). Since it is part of the Rho/Rac pathway (which induces actin cytoskeletal reorganization), it is greatly involved in cell motility and migration. ARHGAP25 was found to play important roles in the immune system and to be involved in the signaling pathways of both FcR and complement receptors (Csépányi-Kömi et al, 2012). In addition, it has been shown that a decrease in ARHGAP25 increases free radical production by neutrophils; higher ARHGAP25 levels interact with the Nox2 complex to reduce its activity (Lőrincz et al, 2014). ARHGAP25, unlike other RacGAPs, has been shown to function as a negative regulator of phagocytosis in neutrophils (Csépányi-Kömi et al, 2012). Moreover, neutrophil transmigration, another process with

high Rac pathway signaling, is also affected by ARHGAP25 activity as loss of function of the latter increases such neutrophil behavior (Csépányi-Kömi et al, 2016).

3. *ARHGAP25 in Cancer*

Importantly, it has been recently discovered that ARHGAP25 can play important tumor modulatory roles. Studies in rhabdomyosarcoma, a smooth muscle tumor that mainly affects children and adolescents, have shown that the mode of tumor cellular migration may be affected by ARHGAP25 expression levels (Thuault et al, 2016). The protein is upregulated in alveolar rhabdomyosarcoma (ARMS), contributing to its more invasive amoeboid migration (as compared to the mesenchymal mode of migration) by inhibiting Rac1 activity (Thuault et al, 2016). In addition, recent studies have illustrated how higher ARHGAP25 expression levels reduce colorectal cancer metastasis (Tao et al, 2019) and lung cancer cell growth, migration and invasion (Xu et al, 2019) through the Wnt/β-catenin pathway. In both cases, ARHGAP25 was negatively correlated with tumor size, differentiation and stage. Tao et al. showed that ARHGAP25 expression is downregulated in colorectal cancer and negatively correlated with tumor cell proliferation, migration and invasion, as well as Wnt/β-catenin signaling which is responsible for colorectal cancer metastasis. Overexpression of ARHGAP25 in colorectal cancer cell lines resulted in a decreased production of several MMPs (matrix metalloproteinases), EMT (epithelial-mesenchymal transition) markers and β-catenin, as well as an increase in E-cadherin (Tao et al, 2019). Overall, ARHGAP25 reduces invasion and metastatic spread of colorectal cancer by inhibiting the Wnt/β-catenin pathway (Tao et al, 2019). Similarly, in the case of

lung cancer, tumor cell proliferation, invasion and metastasis decreased upon ARHGAP25 overexpression which also resulted in reduced Wnt/β-catenin pathway activity. In addition, the HOXA4 transcription factor has been shown to be a key positive regulator of ARHGAP25 expression (Xu et al, 2019).

F. Gap in Knowledge, Rationale & Hypothesis

Although splicing factor alterations in cancer have received a lot of attention, RBM20, in specific, hasn't been studied in such a context before. On the other hand, there have been relatively fewer studies conducted on ARHGAP25 but it has already been implicated to play very important roles in three types of cancer. In addition, as part of a study to better understand the role of RBM20 and proteins it possibly interacts with, our lab previously demonstrated that a homologous sequence in ARHGAP25 matched a motif that directly interacts with C-RBM20. Specifically in the context of breast cancer, several gaps remain unanswered in relation to these 2 proteins and their roles: What are the expression profiles of each of RBM20 and ARHGAP25 in breast cancer? Can RBM20 and ARHGAP25 interact in direct protein-protein binding? What are the effects of altered RBM20 and ARHGAP25 expression and activity levels in breast cancer cells on the tumorigenic process? What pathways regulate RBM20 expression and activation (phosphorylation) in breast cancer?

Therefore, we hypothesized that the expression of RBM20 and ARHGAP25 is altered in the context of breast cancer, leading the two proteins to directly interact which might result in the amplification of some cancer hallmarks.

G. Objective & Specific Aims

Our overall objective is to prove the alteration in the expression of and interaction between RBM20 and ARHGAP25 and study the effects that has on the breast cancer phenotype. We also mean to look for the pathways that result in the altered RBM20 expression and activity in breast cancer cells. This will be achieved through investigating the following specific aims:

Specific Aim 1: To assess the expression of RBM20 and ARHGAP25 on both, the protein and transcript levels, in MDA-MB-231 and MCF-7 cell lines as *in vitro* models of more and less aggressive models of breast cancer, respectively.

Specific Aim 2: To validate that RBM20 and ARHGAP25 can directly interact in each of the MDA-MB-231 and MCF-7 breast cancer cell lines.

Specific Aim 3: To test how AMPK inhibition affects RBM20 expression and subcellular localization in breast cancer *in vitro* models and in relevance to cancer hallmarks.

H. Significance of the Study

This endeavor sheds light on RBM20 as another culprit protein in the splicing deregulation we see in cancer and further highlights the newly discovered involvement of ARHGAP25 in this disease. Our study suggests a new mechanism of interaction between two otherwise independent proteins, proposing ARHGAP25 as another one of the RBM20-interacting proteins, in the context of breast cancer which could contribute to the initiation and progression of tumorigenesis by affecting the hallmarks of cancer. It also deciphers possible pathways of control and expression of RBM20, at least in the context of breast cancer. In addition to understanding such underlying mechanisms, it opens the space for new potential therapeutics that target splicing deregulation or aberrant protein-protein interaction.

CHAPTER II

MATERIALS & METHODS

A. Cell Lines

Breast cancer, as mentioned before, is highly variable, even within each subtype. For this reason, we chose to work with 2 cell lines that vary from each other to capture a wider spectrum of function of the proteins being studied and their implications. The chosen breast cancer cell lines to model the disease were MDA-MB-231 (HTB-26) and MCF-7 (HTB-22) from the American Tissue Cell Culture (ATCC); these will be discussed below and had been already kindly provided to our lab by Dr. Rabih Talhouk (Biology Department, American University of Beirut).

The MDA-MB-231 cell line comes from a 51 year old patient who had metastatic triple negative breast cancer. The cells are very poorly differentiated and their high aggressiveness and invasiveness are mainly due to their proteolytic degradation of the surrounding extracellular matrix. The cell line is characterized by having low expression levels of Ki-67, the protein that regulates cell growth rate and used as a proliferation marker, mentioned above. In addition, MDA-MB-231 cells have a downregulation of claudins 3 and 4 and an upregulation of EMT and mammary cancer stem cell markers (MDA-MB-231 Cell Line Profile, ECACC).

The MCF-7 cell line was retrieved from a 69 year old patient. Unlike MDA-MB-231, it has features of differentiated mammary epithelium. It is used to look for PI3K and

MAPK activity in breast cancer. It is highly sensitive to estrogen, therefore, ER positive. Like the MDA-MB-231 cell line, it is easy to grow in culture but has a significantly slower doubling time (MCF-7 Cell Information).

We meant to also compare these two cell lines to a non-transformed one, MCF-10A, also present and frozen in the lab, to look for variations in our study between the normal and cancer cells, however, the unavailability of cholera toxin during this period made that unfeasible for now.

1. *Cell Culture*

The MDA-MB-231 and MCF-7 cells were grown and propagated in tissue culture in a 5 % CO_2 and at 37˚C incubator using RPMI-1640 media (Cat. # R7388, *Sigma-Aldrich*), supplemented with 10 % Fetal Bovine Serum (FBS) (Cat. # F9665, *Sigma-Aldrich*) and 1 % penicillin-streptomycin antibiotic solution (Cat. # DE17-602E, *Lonza*). When the cells reached 80-90% confluence, they were washed with 5 mL of Phosphate Buffered Saline (PBS) with no calcium or magnesium (Cat. # BE17-517Q, *Lonza*), diluted to 1X concentration, and detached from the plate with 2 mL of Trypsin. The MDA-MB-231 cells were detached using 2X Trypsin (Cat. # BE17-160E, *Lonza*) diluted in PBS, with which they were incubated for 30 seconds in the CO_2 incubator. The MCF-7 cells were detached with 1X Trypsin + EDTA (0.2 g/500 mL) (Cat. # T3924, *Sigma-Aldrich*), with which they were incubated for 1 minute in the CO_2 incubator. The cells were then suspended with the trypsin in 6 mL of RPMI-1640 media and centrifuged at 200 g and 4˚C

for 5 minutes. The collected cell pellet was resuspended in 1 mL of RPMI-1640 media and placed in a cell culture dish, prefilled with 9 mL of growth media. The cells were frozen in RPMI-1640 media with 10 % FBS and 10 % dimethyl sulfoxide (DMSO) (Cat. # D2650, *Sigma-Aldrich*).

2. *Cell Count*

To determine an accurate number of living cells in a plate, cell counting took place by mixing the cell suspension with Trypan Blue (Cat. # T8154, *Sigma-Aldrich*) in a 1:1 ratio and loading 10 μL of the solution into each side of a Hemocytometer. Cells in the center square and the 4 directly surrounding it (a total of 5 squares) were counted at 20x magnification. The average of the cell counts was taken and multiplied by 2 (since the solution was half cells and half Trypan Blue) and by 10 (since the cells were diluted 1:10 after resuspension in media) and by 10^4 to find the number of cells per 1 mL. Upon that, the suitable volume was used in cell seeding to get the wanted number of cells for the experiment.

B. RNA Extraction, Reverse Transcription & Real-Time Polymerase Chain Reaction

Quantitative Real-Time Polymerase Chain Reaction (qRT-PCR) will be performed on RNA from breast cancer patient samples to quantify the expression levels of RBM20 and ARHGAP25 at the transcript level, relating them to protein expression levels and

studying how that might differ between the two cell lines, based on their level of aggressiveness.

1. *RNA Extraction & Quantification*

Whole RNA was extracted from MDA-MB-231 cells grown under baseline conditions in culture, described above, using the RNeasy Plus Mini Kit (Cat. # 74134, *QIAGEN*). The used protocol was adapted and modified from that of the kit. First, the cells were washed twice with 1X ice-cold PBS. 1 mL of the lysis, RLT, buffer mixed with β-mercaptoethanol (Cat. # M3148, *Sigma-Aldrich*) (for nuclease reduction) was then added to the cells in the 10 cc plate and left for a couple of minutes. The tubes are filled with 600 μL of 70 % ethanol to clean them before moving the cell lysates into them. The mixture was added to the RNeasy columns (up to 700 μL at a time). The columns were centrifuged for 25 seconds at 12,000 rpm and 4˚C and the contents were discarded from the flow-through tube. After that, 700 μL of the RW1 wash buffer were added and the centrifugation repeated. The following 2 washes were made with 500 μL of RPE buffer and centrifuged (for 15 seconds the first time and 2 minutes the second time). The collection tube was changed and the column spun again for 2 minutes to remove any remaining buffer. 30 μL then 10 μL of RNase-free water were added directly to the gel in the column and the latter centrifuged for 1 minute each time. The liquid collected at the bottom of the collection tube (same one used both times) contained the RNA sample which was moved to a clean 1.5 mL eppendorf tube. The extracted RNA samples were stored at -80˚C if not used right away.

They were quantified by measuring 1 μL of each sample against a ddwater (deionized distilled water) control using the NanoDrop 2000c from *Thermo Scientific*.

2. *Reverse Transcription*

The collected RNA was reverse transcribed into cDNA using the iScript cDNA Synthesis Kit (Cat. # 170-8891, *BIO-RAD*). 1 μg of each RNA sample was mixed with 4 μL of 5x iScript reaction mix, 1 μL of iScript reverse transcriptase and completed with molecular grade, RNase- DNase-free sterile water (Cat. # W4502, *Sigma-Aldrich*) for a total volume of 20 μL. The reverse transcription was carried out by the SimpliAmp thermal cycler (from *Thermo Fisher Scientific*) in a cycle that started with 5 minutes of annealing at 25˚C, to 30 minutes at 42˚C for elongation, 5 minutes at 85˚C for inactivation of the reverse transcriptase enzyme and holding the mixture at 4˚C for another 30-45 minutes. The resulting cDNA samples were stored at -20˚C.

3. *Quantitative Real-Time Polymerase Chain Reaction*

Quantitative Real-Time PCR was conducted using SYBR Green JumpStart Taq ReadyMix (Cat. # S4438, *Sigma-Aldrich*) on the cDNA samples from cell lines to assess the concentrations needed to be used from cDNA of patient samples and test the primers for proper functioning, making sure of the working dilutions. Two primer pairs for the RBM20 transcript and one primer pair for the ARHGAP25 transcript (Table 3) were computationally generated by and obtained from *Sigma-Aldrich*.

Gene - Species	Primer Sequence
RBM20 (Human)	Forward Primer: **5' - GCAGCCATACCCAGTACCC - 3'** Reverse Primer: **5' - CATTACCCCAGTGAAAGGATGC - 3'**
RBM20 (Human)	Forward Primer: **5' - TGTGACCTATGAAGGGCACTA - 3'** Reverse Primer: **5' - CTTGGGAGTTGGGTCCGTAA - 3'**
ARHGAP25 (Human)	Forward Primer: **5' - TAAAGCTCTACCTCCGAGACC- 3'** Reverse Primer: **5' - TTTGCCTCATCCGCATTCGT - 3'**

The primers, which were initially lyophilized, were prepared with molecular grade

Table 2 A list of the *RBM20* and *ARHGAP25* primer pairs and their sequences to be used in qRT-PCR to determine their transcript expression levels in cell lines and patient samples.

water to reach a concentration of 100 μM. A working dilution of 1:10 of each primer was

prepared from the original vial also by dilution in molecular grade water. The cDNA

samples were each diluted by a ratio of 1:20 by mixing 5 μL of the sample with 95 μL of

molecular grade water. From this dilution, a series of dilutions of 1:5, 1:10 and 1: 50 were

made. A Mastermix was prepared for the number of samples for each primer by mixing 1

μL of the forward primer, 1 μL of the reverse primer, 12.5 μL of the SYBR Green

Supermix and 6.5 μL of molecular grade water per PCR tube. The RT-PCR was performed

by the C1000 Touch Thermal Cycler from BIO-RAD and analyzed by the CFX96 Real-

Time System with the cycle spanning over initial heating to 50°C for 2 minutes, DNA

double-helix opening at 95˚C for 10 minutes, primer annealing at 60˚C for 1 minute then
extension at 72˚C for 30 seconds.

C. Protein Extraction, SDS-PAGE & Western Blot Analysis

Western blot analysis was performed on whole protein extracts from *in vitro* grown
MDA-MB-231 and MCF-7 breast cancer cell lines to assess the expression levels of
RBM20 and ARHGAP25 on the protein level and how that might differ between the two
cell lines.

1. *Protein Extraction*

Whole protein extracts were obtained from MDA-MB-231 and MCF-7 cells grown
under baseline conditions in culture. The cells were washed twice with ice-old 1X PBS,
after which each 10 cc plate was incubated with 300 μL of RIPA lysis buffer (Cat. #
R0278, *Sigma-Aldrich*) supplemented with 1 % phosphatase inhibitor cocktail II (Cat. #
4160092-1, *bioWORLD*), 1 % phosphatase inhibitor cocktail III (Cat. # 4160096,
bioWORLD) and 1 % protease inhibitor cocktail III (Cat. # 22020008-1, *bioWORLD*). The
cells were incubated with the lysis buffer on a slow shaker on ice for 30 minutes. Then, the
cells were scraped off the culture plate with a plastic scraper and collected into eppendorf
tubes that were then centrifuged at 14,000 rpm and 4˚C for 10 minutes in a Heraeus Fresco
17 centrifuge from *Thermo Scientific*. The resulting supernatant was collected and moved
to another tube that was stored at -20˚C until used.

2. *Protein Quantification*

To quantify the extracted proteins, we made dilutions of 1 mg/mL of Bovine Serum Albumin (BSA) (Cat. # A2153, *Sigma-Aldrich*) in a 96-well plate. The BSA dilutions would form the standards to which the protein samples would be compared. The BSA concentrations in the standards were duplicates of 0.0, 0.2, 0.4, 0.6, 0.8 and 1 μg/μL diluted in deionized distilled water (ddH$_2$O). The next wells contained 5 μg of the sample proteins in duplicates. Then, 200 μL of Bradford Reagent (Cat. # ab119216, *abcam*) were added to each well and covered to prevent exposure to light until the plate was placed in the ELISA Reader, Multiskan EX from *Thermo LabSystems*, which measured the protein contents of the wells that were analyzed by and retrieved from Ascent Software version 2.6 and copied to an Excel SpreadSheet to be used when the R^2 value was not lower than 0.9.

3. *SDS-PAGE*

a. <u>Preparing & Casting the Gels</u>

The gels used to separate proteins in the samples were 4 % Tris. HCl polyacrylamide gels. The lower or resolving gel was prepared by mixing 2.67 mL of 30% acrylamide (Cat. # 161-0158, *BIO-RAD*), 2 mL of 1.5 M Tris.HCl pH=8.8 (Cat. # 161-0798, *BIO-RAD*) and 3.17 mL double distilled water. 80 μL of 10 % ammonium persulfate (APS) (Cat. # 7721-54-0, *Sigma-Aldrich*) and 8 μL of Tetramethylethylenediamine

(TEMED) (Cat. # 0761, *amresco*) were added at the same time to the solution right before

casting the gel. The mixture was then filled into a 1.5 mm cast, using a 2 mL Pasteur

pipette, reaching the lower green limit and left for 45 minutes to solidify after being

covered with a layer of isobutanol (Cat. # 538132, *Sigma-Aldrich*) for a smooth upper

surface, to be removed after solidification. The upper or stacking gel was prepared by

mixing 0.67 mL of 30 % acrylamide, 1.25 mL of 0.5 M Tris.HCl pH=6.8 (Cat. # 161-0799,

BIO-RAD) and 3 mL of ddwater. 50 µL of 10 % APS and 5 µL of TEMED were added to

the solution at the same time right before casting the gel. After being poured in the

remaining space of the cast, a 10-well 1.5 mm comb was inserted and the gel was left for 15

minutes in the cast to polymerize and solidify. In the meantime, the 1X running buffer was

prepared by mixing 14.4 g of glycine (Cat. # 161-0718, *BIO-RAD*), 2.5 g of Tris-base (Cat.

161-0716 *BIO-RAD*) and 1 g of sodium dodecyl sulfate (SDS) (Cat. # 161-0302, *BIO-

RAD*) and dissolving them in 1 L of deionized distilled water. After the gels solidified, they

were placed in the running chamber and submerged in running buffer.

b.   Preparing & loading the Protein Samples & running them in the gel

The protein samples needed to be prepared with loading buffer on ice before being

loaded into the gel wells. Based on the protein sample concentration, the volume that

needed to be loaded into the well was calculated, given that 20 % of the total volume to be

loaded was that of the loading buffer. In addition to the protein samples and loading buffer,

10 % β-mercaptoethanol was added to the mixture under a fume hood. Then, the tube was

placed in a TS-100 Thermo Shaker from *BOECO* at 95˚C for 5 minutes to denature the

proteins that were moved right back to ice after that. The samples were loaded into their respective wells in the gel, given that the first well was reserved for the protein ladder (Cat. # ab116028, *abcam*). All the samples were loaded in duplicates. The remaining running buffer was added to the running chamber and the samples were run in the gel at 200 V for 1 hour.

c. Transferring the Proteins from Gels to Blots

During the run, the transfer buffer was prepared by dissolving 14.4 g of glycine and 2.5 g of Tris-base in 200 mL of methanol, to which 800 mL of ddH$_2$O was added to reach 1 L. A polyvinylidene fluoride (PVDF) membrane (Cat. # 1620177, *BIO-RAD*) was activated by submerging it in methanol for 1 minute and moving it to a tray filled with the transfer buffer. In the same tray, 2 sponges and 2 blotting papers per membrane were soaked with the transfer buffer. When the run was over, the gel was removed from the running chamber and placed on the same tray to detach the top layer and keep the lower gel (carrying the proteins) which was then sandwiched with the activated PVDF membrane between the blotting papers (on the inside) and the sponges (on the outside) in a transfer cassette. The membrane had to be placed on the white side of the cassette while the gel on the black side since the transfer took place from the black to the white side. After the cassette was placed in the transfer chamber, it was submerged with the transfer buffer and the transfer took place at 100 V for 2 hours.

d. Blocking the Membrane & Incubating it with Antibodies

During the transfer, the wash buffer was prepared by mixing 0.01 % Tween20 in 1X PBS. Once the transfer was over, the membrane was removed from the cassette directly into 25 mL of 5 % non-fat milk (Regilait) in wash buffer, the blocking solution, in which the membrane was incubated for 1 hour on a slow shaker to block any nonspecific binding sites left on the membrane. The membrane was then washed 3 times with 20-25 mL of wash buffer for 10 minutes (per wash) before being incubated in a sealed nylon sack with a polyclonal anti-RBM20 rabbit antibody (Cat. # PAN021Hu01, *Cloud Clone Corp*), monoclonal anti-ARHGAP25 rabbit antibody (Cat. # ab181202, *abcam*) or anti-actin mouse antibody (Cat. # sc-7210, *Santa Cruz Biotechnology*), all of which were diluted in 5 % milk at dilutions of 1:200, 1:10,000 and 1:200, respectively. The incubation took place overnight at 4˚C on a slow shaker. The following day, the membrane was washed from the excess primary antibody and incubated in the same way as before with a secondary HRP-conjugated goat anti-rabbit or anti-mouse antibody (Cat. # 111-035-144, *Jackson Immuno Research*) at a dilution of 1:5000 but at room temperature for 1 hour.

e. Stripping & Reprobing

After imaging analysis (below), the membrane was washed and stripped with 10 mL of 0.1 M NaOH (stored as 1 M at 4˚C and diluted in ddwater) for 45 minutes at slow speed to remove the attached primary and secondary antibodies. It was then placed in blocking solution again before being reprobed with the other primary antibody.

4. *Western Blot Analysis*

a. Chemidoc Analysis of western Blots

The excess secondary antibody was washed away from the membrane. A 1:1 solution of ECL reagents A and B (Cat. # ab65623, *abcam*) was prepared and 2 mL was added to a membrane and incubated with it for 5 minutes in a closed dark space. In the meantime, the Chemidoc MP Imaging System from *BIO-RAD* was turned on and prepared to be used for visualization and image capturing after the membrane was removed from the ECL. The images were captured by the Image Lab version 5.2.1 software.

b. Densitometry Analysis

The images obtained from the Chemidoc visualization were analyzed using the ImageJ free Java image processing software (downloaded from: https://imagej.nih.gov/ij/download.html). The images were converted to an 8-bit scale and the peaks were represented in a percentage format. The values were copied to a Microsoft Excel Worksheet and averaged for each sample type. The RBM20 and ARHGAP25 values were normalized against the actin loading controls.

D. Immunofluorescence Staining & Imaging

Immunofluorescence was performed on *in vitro* grown MDA-MB-231 and MCF-7 breast cancer cell lines to assess the intracellular localizations of RBM20 and ARHGAP25 proteins and how they might differ between the two cell lines and from what has been reported in the literature.

1. *Immunofluorescence Staining using the PFA-Triton-X method*

Cells were counted as described above and 7×10^4 cells of the MDA-MB-231 cell line and 8×10^4 cells of the MCF-7 cell line were seeded on autoclaved square coverslips in 6-well plates. The cells were incubated with 3 mL of RPMI-1640 media over 2 nights in the CO_2 incubator to reach 80 % confluence. They were then washed twice with 2 mL of 1X PBS per well for 5 minutes. All washing steps were done by gently moving the plate in a square motion by hand. The cells were then fixed with 2 mL 4 % paraformaldehyde (PFA) (Cat. # 47608, *Sigma-Aldrich*) (freshly prepared from the stock of 40 % by diluting it in PBS) per well for 20 minutes at room temperature under a fume hood. Once the incubation was over, the cells were washed 4 times with PBS and permeabilized with 1.5 mL of 0.2 % Triton-X (Cat. # T9284, *Sigma-Aldrich*) diluted in PBS for 10 minutes at room temperature, after which the cells were washed 3 times with PBS. Blocking was done using 2 mL of 3 % BSA in PBS per well for 2 hours. Trial and error showed that blocking with 2 % BSA was not enough to remove all the background signal. The cells were then washed with PBS and incubated with 200 μL of primary anti-RBM20 and anti-ARHGAP25 antibodies (mentioned above) in 1 % BSA at a dilution of 1:50. The antibodies were placed on Parafilm in closed trays with the coverslips inverted over them and incubated overnight

at 4˚C. The following day, the coverslips were returned to the wells right side up and washed 3 times with PBS before being incubated in the same manner with secondary donkey anti-rabbit Alexa 488-conjugated antibody (Cat. # ab150073, *abcam*) in PBS at a dilution of 1:500 for 1 hour at room temperature. The coverslips were then washed twice with PBS and mounted on clean glass slides with UltraCruz Hard-set mounting media (Cat. # sc-359850, *Santa Cruz Biotechnology*), containing DAPI to stain nuclei, in a dark area. The slides were sealed with nail polish and placed horizontally in the dark at room temperature for a few hours before being moved to 4˚C.

2. *Immunofluorescence Staining using the Acetone-Methanol method*

Another method used to fix the cells and pemeabilize them before staining and imaging was the acetone-methanol method. The same cell concentration was used here as the one used in the PFA fixation method. The 2 protocols differ solely in the reagents used for fixation and permeabilization. After washing the cells, they were fixed with 200 μL of methanol (stored at -20˚C) for 10 minutes at -20˚C. Permeabilization was done using 200 μL of acetone (stored at -20˚C) for 1 minute at -20˚C. The remaining washing, blocking, staining and mounting steps were the same as in the PFA method.

3. *Microscopic Visualization & Imaging*

Imaging took place 48-72 hours after mounting. An upright fluorescent microscope, Leica DFC 7000 T (CRSL facility, AUB), was used for the visualization and image acquisition of the stained slides by the Leica Application Suite X software. A 40X magnification was used and 6 frames were captured per slide. The staining was done in duplicates for each protein and cell line. Each cell was exposed once to the fluorescence. The fluorescence filters used were the 440 nm one for DAPI and the 580 nm one for Alexa 488 fluorochrome. The acquired images were analyzed based on the subjective distribution of each of the RBM20 and ARHGAP25 proteins between the nucleus and/or cytoplasm.

4. *Co-localization of RBM20 & ARHGAP25*

As we aim to verify the interaction between RBM20 and ARHGAP25, we intended to look for their co-localization within the cell compartments. However, due to the unavailability of anti-RBM20 and anti-ARHGAP25 antibodies that both work in IF staining experiments but raised in different animal species (so the specificity of the secondary antibodies does not recognize both at the same time), we pushed that to the future perspectives until the said antibodies arrive.

E. Co-Immunoprecipitation

Co-Immunoprecipitation will be utilized to verify protein-protein interaction between RBM20 and ARHGAP25 using the MDA-MB-231 and/or MCF-7 cell lines. We have optimized the protocol for the antibodies and cell lines of interest to us. The kit used was Dynabeads Co-ImmunoPrecipitation Kit (Cat. # 14321D, *invitrogen* by *Thermo Fisher Scientific*) and the protocol followed was that of the manufacturer.

1. *Antibody Conjugation*

Previous members of our team had tested the conjugation of the anti-RBM20 antibody mentioned above to the beads and concluded that it is best to do the pull-down using the antibody for the potentially interacting protein not that for RBM20 (unpublished data). Therefore, we decided to directly start by conjugating the ARHGAP25 antibody discussed above to the beads. We started by disinfecting the magnet and weighing 1.5 mg of Dynabeads which were washed with 1 mL of C1 buffer. The wash was done by pipetting gently and pacing the tube on the magnet for 1 minute until the solution was clear then removing the supernatant. 7 μg of ARHGAP25 antibody were used per mg of beads so we mixed 31.5 μL of it with 43.5 μL of C1 and 75 μL of C2. The mixture was incubated on a rotator at 37°C overnight. Another tube was prepared with normal rabbit IgG (Cat. # 5732S, *Cell Signaling Technology*) in the same way (17.5 μL were conjugated to the beads) to be used as a negative control. The following day, each tube was washed once with 800 μL of HB, once with 800 μL of LB, twice with 800 μL of SB and then incubated with 800 μL of SB on a rotator for 15 minutes at room temperature. The beads were then suspended in 100 μL of SB and stored at 4°C for 2 days (until use).

2. Cell & Protein Collection

The extraction buffer (total volume of 18 mL) was prepared on the same day by diluting the 5X IP to 1X and mixing 3.6 mL of it with 14.4 mL of ddwater, 60 μL of protease inhibitor cocktail III, 60 μL of phosphatase inhibitor cocktail II and 60 μL of phosphatase inhibitor cocktail III (mentioned above). The tube in which the cells were to be collected was pre-weighed by placing it in the center of a small beaker with tissues on the balance to align the center of gravity. The cells, which had been grown to around 85 % confluence in 5 175 cc flasks, were washed with 2 mL of 1X PBS and detached from their flasks, as described above, using 5 mL of Trypsin per 175 cc flask and suspended in an additional 15 mL of RPMI-1640. The suspension was placed in the pre-weighed tube which was centrifuged at 200 g and 4˚C for 5 minutes. The cell pellet was then washed twice with 2 mL of 1X ice-cold PBS and centrifuged. The pellet was dried and the tube weighed again to find that the weight of the harvested cells was 0.17 mg (in the first trial) and 0.8 mg (in the second trial) (within the required 0.05-1.5 mg range). The cells were resuspended in the prepared extraction buffer in a 1:9 ratio and incubated on ice for 15 minutes and vortexed halfway through the incubation period. The tube was then centrifuged for 5 minutes at 2600g and 4˚C. The supernatant, containing the protein sample, was collected into a new tube on ice, out of which a volume of 100 μL was placed aside in another tube at -20˚C to be used as a positive control in subsequent steps.

3. Co-IP

The Last Wash Buffer (LWB) was prepared by mixing 100 μL of LWB, 400 μL of ddwater and 0.1 μL of Tween-20 for a total of 500 μL. The beads were moved to a new tube and blocked with 0.1 % BSA in PBS for 5 minutes on a rotator at room temperature. They were then washed with 900 μL of extraction buffer and the cell lysate was distributed equally between the tubes containing the anti-ARHGAP25 and IgG coupled beads which were incubated on a rotator in a 4˚C fridge for 1 hour. After the incubation, the beads were washed 3 times with 200 μL of extraction buffer and once with 200 μL of the prepared 1X LWB with which they were incubated for 5 minutes at room temperature. The solution was moved to a clean tube in which the beads were isolated as a pellet and incubated in 60 μL of EB on a rotator for 5 minutes at room temperature. The purpose of this step was to separate the protein complexes from the antibody-conjugated beads. The supernatant, containing the pure protein complexes, was transferred to a new tube placed on ice.

4. SDS-PAGE

Right after completion of the above steps, the isolated protein complexes and pure protein sample, put aside earlier (loaded 10, 35 and 55 μL in three respective wells), were prepared as the protein samples in the above SDS-PAGE protocol and loaded into the polyacrylamide gels. The steps are the same as above with the exception of probing for actin loading controls.

5. Western Blot Analysis

The blots were visualized and imaged in the same Chemidoc MP Imaging System to look for IP's of RBM20 with ARHGAP25.

F. Novus Biologicals Anti-RBM20 Antibody Optimization

To make sure the same lot of antibody is used for all the Co-IP repeats, the antibody from *Novus Biologicals* (Cat. # NBP2-27509) was tested and optimized to validate that it works on blots of proteins extracted from the cell lines we work with. For that purpose, we extracted the proteins from MDA-MB-231 cells, prepared them, ran them in a gel and analyzed them using western blotting and imaging, as described above in the sections on SDS-PAGE and western blot analysis. The antibody dilution used was 1:500 in 5 % milk, also prepared as described above.

G. Drug Inhibition of AMPK Using Dorsomorphin Dihydrochloride & Analysis

1. Cell Seeding

MDA-MB-231 breast cancer cells, cultured at 37°C and 5 % CO_2 and reaching 85 % confluence in 10 cc plates, were seeded into 6-well plates. After washing the cells, centrifuging them and resuspending them in 1 mL of RPMI media, 100 μL were added to every well. Each well contained a total volume of 4 mL (cells + media).

2. *Treatment with Dorsomorphin Dihydrochloride*

Around 24 hours after seeding, the cells were treated with the AMPK inhibitor, Dorsomorphin Dihydrochloride (Cat. # sc-361173, *ChemCruz*). The 10 mg of lyophilized powder were dissolved in 1058.4 μL of DMSO to obtain a concentration of 20 mM. A working dilution of 1 mM was prepared by adding 85 μL of 20 mM drug solution to 1615 μL of DMSO. The drug concentrations used on the cells were 10 nM and 10 μM, prepared by dilution of the 1 mM working dilution in RPMI media. The conditions in the wells were as follows: media only control, DMSO control, 10 nM of Dorsomorphin Dihydrochloride, 10 μM of Dorsomorphin Dihydrochloride and an extra well of untreated cells (media only) that was used for cell counting to ensure similarity between the repeats. The total volume per well was 3 mL. The treatment was administered at two different time points: 2 hours (delayed/IC-$_{50}$ response) and 15 minutes (for immediate-early response). Addition of drug was done in such a way that analysis of both time points was performed at the same time.

3. *Analysis*

a. Protein Extraction, SDS-PAGE & Western Blot Analysis

Protein extraction and quantification, SDS-PAGE and western blot analysis were performed on the MDA-MB-231 cells treated with Dorsomorphin Dihydrochloride and their extracts, as described in earlier sections. The anti-RBM20 antibody from *Novus Biologicals* was used. RBM20 protein levels were compared to β-actin loading control. The extracted proteins were run in SDS-PAGE right away.

b. Immunofluorescence Staining & Imaging

Immunofluorescence staining and imaging were performed on the MDA-MB-231 cells treated with Dorsomorphin Dihydrochloride, as described in earlier sections. The cells were fixed right after removing the drug and washing.

H. Statistical Analysis

All the data was collected from 3 independent repeats of each experiment with duplicates of each sample (except for IF after AMPK inhibition) and represented as a mean $\pm$ SEM (standard error of the mean). Statistical analysis was done through Microsoft Excel and R using the two-tailed Student t-test. One-way ANOVA was used for analysis of western blot results following drug treatment of cells. The statistical results were considered significant at $P < 0.1$ (*) and $P < 0.05$ (**). The graphs representing data quantifications were constructed using GraphPad Prism 8 GraphPad Prism 9.

CHAPTER III

RESULTS

A. Specific Aim 1: To assess the expression of RBM20 and ARHGAP25 on both, the protein and transcript levels, in MDA-MB-231 and MCF-7 cell lines as in vitro models of more and less aggressive models of breast cancer, respectively.

1. Western blot analysis to study the expression profiles of RBM20 and ARHGAP25 proteins in each of the MDA-MB-231 and MCf-7 cell lines.

a. <u>ARHGAP25 expression levels were significantly higher in MDA-MB-231 cells compared to MCF-7 cells, both grown under baseline conditions.</u>

In order to study the protein expression levels of ARHGAP25 in our breast cancer *in vitro* cell line models, MDA-MB-231 and MCF-7 grown under baseline conditions, we performed western blot analysis of whole protein extracts from these cells (protocol detailed above). Since we considered these cell lines to represent more aggressive and less aggressive forms of breast cancer, respectively, we also aimed to compare the expression levels of ARHGAP25 between the two.

We observed that the ARHGAP25 protein expression level in the MDA-MB-231 cell line was around 2-fold that in MCF-7 cells, when normalized to actin loading controls (Figure 6a). However, statistical analysis of the quantitative data revealed that the

difference in expression is significant only at P < 0.1 (*) and not lower (Figure 6b). These results suggest that the ARHGAP25 protein may be expressed at higher levels in more aggressive forms of breast cancer.

Figure 5 Looking at ARHGAP25 protein expression using western blot analysis. (a) Immunoblots of ARHGAP25 against actin controls from protein samples of MDA-MB-231 and MCF-7 cells (b) Graphical and quantitative representation of ARHGAP25 protein expression in MDA-MB-231 and MCF-7. Statistical analysis done using the two-tailed Student t-test. Data is represented as the mean ± SEM. Difference in expression was significant at P < 0.1. Graph made using GraphPad Prism 8.

b. <u>RBM20 showed differential isoform expression and levels between MDA-MB-231 and MCF-7 cell lines.</u>

Since several RBM family proteins have been implicated in different cancer types, we decided to investigate whether RBM20 expression levels are also altered in breast cancer. For that purpose, we performed western blot analysis of whole protein extracts from the MDA-MB-231 and MCF-7 cell lines, as *in vitro* breast cancer models, again considering that they represent more and less aggressive forms of breast cancer, respectively.

The total RBM20 protein did not show any significant difference in expression between the two cell lines (Figure 7b). However, they showed variations in the expressed isoforms (Figure 7a). Importantly, the 135 kDa isoform did not consistently appear in either cell line. The 2 isoforms at around 75 kDa appeared in both the MDA-MB-231 and MCF-7 cells but with higher expression in the MCF-7 cell line. Still, however, this higher expression was not statistically significant. What is really interesting is the presence of a novel RBM20 isoform close to 35 kDa that only appeared in the MDA-MB-231 cell line (Figure 7a). This isoform has never been reported in any RBM20 study before; however, our team has been continuously finding it to be expressed in the more aggressive forms of different types of cancer, including leukemia and rhabdomyosarcoma (unpublished data), in addition to breast cancer. Together, these results suggest that RBM20 splicing patterns may

be altered in the context of breast cancer with the emergence of a new smaller isoform that seems to be solely expressed in more aggressive cancer subtypes.
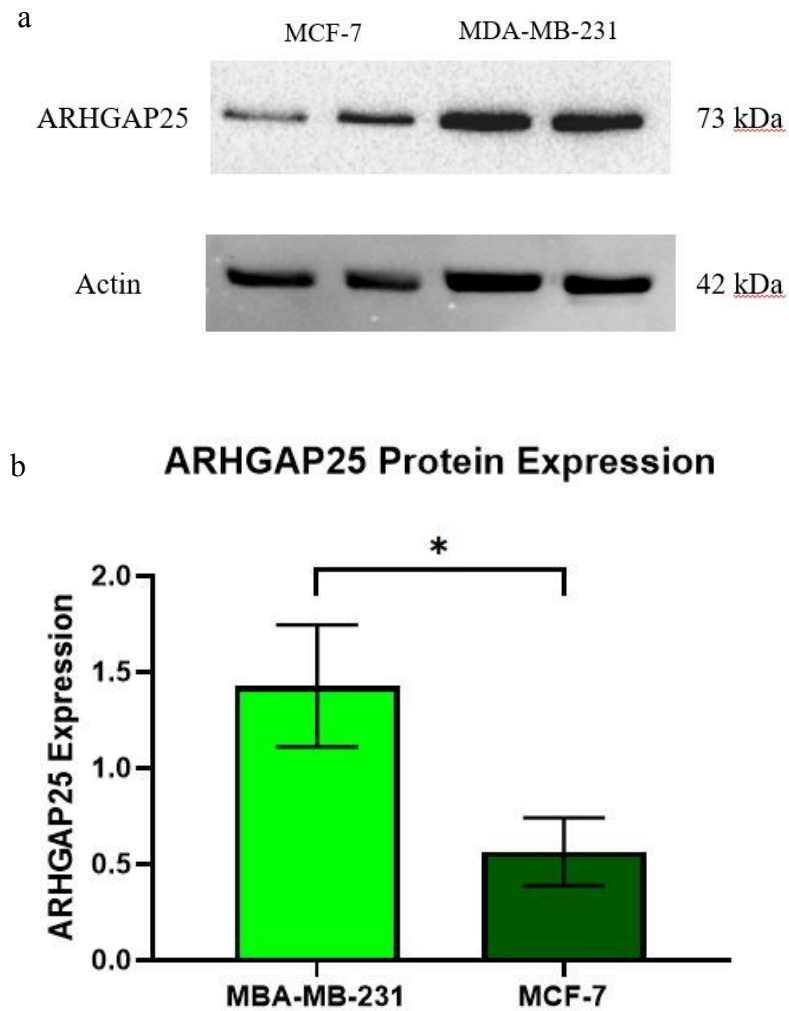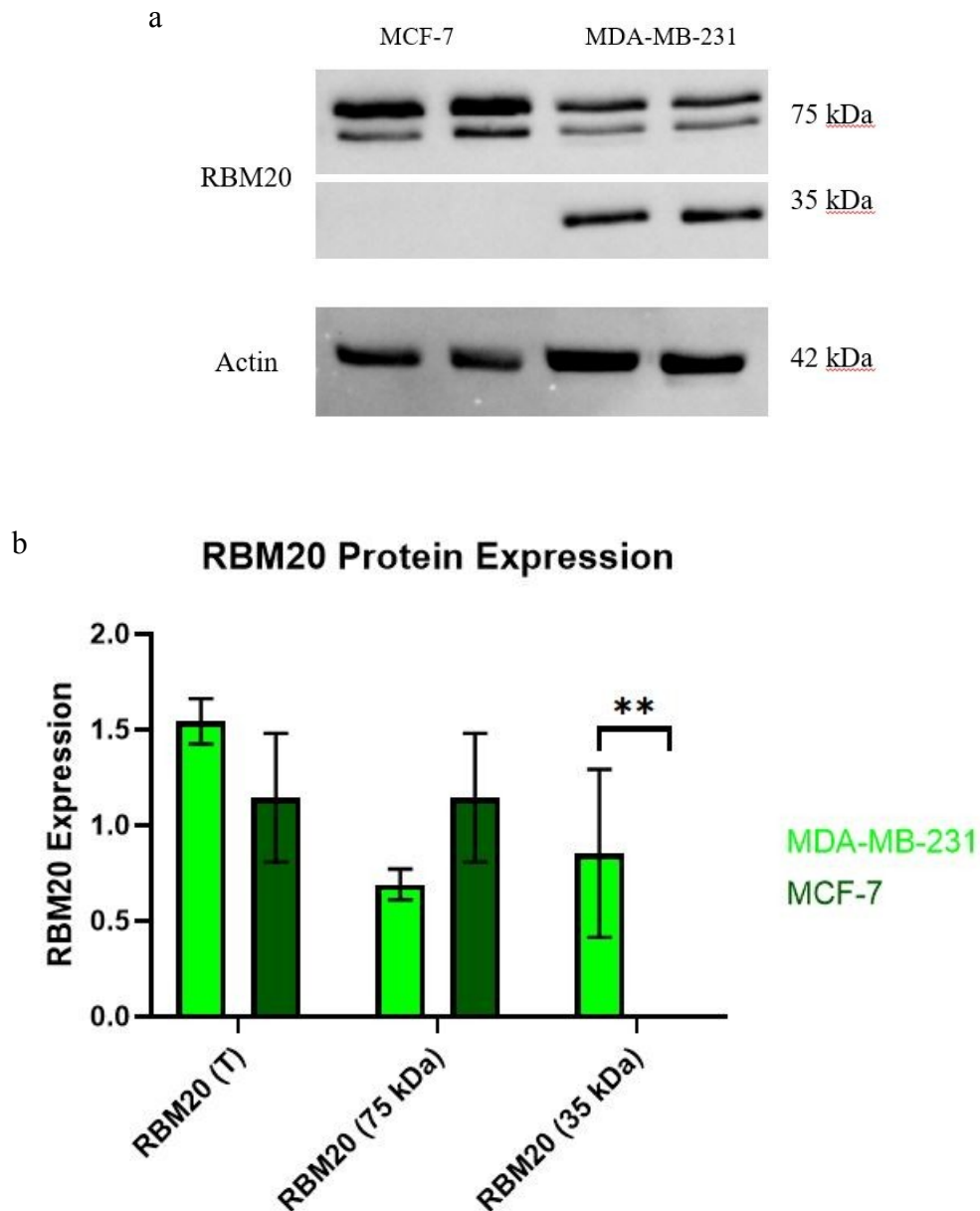

Figure 6 Looking at RBM20 protein expression using western blot analysis. (a) Immunoblots of RBM20 against actin controls from protein samples of MDA-MB-231 and MCF-7 cells (b) Graphical and quantitative representation of RBM20 protein expression in MDA-MB-231 and MCF-7. Statistical analysis done using the two-tailed Student t-test. Data is represented as the mean ± SEM. Difference in expression was significant at $P < 0.05$. Graph made using GraphPad Prism 8.

2. *Immunofluorescence to visualize whether the subcellular localizations of RBM20 and ARHGAP25 proteins is normal or altered in the context of breast cancer in both the MDA-MB-231 and MCF-7 cell lines.*

a. The Acetone-Methanol fixation method did not work as efficiently for the MDA-MB-231 and MCF-7 cell lines with the anti-RBM20 and anti-ARHGAP25 antibodies used.

Over 2 trials, we performed immunofluorescence staining on MDA-MB-231 and MCF-7 cells using the same anti-RBM20 and anti-ARHGAP25 antibodies discussed above in both cases. In addition to protocol optimization gained from these trials, it was evident that the PFA-Triton-X method was more efficient than the acetone-methanol method in fixing and permeabilizing the cells for adequate staining and visualization (Figure 7). In some cases, the cells were even damaged and their structures were not preserved when viewed under phase contrast microscopy (data not included).

DAPI   ARHGAP25

MDA-MB-231

Figure 7 A representative panel of the unsuccessful immunofluorescence staining of RBM20 and ARHGAP25 in MDA-MB-231 cells. The nuclei were stained with DAPI (blue) and RBM20 and ARHGAP25 with Alexa 488 (green) and visualized with an upright fluorescent microscope under 40x magnification.

b. <u>ARHGAP25 subcellular localization in both MDA-MB-231 and MCF-7 cells was the same as in normal cells and present in the cytoplasm.</u>

Under normal conditions, ARHGAP25 is localized in the cytoplasm with some presence on the cytoplasmic side of the plasma membrane (ARHGAP25 GeneCard). To check whether this localization remains as is in breast cancer, we immunostained ARHGAP25 as described in previous sections and visualized it under an upright fluorescent microscope (Figure 8a).

We observed that ARHGAP25 was localized solely in the cytoplasm in both MDA-MB-231 and MCF-7 cells (Figure 8b – Table 1) as 100 % of the MDA-MB-231 cells and 98.6 % of MCF-7 cells showed ARHGAP25 localization strictly in the cytoplasm (Figure 8c). We predict that its minor localization in the plasma membrane was also be present but

we could not directly observe that and distinguish between the cytoplasm and plasma membrane using the mentioned imaging techniques. For further verification, we will be using confocal imaging than upright fluorescence imaging. In addition, the nuclei appear to have holes in several instances (Figure 8a). The green immunostaining for ARHGAP25 appears within these holes as well which further supports its localization in the cytoplasm and not the nucleus. We conclude from these results that ARHGAP25 subcellular localization is not altered from the normal context in breast cancer.

a

MCF-7   MDA-MB-231

Merge

ARHGAP25

DAPI

b
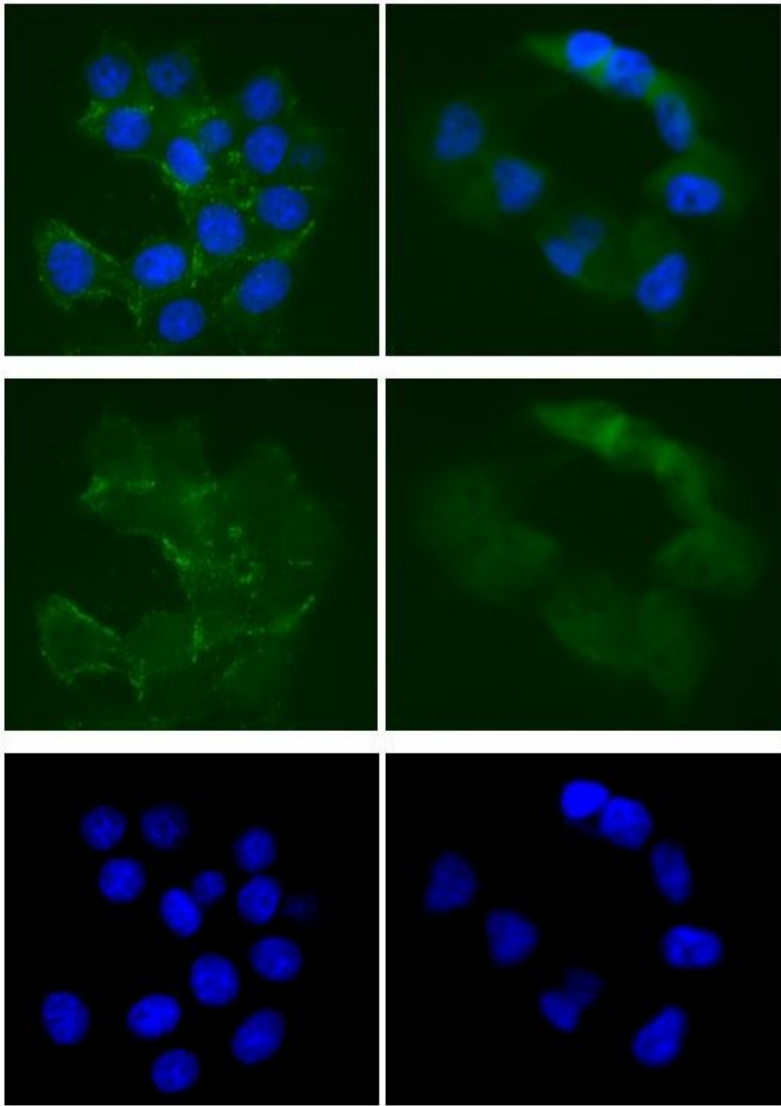
ARHGAP25 Subcellular Localization in MDA-MB-231 and MCF-7 Cells
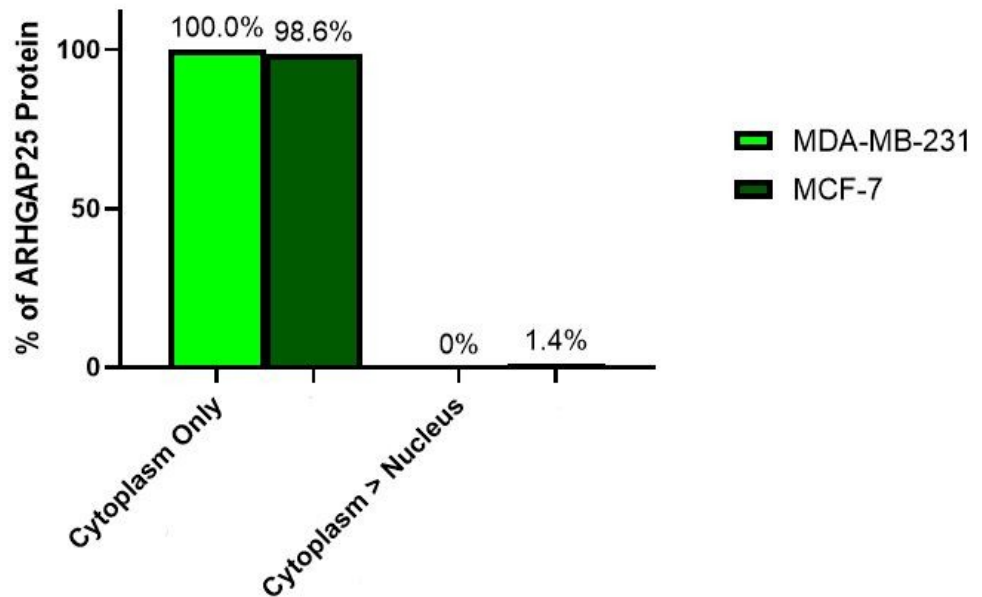
Figure 8 Immunofluorescence imaging of ARHGAP25 subcellular localization. (a) Upright fluorescent microscope images of MDA-MB-231 and MCF-7 cells stained with DAPI for nuclei visualization and in green for ARHGAP25 (b) A graphic representation of the % distribution of ARHGAP25 subcellular localization, also comparing between the 2 cell lines. Graph made using GraphPad Prism 8.

ARHGAP25 Subcellular Localization

	Cytoplasm Only	Cytoplasm > Nucleus
MDA-MB-231	100 %	0 %
MCF-7	98.6 %	1.4 %

Table 3 Subjective categorization of the percentage of ARHGAP25 protein in the cytoplasm only or cytoplasm more than the nucleus. The presented results are the average of cell counts from 3 independent repeats.

c. <u>RBM20 subcellular localization in the MDA-MB-231 cell line was similar to that in normal cells, with expression mainly in the nucleus and only some in the cytoplasm. However, in the MCF-7 cell line, RBM20 was more evenly distributed between the nucleus and cytoplasm with significantly higher cytosolic presence than normal contexts.</u>

In normal cells, RBM20 is mainly localized in the nucleus where it carries out its role, with evident but little presence in the cytoplasm, probably to regulate its function and level of activity (Filippello et al, 2013). To visualize whether its subcellular localization remains the same or is altered in the context of cancer, we immunostained RBM20 and visualized it under an upright fluorescent microscope.

Unlike ARHGAP25, RBM20 subcellular localization varied between MDA-MB-231 and MCF-7 cells. In the MDA-MB-231 cell line, RBM20 subcellular localization remained primarily in the nucleus with less presence in the cytoplasm, as should be (Figure 9a). The majority of the MDA-MB-231 cells, 81.9 %, had RBM20 localized in the nucleus more than in the cytoplasm, 11.1 % had equal localization between the nucleus and the cytoplasm and 7 % had RBM20 localized in the nucleus alone (Figure 9b – Table 4). In the MCF-7 cell line, on the other hand, the opposite was observed, as RBM20 was comparably largely found in the cytoplasm with little localization in the nucleus (Figure 9b – Table 4). The large share of MCF-7 cells exhibited equal RBM20 localization between the nucleus and the cytoplasm, making up 61.2 % of the cells, while 34.5 % showed more expression in the nucleus than the cytoplasm and 4.3 % had RBM20 restricted to the nucleus alone (Figure 9c). In addition, RBM20 localization in the cleavage furrow of several dividing

cells within the frames was evident, which is another characteristic of RBM20; however, the function it plays there is still unknown. These results predict that RBM20 subcellular localization remains intact in the MDA-MB-231 cell line, while it may be significantly altered in MCF-7 cells, pointing towards a possible relationship between the isoform variation and subcellular localization of the protein in breast cancer.

a

MCF-7   MDA-MB-231

Merge

RBM20

DAPI

b
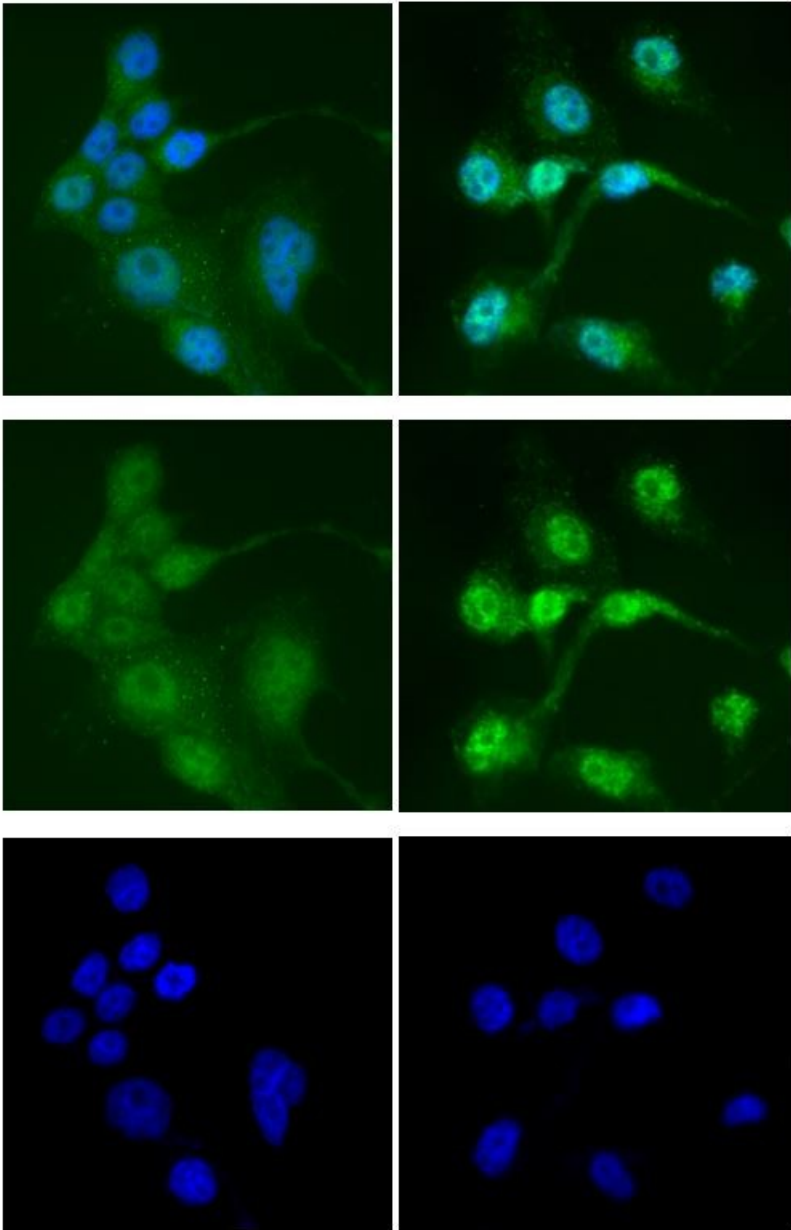
RBM20 Subcellular Localization in MDA-MB-231 and MCF-7 Cells
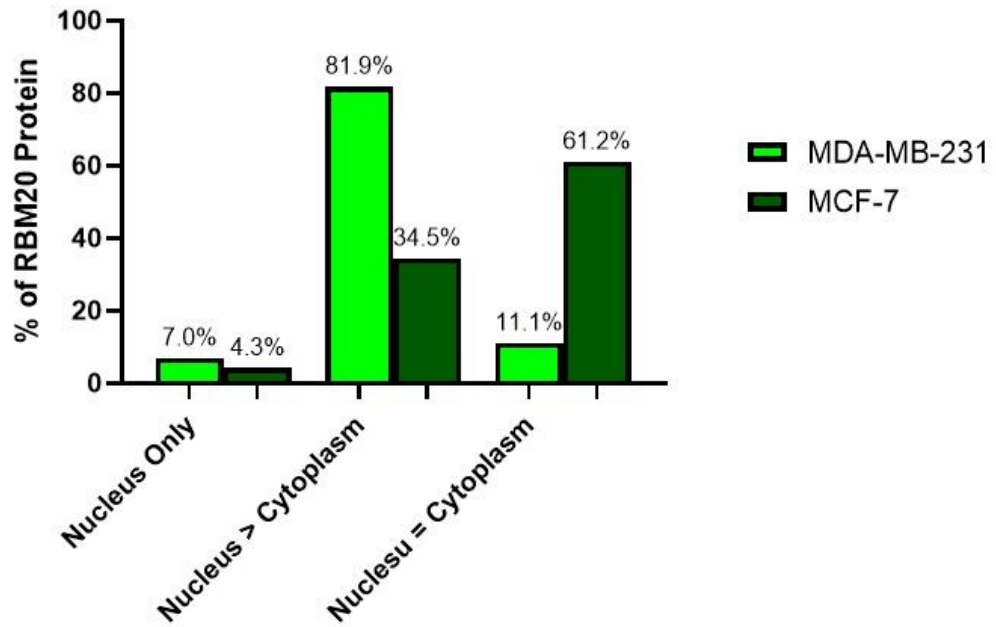
Figure 9 Immunofluorescence imaging of RBM20 subcellular localization. (a) Upright fluorescent microscope images of MDA-MB-231 and MCF-7 cells stained with DAPI for nuclei visualization and in green for RBM20 (b) A graphic representation of the % distribution of RBM20 subcellular localization, also comparing between the 2 cell lines. Graph made using GraphPad Prism 8.

RBM20 Subcellular Localization

	Nucleus Only	Nucleus > Cytoplasm	Nucleus = Cytoplasm
MDA-MB-231	7 %	81.9 %	11.1 %
MCF-7	4.3 %	34.5 %	61.2 %

Table 4 Subjective categorization of the percentage of RBM20 protein in the nucleus only, nucleus more than the cytoplasm or nucleus same as the cytoplasm. The presented results are the average of cell counts from 3 independent repeats

3. PCR dilution plots to test the primers and optimize their dilutions for use on patient samples.

In order to assess the expression of RBM20 and ARHGAP25 at the transcript level, we mean to perform qRT-PCR on breast cancer patient-derived samples and compare it to the protein levels of each, respectively. Prior to doing that, we needed to test the primers and optimize the protocol. For that purpose, we performed dilution plots of the 3 primers we mean to use; 2 primer pairs for the RBM20 mRNA transcript and 1 primer pair for the ARHGAP25 mRNA transcript (Table2).

RNA was extracted from the MDA-MB-231 and MCF-7 cells, as described in the Methods section, and a cDNA library was created using reverse transcription.

From the 2 performed dilution plots, we chose a maximum number of cycles for all samples that satisfies the three primer pairs and the two cells lines. We picked 45 cycles

since, beyond that point, the curves start to plateau. We plan to analyze the qRT-PCR results using the simple 2-tailed Student t-test.

B. <u>Specific Aim 2</u>: To validate that RBM20 and ARHGAP25 can directly interact in each of the MDA-MB-231 and MCF-7 breast cancer cell lines.

1. Co-Immunoprecipitation Optimization

a. <u>Co-IP Optimization</u>

In order to verify the potential interaction between RBM20 and ARHGAP25 from the hits observed in the phage display biopanning assay (Table 1), we decided to perform a co-immunoprecipitation of the two proteins. The optimization trial, detailed previously, revealed that we can only pull-down with the anti-ARHGAP25 antibody conjugated to the beads but not with the anti-RBM20 antibody. The same volumes and amounts detailed above will be used with the exception of the original cell weight which needs to be increased. We will aim at doubling the number of 175 cc flasks (80-85 % cell confluence) from 5 to 10 flasks in order to get a larger cell weight, increasing the yield.

b. <u>Novus Anti-RBM20 Antibody Trial & Optimization</u>

To make sure we use the same lot of antibody for the Co-IP pull-down, we tested a monoclonal anti-RBM20 antibody from *Novus Biologicals* to be used for all the Co-IP repeats. After Chemidoc imaging, one band was visible at 75 kDa with a slight streaking (Figure 10). This verifies that this antibody detects one RBM20 isoform with high sensitivity at a 1:500 dilution; however, for future western blot experiments, a higher dilution would be better to avoid the streaking observed. For Co-IP, on the other hand, the same dilution would be suitable to detect the relatively lower amount of protein that may be collected in this case. As mentioned, we will be using the ARHGAP25 antibody and not this RBM20 antibody for conjugation to beads and pull down.

RBM20   Ladder

━━ 75 kDa

Figure 10 Immunoblot of RBM20
from MDA-MB-231 protein
extracts using the Novus antibody
at a dilution of 1:500.
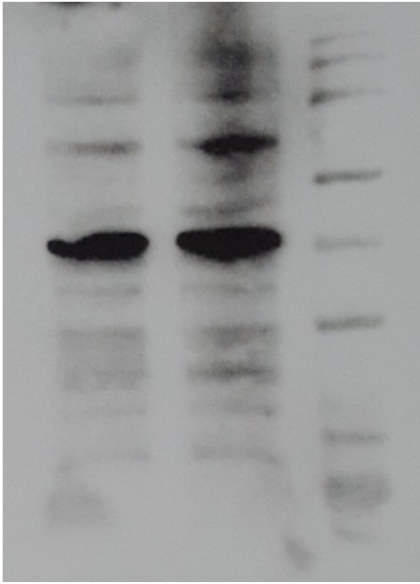
C. Specific Aim 3: To test how AMPK inhibition affects RBM20 expression and subcellular localization in breast cancer *in vitro* models in relevance to cancer hallmarks.

1. AMPK inhibition by Dorsomorphin Dihydrochloride, followed by western blot or immunofluorescence imaging, to check for variations in RBM20 protein expression and localization.

Since AMPK is a component of the PI3K/mTOR pathway that has been shown to be involved in RBM20 protein expression and phosphorylation, we aimed at testing whether inhibiting it using the drug Dorsomorphin Dihydrochloride, which acts by competing with ATP binding to AMPK, would alter the expression and phosphorylation levels of RBM20.

a. RBM20 expression levels did not not seem to be affected by inhibiting AMPK in MDA-MB-231 cells using Dorsomorphin Dihydrochloride, 2 hours or 15 minutes after the treatment.

To test the effect of AMPK inhibition on RBM20 expression levels, we treated MDA-MB-231 cells with Dorsomorphin Dihydrochloride, as described earlier, and semi-quantified the RBM20 protein levels using western blot analysis.

Densitometry and statistical analyses of the protein blots from whole protein extracts of MDA-MB-231 cells showed no significant difference (neither at $p = 0.05$ nor $p = 0.1$) in the expression levels of the 3 RBM20 isoforms that appeared, neither after 15 minutes nor 2 hours of treatment. In addition, there was no difference between the expression of the three isoforms within the cells treated with the same drug concentration and for the same period of time (Figure 11a). Importantly, the smallest and novel isoform close to 35 kDa, previously discussed, persisted under all conditions (Figure 11b). We conclude that under the described experimental conditions and analysis tools, AMPK inhibition does not affect RBM20 protein expression levels.

a

	15 mins Media	15 mins 10 nM	15 mins 10 µM	2 hrs Media	2 hrs 10 nM	2 hrs 10 µM	

RBM20

75 kDa

35 kDa

Actin

42 kDa

b

Effect of AMPK Inhibition on RBM20 Expression

Legend:
- 15 mins media
- 15 mins 10 nM
- 15 mins 10 µM
- 2 hrs media
- 2 hrs 10 nM
- 2 hrs 10 µM

RBM20 isoforms: above 75, below 75, around 35
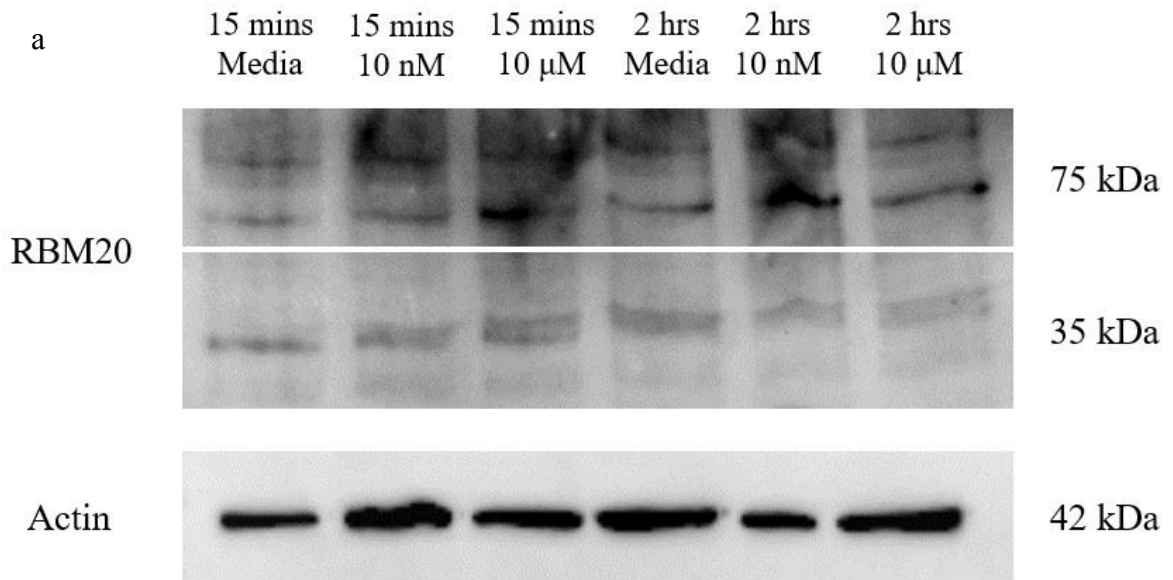
RBM20 Isoform Expression 15 mins After Treatment
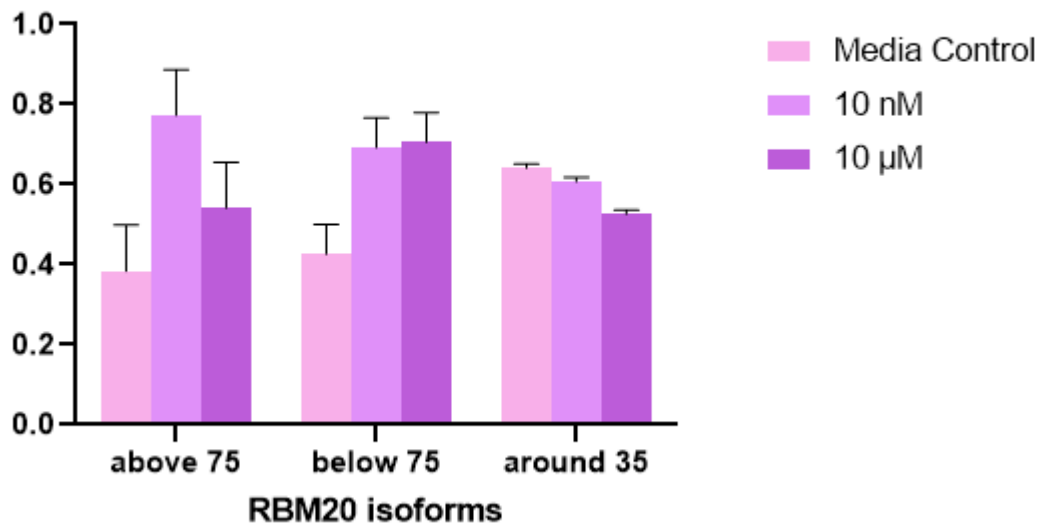
RBM20 Isoform Expression 2 hrs After Treatment
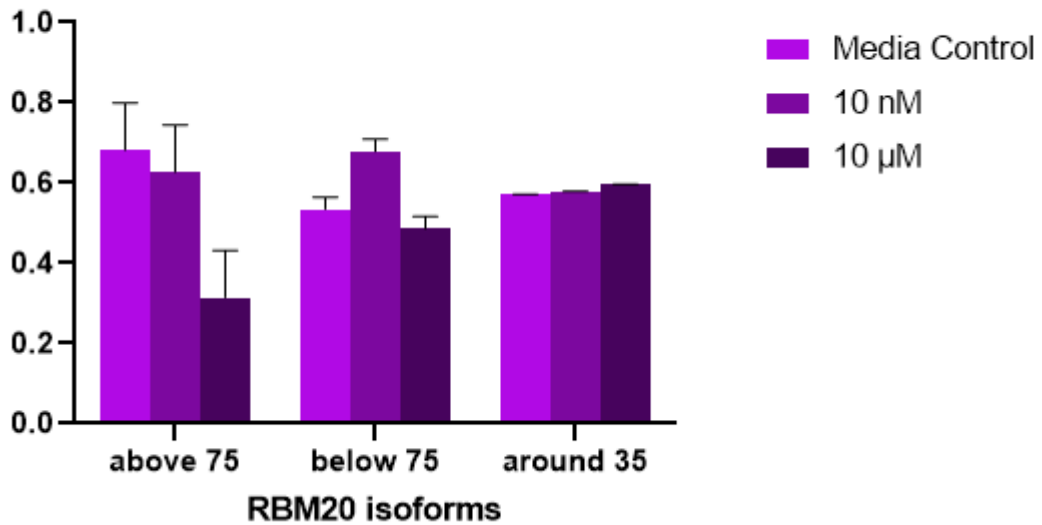
Figure 11 RBM20 protein expression after treatment with 10 nM and 10 μM of Dorsomorphin Dihydrochloride for 15 minutes and 2 hours. (a) Semi-quantification of RBM20 expression levels using western blot. (b) Immunofluorescence staining: upright fluorescent microscope images of MDA-MB-231 cells stained with DAPI for nuclei visualization and in green for RBM20. Graphs made using GraphPad Prism 9.

b. <u>RBM20 subcellular localization suggested preference towards more cytoplasmic localization after treatment of MDA-MB-231 cells with 10 nM and 10 μM of Dorsomorphin Dihydrochloride for 15 minutes, compared to the experimental controls used.</u>

The subcellular distribution of RBM20 shows that it is mainly localized in the nucleus with some presence in the cytoplasm (Filipp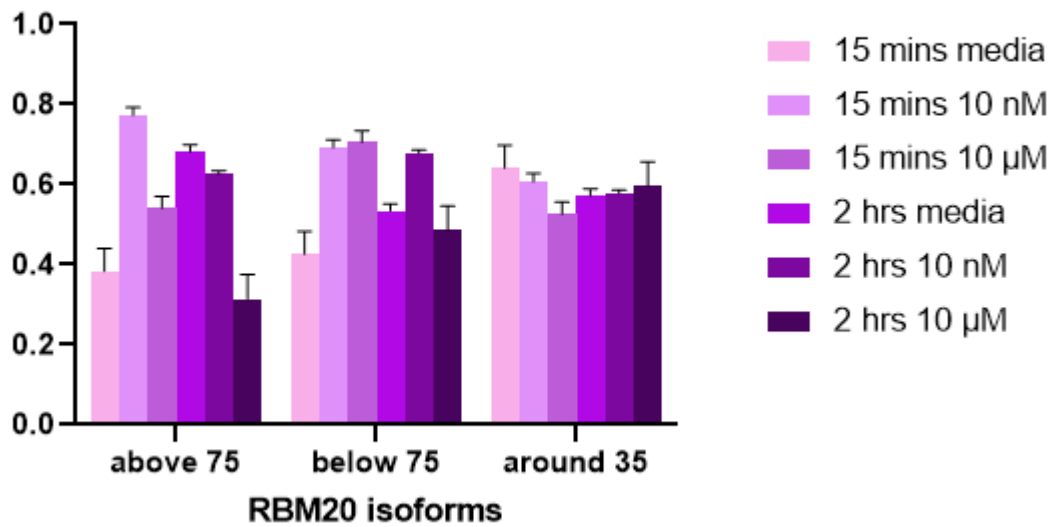ello et al, 2013). As a preliminary experiment to test the effect of AMPK inhibition on RBM20 phosphorylation levels, we treated MDA-MB-231 cells with Dorsomorphin Dihydrochloride, as described earlier, and visualized the subsequent subcellular localization of RBM20 since its phosphorylation state directly interferes with its nuclear vs. cytoplasmic localization.

Under the described experimental conditions, we observed an increase in RBM20 cytoplasmic localization in the treated cells (Figure 12a). This increase is more evident after 15 minutes with the drug (from 78.1 % to 92.3 % and 97.1 % at 10 nM and 10 μM of drug, respectively) than in the case with 2 hours of incubation (from 88% to 92.5 % and 93.3 % at 10 nM and 10 μM of drug, respectively) (Figure 12b – Table 5). These results suggest that, although RBM20 expression levels may not be affected by AMPK inhibition, its phosphorylation state could possibly be altered, resulting in a shift towards more cytoplasmic localization. It is worthy to note, however, that the presented results are from one repeat of the experiment. It is important to repeat the immunofluorescence staining and imaging after more rounds of treatment to confirm these results using confocal fluorescence imaging and subcellular fractionation.

| | Merge | RBM20 | DAPI |

a

15 mins – media

15 mins – 10 nM
c

15 mins – 10 μM

2 hrs – media

2 hrs – 10 μM

2 hrs – media
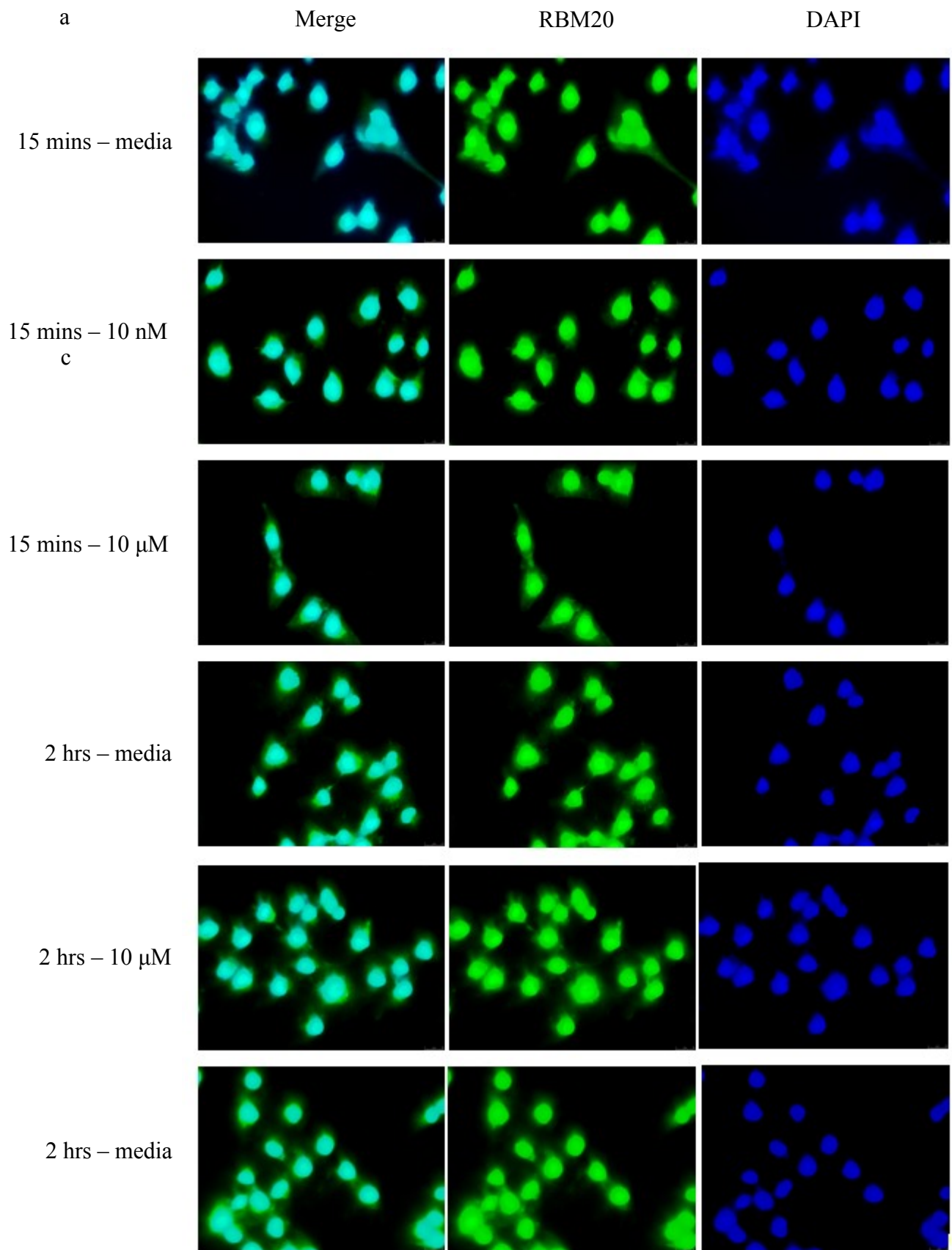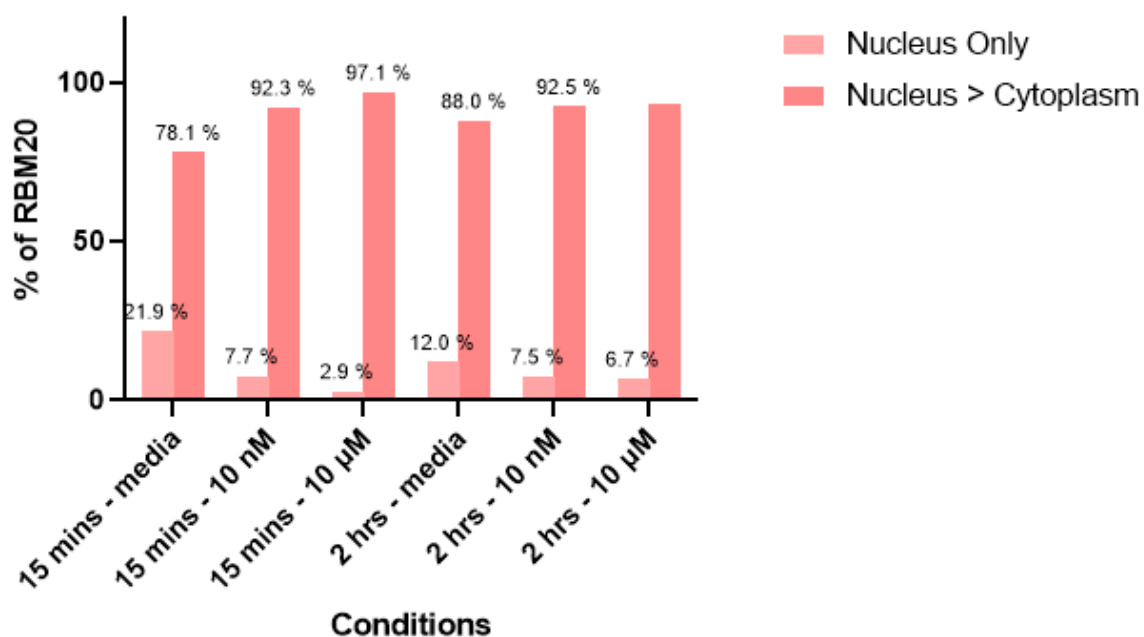
Figure 12 Immunofluorescence imaging of RBM20 subcellular localization in MDA-MB-231 cells after treatment with Dorsomorphin Dihydrochloride. (a) Upright fluorescent microscope images of MDA-MB-231 cells stained with DAPI for nuclei visualization and in green for RBM20 (b) A graphic representation of the % distribution of RBM20 subcellular localization after treatment. Graph made using GraphPad Prism 9.

RBM20 Subcellular Localization After Treatment

Treatment	Nucleus Only	Nucleus > Cytoplasm
15 mins – media	21.9 %	78.1 %
15 mins – 10 nM	7.7 %	92.3 %
15 mins – 10 µM	2.9 %	97.1 %
2 hrs – media	12 %	88 %
2 hrs – 10 nM	7.5 %	92.5 %
2 hrs – 10 µM	6.7 %	93.3 %

Table 5 Subjective categorization of the percentage of RBM20 protein in the nucleus only or nucleus more than the cytoplasm.

CHAPTER IV

DISCUSSION

RBM20 is a splicing factor implicated in cardiomyopathies and other cardiac diseases, especially that the majority of its substrate transcripts are those of proteins that play key roles in cardiac development and function, such as titin. Although members of its family, including, but not limited to, RBM3, RBM5 and RBM10, have been shown to have altered expression and function in different types of cancer, RBM20 has not been reported to play any role in cancer to date. ARHGAP25, on the other hand, is a RhoGTPase-activating protein that has been recently implicated in three different types of tumors, namely rhabdomyosarcoma, lung and colorectal cancers. However, its function differed between different cancers, as described earlier. We aimed to study the roles of RBM20 and ARHGAP25 in the context of breast cancer.

This study shows that total RBM20 protein expression levels do not seem to vary significantly between the MDA-MB-231 and MCF-7 cell lines, under the used experimental conditions. However, what differs is the isoform expression as the 135 kDa isoform is not consistently expressed in either cell line while the 75 kDa isoforms are present in both but at higher levels in MCF-7 cells. Of interest, is the novel isoform seen at around 35 kDa. This isoform has never been reported in the literature before and is found to be only expressed in the MDA-MB-231 cell line and not in the MCF-7 cell line. It is an isoform that consistently appears in western blot analyses of protein extracts from the more

aggressive models of breast cancer, leukemia and rhabdomyosarcoma tested in our lab. The discrepancies observed in the expressed isoforms of RBM20 may be part of the reason the protein's subcellular localization is altered, since it could potentially affect the translation of the RSRSP region and its spatial availability for phosphorylation which implies aberrant nuclear localization. This is especially the case in the MCF-7 cell line where the majority of RBM20 protein is localized equally between the nucleus and cytoplasm as opposed to the MDA-MB-231 cell line that shows more RBM20 localization within the nucleus, as should be. So, for future activities aiming to validate the interaction between RBM20 and ARHGAP25, which remains to be found in the cytoplasm in both cell lines, MCF-7 may be a better choice since both proteins would be localized in the cytoplasm, suggesting a higher chance of observing said interaction. If the interaction is to be verified, it would most probably occur in the cytoplasm of the cell. In addition to the protein-protein interaction between them, RBM20 may affect ARHGAP25 expression by carrying out its role as a splicing factor. This is a possibility since the ARHGAP25 pre-mRNA transcript contains 29 UCUU tetramers in its sequence when read forward and 26 when read backward, to which RBM20 can bind.

Our study suggests that ARHGAP25 is expressed at higher levels in MDA-MB-231 relative to MCF-7 cells. Therefore, we propose a model in which ARHGAP25 is expressed in lower levels in less aggressive forms of breast cancer as opposed to the more aggressive forms. Here, it would potentially be behaving as it does in lung and colorectal cancer by inhibiting the Wnt/β-catenin pathway, thereby reducing cancer cell invasion and metastasis. However, it would possibly be upregulated in more aggressive forms of breast cancer and

function in a manner similar to how it does in alveolar rhabdomyosarcoma, inactivating

Rac1 and inducing the switch to an amoeboid mode of migration. This is also supported by

the morphology taken on by the MDA-MB-231 cancer cells that model the more aggressive

breast tumor. It resembles the physiology taken on by cells using the amoeboid mode of

migration rather than the mesenchymal one. The cells are ellipsoidal with weak adhesions

to the substrate, whereas the MCF-7 cells assume a shape that better matches the

mesenchymal mode of migration with strong cell-substrate adhesions and well-defined

protrusions.

As presented earlier, ARHGAP25 inhibits Wnt/β-catenin signaling (Tao et al, 2019,

Xu et al, 2019), the pathway which happens to be pivotal in MaSC renewal, expansion and

long-term maintenance, in addition to playing a role in the potential transition from MaSC

to BCSC (Zhou et al. 2019). It is possible that when ARHGAP25 levels are high, it might

play a role in inhibiting steps of the tumorigenic process in damaged MaSCs. An

upregulation of ARHGAP25 in transforming MaSCs may be able to inhibit or slow down

the proliferation and expansion of these cells by blocking the Wnt/β-catenin pathway

earlier on during transformation. This also supports the notion that, by stimulating higher

levels of ARHGAP25 expression in the less aggressive breast tumors (represented by the

MCF-7 cell line in our study), the resultant inhibition of Wnt/β-catenin signaling would

lead to a reduction in tumor growth and cancer cell proliferation. That is in parallel to

reducing cancer cell invasion and metastasis as higher levels of ARHGAP25 in lung and

colorectal cancer do (Tao et al, 2019, Xu et al, 2019). In addition, it is important to note

that the triple negative subtype of breast cancer generally has an upregulation of Wnt/ β-

catenin signaling (Raghunath at al, 2019) but we showed that ARHGAP25 levels are relatively high in the MDA-MB-231 cell line which is classified as TNBC. This further supports that these higher levels of ARHGAP25 in this context might behave as they do in alveolar rhabdomyosarcoma, not only because of the cell shape, but also because of the possibly high remaining Wnt/ β-catenin activity.

As shown in the literature, members of the same family as RBM20, like RBM5, RBM6 and RBM10, when mutated in sequence or altered in function in the context of cancer, act as tumor suppressors (Anczuków & Krainer, 2016, Wampfler et al, 2016). We speculate that this is also the case for RBM20, especially since we observe smaller and novel isoforms (notably the one at around 35 kDa) in the more aggressive breast cancer cell line, MDA-MB-231. This fragmentation of RBM20 in the more aggressive forms of cancer may most probably result in a loss of function of the protein. Again, the role of alternative splicing in cancer cells becomes evident as this new short RBM20 isoform appears. Even more interesting is the possibility that RBM20 may play a role in its own splicing. This speculation arises from the study of the RBM20 pre-mRNA transcript that shows the presence of the UCUU tetramer 56 times when it is read forward and 50 times when read backward (RBM20 GeneCard), acting as potential binders of the RBM20 protein.

By studying the pyruvate kinase – muscle (PKM) pre-mRNA sequence (PKM GeneCard), we found that the UCUU tetramer, recognized and bound by the RRM domain of RBM20, is present 24 times when the pre-mRNA sequence is read forward and 16 times when read backward, which raises the question of whether RBM20 controls the switching between the PKM1 and PKM2 isoforms during different stages of tumorigenesis. The

difference between the PKM1 and PKM2 isoforms is that the former contains exon 9 but lacks exon 10 while the latter contains exon 10 but lacks exon 9 (Noguchi et al, 1986). PKM is a key player in cellular metabolism and PKM2, specifically, mediates aerobic glycolysis which is a feature of all cancer cells, contributing to the Warburg Effect (Desai et al, 2014). It used be thought that the higher PKM2 levels in cancer cells as compared to PKM1 levels were due to isoform switching. However, it was shown that many normal differentiated tissue types already expressed PKM2 as the major PKM isoform (Desai et al, 2014). Nevertheless, PKM2 levels always increased in cancer cells as compared to normal cells of the same tissue type even if it had already been the major isoform (Desai et al, 2014). What is noteworthy is how cancer cells can switch between dominant PKM1 and dominant PKM2 phases during tumor cell migration and growth, respectively. This process is what could be the result of PKM isoform switching (Filippello et al, 2013). The isoform switching is known to be controlled by the hnRNP family of splicing factors which is under the control of c-myc (David et al, 2010), making PKM highly probable to exhibit expression alterations in the presence of mutated myc in cancer. Interestingly, RBM20 is a submember of the hnRNP family of splicing factors (Filippello et al, 2013) and also happens to be a target of myc. We will be investigating the role of RBM20 in regulating this cyclic PKM isoform switching in future studies.

A class of compounds, known as aryl sulfonamides, have been successfully used as cancer treatments and their mode of action has been recognized as inducing the proteasomal degradation of some RBM proteins through ubiquitination (Kim & Abdel-Wahab, 2017). They are now used in patients with known alterations in several RBM

proteins including RBM39 and RBM23 (Ting at al, 2019). Given the structural similarity

between members of the RBM family of splicing factors, it would be worthy to test the

effect of aryl sulfonamides on RBM20 in cancer. In addition, many RBM proteins,

including RBM3, RBM5, RBM6, RBM10 and RBMX, have been implicated in apoptosis

and play an important regulatory role in that process (Martin-Garabato et al, 2008,

Sutherland et al, 2005). It would also be of interest to investigate whether RBM20 interacts

with any of the pro-apoptotic or anti-apoptotic proteins and contributes to programmed cell

death.

Another role RBM20 can potentially play in both, the normal and transformed

contexts, is the regulation of microRNA activity similar to RBM38 which can interfere

with and block microRNA-mRNA interactions (Wampfler et al, 2016). Additionally,

RBM20 can possibly be involved in cancer by controlling p53 and p21, or less commonly

p73, expression as do RBM38 and RBMS2 which do so by stabilizing their mRNAs

(Wampfler et al, 2016, Sun et al, 2018). The latter potential function supports RBM20

being a tumor suppressor like many other RBM proteins studied in cancer, especially that

our study shows fragmentation, and potential loss of function, of RBM20 isoforms.

As discussed earlier in the introduction, RBM20 expression is mediated by the

PI3K/mTOR pathway (Zhu et al, 2017), with AMPK being an important player in the

process by regulating and being regulated by mTOR (Ling et al, 2020). If the loss of

function of RBM20 in breast cancer were due to a reduction in expression, targeting AMPK

and inactivating it would be expected to lift its inhibition over mTOR, leading to an

increase in the expression of RBM20. However, according to what we observed in the

above results, the RBM20 loss of function seems to be caused by alternative splicing, resulting in much smaller and probably inactive isoforms. Nevertheless, since the phosphorylation levels of RBM20 are also affected by this pathway, we aimed to elucidate the effects of AMPK inhibition by drugs on the expression levels of RBM20, as well as its phosphorylation levels and, by consequence, its subcellular localization. As seen in the above results, RBM20 expression levels were not affected by AMPK drug inhibition (by Dorsomorphin Dihydrochloride) under the described conditions, implicating that the latter kinase may not play a direct role in RBM20 expression regulation. The AMPK branch may not be the dominantly active route of the PI3K/mTOR pathway during RBM20 expression. However, our immunofluorescence staining and imaging trial proposed that, although the bulk of RBM20 protein was still contained in the nucleus, there was a shift in RBM20 subcellular localization towards the cytoplasm in the MDA-MB-231 cells treated with Dorsomorphin Dihydrochloride. These variations were especially evident as an immediate-early response to the drug when the cells were fixed, stained and imaged after 15 minutes of treatment. Although we observed an increase in RBM20 cytoplasmic presence in the cells incubated with the drug for 2 hours (delayed response), its distribution was relatively much closer to the control cells (Figure 12). This suggests that AMPK may not be directly involved in RBM20 expression but potentially affects its phosphorylation levels and sites, which may result in variations and shifts in the protein's subcellular localization. In subsequent experiments, we aim to repeat immunofluorescence staining and imaging of MDA-MB-231 cells treated with Dorsomorphin Dihydrochloride to confirm the results observed in the primary test, presented above, but using confocal imaging. If the results are found to be consistent, we will be using analysis by western blot after drug treatment of

cells to look for variations in RBM20 phosphorylation levels by comparing phospho-RBM20 to total RBM20 protein levels. The described procedures will be repeated for the MCF-7 cell line as well to elucidate whether the same mode of expression and activation will be found after treatment as in MDA-MB-231 cells. In addition, other kinases of the PI3K/mTOR pathway will also be inhibited and tested to study their role in RBM20 expression and activation, specifically in the context of breast cancer.

Importantly, it has been also shown (Campa et al, 2015) that there is an interplay between the PI3K and Rac pathways. Second messengers generated by the PI3K pathway bind to PH domains that have been found in GEFs, a mechanism that may be behind the recruitment of these cytosolic proteins to the plasma membrane, leading to the activation of Rac (Campa et al, 2015). On the other hand, some studies have shown that the lipid products of the PI3K pathway can also act as inhibitors of Rac through the activation of GAPs specific to it, such as ARHGAP15 (Figure 13). However, this relationship applies both ways as activated Rac can, in turn, activate PI3K which interacts with the GTP-bound forms of Rac1 and Rac2. The feedback loop between the 2 has some important roles like the maintenance of leukocyte polarity and actin polymerization at the leading edge of migrating cells (Campa et al, 2015). This is yet another way by which RBM20 and ARHGAP25 could be linked. Although there would not seem to be a direct interaction between them in this context, but ARHGAP25 could actually possibly lead to the reduction of RBM20 activity through modulation of the PI3K pathway.
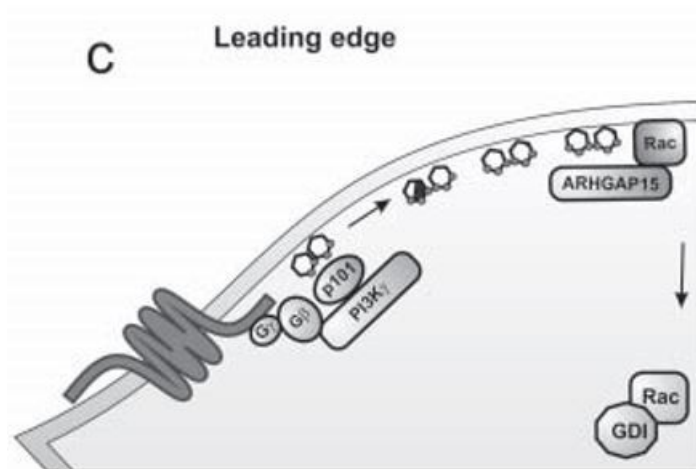
Figure 13 Inactivation of Rac by the PI3K pathway (adapted from Thault et al. 2016).

In summary, our data suggest that RBM20 undergoes isoform switching in breast cancer, resulting in the expression of the 75 kDa isoforms without the main large isoform of 135 kDa. Strikingly, a novel and much smaller isoform of around 35 kDa has been identified to be expressed only in the more aggressive cell line (MDA-MB-231), a pattern observed in more aggressive types of other cancers in our lab. ARHGAP25, previously reported in rhabdomyosarcoma, colorectal and lung cancer, has been found to be possibly expressed at higher levels in MDA-MB-231 compared to MCF-7. The variation in RBM20 subcellular localization in the breast cancer models studied and its significant presence in the cytoplasm raises the possibility of the two proteins interacting in direct protein-protein interaction in the cytoplasm, supporting the binding previously suggested in the phage display biopanning assay. Although AMPK was not found to affect the expression levels of RBM20, it seems to be involved in the regulation of its phosphorylation levels, both of which to be verified.

As for future endeavors to continue the investigation on this topic, we may start by comparing the RBM20 and ARHGAP25 expression levels and subcellular localizations in the non-transformed MCF-10A cell line to clarify the potential alterations observed in the MDA-MB-231 and MCF-7 cell lines. To further support our data on the subcellular localization of RBM20 and ARHGAP25 from the immunofluorescence staining, we will perform subcellular fractionation and look for the protein localizations and expression levels in the cytoplasm vs. the nucleus. This will also put the notion of ARHGAP25 and RBM20 interacting into more perspective based on their subcellular distribution relative to one another in each of the two cell lines. To verify the interaction between RBM20 and ARHGAP25, we will be performing the three complete independent repeats of Co-ImmunoPrecipitation, as described previously, and backing it up with Proximity Ligation Assays (PLA) for more solid results. In addition, drug inhibition of kinases in the PI3K/Akt/mTOR in both cell lines will allow us to experimentally assess the role AMPK plays in RBM20 phosphorylation and the effect of other kinases on the protein expression and activation of RBM20.

Finally, our results reinstate the importance of alternative splicing in tumorigenesis and raise the possibility of yet new proteins playing key roles in the different steps of initiation and development of cancer, specifically breast tumors. This study raises questions about the structures and functions taken on by RBM20 and ARHGAP25 in tumor cells, not just individually, but also together as a complex. Due to this new understanding, these proteins and the upstream signaling pathways that regulate their expression and activation could act as potential therapeutic targets in the future.

REFERENCES

Anczuków, O., & Krainer, A. R. (2016). Splicing-factor alterations in cancers. *Rna*, 22(9), 1285-1301.

Beqqali, A., Bollen, I. A., Rasmussen, T. B., van den Hoogenhof, M. M., van Deutekom, H. W., Schafer, S., ... & Mogensen, J. (2016). A mutation in the glutamate-rich region of RNA-binding motif protein 20 causes dilated cardiomyopathy through missplicing of titin and impaired Frank–Starling mechanism. *Cardiovascular research*, 112(1), 452-463.

Blech-Hermoni, Y., & Ladd, A. N. (2013). RNA binding proteins in the regulation of heart development. *The international journal of biochemistry & cell biology*, 45(11), 2467-2478.

Bonilla, J. M., Tabanera, M. T., & Mendoza, L. R. (2017). Breast cancer in the 21st century: from early detection to new therapies. *Radiología* (English Edition), 59(5), 368-379.

Brauch, K. M., Karst, M. L., Herron, K. J., De Andrade, M., Pellikka, P. A., Rodeheffer, R. J., ... & Olson, T. M. (2009). Mutations in ribonucleic acid binding protein gene cause familial dilated cardiomyopathy. *Journal of the American College of Cardiology*, 54(10), 930-941.

Campa, C. C., Ciraolo, E., Ghigo, A., Germena, G., & Hirsch, E. (2015). Crossroads of PI3K and Rac pathways. *Small GTPases*, 6(2), 71-80.

Capuco, A. V., & Ellis, S. E. (2013). Comparative aspects of mammary gland development and homeostasis. *Annu. Rev. Anim. Biosci.*, 1(1), 179-202.

Chai, N. N., Salido, E. C., & Yen, P. H. (1997). Multiple functional copies of theRBMgene family, a spermatogenesis candidate on the human Y chromosome. *Genomics*, 45(2), 355-361.

Chami, N., Tadros, R., Lemarbre, F., Lo, K. S., Beaudoin, M., Robb, L., ... & Lettre, G. (2014). Nonsense mutations in BAG3 are associated with early-onset dilated cardiomyopathy in French Canadians. *Canadian Journal of Cardiology*, 30(12), 1655-1661

Csépányi-Kömi, R., Sirokmány, G., Geiszt, M., & Ligeti, E. (2012). ARHGAP25, a novel Rac GTPase-activating protein, regulates phagocytosis in human neutrophilic granulocytes. *Blood*, 119(2), 573-582.

Csépányi-Kömi, R., Wisniewski, É., Bartos, B., Lévai, P., Németh, T., Balázs, B., ... & Ligeti, E. (2016). Rac GTPase activating protein ARHGAP25 regulates leukocyte transendothelial migration in mice. *The Journal of Immunology*, 197(7), 2807-2815.

Dai, M., Liu, Y., Nie, X., Zhang, J., Wang, Y., Ben, J., ... & Sang, A. (2015). Expression of RBMX in the light-induced damage of rat retina in vivo. *Cellular and Molecular Neurobiology*, 35(4), 463-471.

Dai, X., Xiang, L., Li, T., & Bai, Z. (2016). Cancer hallmarks, biomarkers and breast cancer molecular subtypes. *Journal of Cancer*, 7(10), 1281.

David, C. J., Chen, M., Assanah, M., Canoll, P., & Manley, J. L. (2010). HnRNP proteins controlled by c-Myc deregulate pyruvate kinase mRNA splicing in cancer. *Nature*, 463(7279), 364-368.

Desai, S., Ding, M., Wang, B., Lu, Z., Zhao, Q., Shaw, K., ... & Yao, J. (2014). Tissue-specific isoform switch and DNA hypomethylation of the pyruvate kinase PKM gene in human cancers. *Oncotarget*, 5(18), 8202.

Filippello, A., Lorenzi, P., Bergamo, E., & Romanelli, M. G. (2013). Identification of nuclear retention domains in the RBM20 protein. *FEBS letters*, 587(18), 2989-2995.

Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *cell*, 144(5), 646-674.

Hsiehchen, D., Goralski, M., Kim, J., Xie, Y., & Nijhawan, D. (2020). Biomarkers for RBM39 degradation in acute myeloid leukemia. *Leukemia*, 1-5.

Human ARHGAP25 GeneCard https://www.genecards.org/cgi-bin/carddisp.pl?gene=ARHGAP25

Human PKM GeneCard https://www.genecards.org/cgi-bin/carddisp.pl?gene=PKM

Human RBM20 GeneCard. https://www.genecards.org/cgi-bin/carddisp.pl?gene=RBM20

Kalimutho, M., Nones, K., Srihari, S., Duijf, P. H., Waddell, N., & Khanna, K. K. (2019). Patterns of genomic instability in breast cancer. *Trends in pharmacological sciences*.

Khan, M. A., Reckman, Y. J., Aufiero, S., van den Hoogenhof, M. M., van der Made, I., Beqqali, A., ... & Pinto, Y. M. (2016). RBM20 regulates circular RNA production from the titin gene. *Circulation research*, 119(9), 996-1003.

Kim, Y. J., & Abdel-Wahab, O. (2017, July). Therapeutic targeting of RNA splicing in myelodysplasia. *In Seminars in hematology* (Vol. 54, No. 3, pp. 167-173). WB Saunders.

Li, X. X., Shi, L., Zhou, X. J., Wu, J., Xia, T. S., Zhou, W. B., ... & Ding, Q. (2017). The role of c-Myc-RBM38 loop in the growth suppression in breast cancer. *Journal of Experimental & Clinical Cancer Research*, 36(1), 1-11.

Liang, R., Lin, Y., Ye, J. Z., Yan, X. X., Liu, Z. H., Li, Y. Q., ... & Ye, H. H. (2017). High expression of RBM8A predicts poor patient prognosis and promotes tumor progression in hepatocellular carcinoma. *Oncology Reports*, 37(4), 2167-2176.

Ling, N. X., Kaczmarek, A., Hoque, A., Davie, E., Ngoei, K. R., Morrison, K. R., ... & Dite, T. A. (2020). mTORC1 directly inhibits AMPK to promote cell proliferation under nutrient stress. *Nature Metabolism*, 2(1), 41-49.

Lőrincz, Á. M., Szarvas, G., Smith, S. M., & Ligeti, E. (2014). Role of Rac GTPase activating proteins in regulation of NADPH oxidase in human neutrophils. *Free Radical Biology and Medicine*, 68, 65-71.

Maragh, S., Miller, R. A., Bessling, S. L., Wang, G., Hook, P. W., & McCallion, A. S. (2014). Rbm24a and Rbm24b are required for normal somitogenesis. *PloS one*, 9(8), e105460.

Martínez-Arribas, F., Agudo, D., Pollán, M., Gómez-Esquer, F., Díaz-Gil, G., Lucas, R., & Schneider, J. (2006). Positive correlation between the expression of X-chromosome RBM genes (RBMX, RBM3, RBM10) and the proapoptotic Bax gene in human breast cancer. *Journal of cellular biochemistry*, 97(6), 1275-1282.

Martin-Garabato, E., Martinez-Arribas, F., Pollan, M., Lucas, A. R., Sanchez, J., & Schneider, J. (2008). The small variant of the apoptosis-associated X-chromosome RBM10 gene is co-expressed with caspase-3 in breast cancer. *Cancer Genomics-Proteomics*, 5(3-4), 169-173.

Masood, S. (2016). Breast cancer subtypes: morphologic and biologic characterization. *Women's Health*, 12(1), 103-119.

MCF-7 Cell Information https://www.mcf7.com/

MDA-MB-231 Cell Line Profile, European Collection of Authenticated Cell culture (ECACC)

Niimi, K., Usami, K., Fujita, Y., Abe, M., Furukawa, M., Suyama, Y., ... & Sato, S. (2018). Development of immune and microbial environments is independently regulated in the mammary gland. *Mucosal immunology*, 11(3), 643-653.

Noguchi, T., Inoue, H., & Tanaka, T. (1986). The M1-and M2-type isozymes of rat pyruvate kinase are produced from the same gene by alternative RNA splicing. *Journal of Biological Chemistry*, 261(29), 13807-13812.

Raghunath, A., Desai, K., & Ahluwalia, M. S. (2019). Current treatment options for breast cancer brain metastases. *Current treatment options in oncology*, 20(3), 1-18.

Refaat, M. M., Lubitz, S. A., Makino, S., Islam, Z., Frangiskakis, J. M., Mehdi, H., ... & Dudley, S. C. (2012). Genetic variation in the alternative splicing regulator RBM20 is associated with dilated cardiomyopathy. *Heart Rhythm*, 9(3), 390-396.

Schreiber, L., Lifschitz-Mercer, B., Paz, G., Yavetz, H., Elliott, D. J., Kula, K., ... & Maymon, B. B. S. (2003). Double immunolabeling by the RBM and the PLAP markers for identifying intratubular (in situ) germ cell neoplasia of the testis. *International journal of surgical pathology*, 11(1), 17-20.

Schreiber, L., Lifschitz-Mercer, B., Paz, G., Yavetz, H., Zadik, Z., Kula, K., ... & Maymon, B. B. S. (2003). Lack of RBM expression as a marker for carcinoma in situ of prepubertal dysgenetic testis. *Journal of andrology*, 24(1), 78-84.

Shemanko, C. S., Cong, Y., & Forsyth, A. (2016). What is breast in the bone? *International journal of molecular sciences*, 17(10), 1764.

Simple Module Architecture Research Tool http://smart.embl.de/smart/do_annotation.pl?DOMAIN=SM00233

Sun, X., Hu, Y., Wu, J., Shi, L., Zhu, L., Xi, P. W., ... & Ding, Q. (2018). RBMS2 inhibits the proliferation by stabilizing P21 mRNA in breast cancer. *Journal of Experimental & Clinical Cancer Research*, 37(1), 298.

Sutherland, L. C., Rintala-Maki, N. D., White, R. D., & Morin, C. D. (2005). RNA binding motif (RBM) proteins: a novel family of apoptosis modulators? *Journal of cellular biochemistry*, 94(1), 5-24.

Tao, L., Zhu, Y., Gu, Y., Zheng, J., & Yang, J. (2019). ARHGAP25: A negative regulator of colorectal cancer (CRC) metastasis via the Wnt/β-catenin pathway. *European journal of pharmacology*, 172476.

Thomas, R., & Majeti, R. (2019). No Matter How You Splice It, RBM39 Inhibition Targets Spliceosome Mutant AML. *Cancer cell*, 35(3), 337-339.

Thuault, S., Comunale, F., Hasna, J., Fortier, M., Planchon, D., Elarouci, N., ... & Gauthier-Rouvière, C. (2016). The RhoE/ROCK/ARHGAP25 signaling pathway controls cell invasion by inhibition of Rac activity. Molecular biology of the cell, 27(17), 2653-2661.Murayama, R., Kimura-Asami, M., Togo-Ohno, M., Yamasaki-Kato, Y., Naruse, T. K., Yamamoto, T., ... & Vatta, M. (2018). Phosphorylation of the RSRSP stretch is critical for splicing regulation by RNA-Binding Motif Protein 20 (RBM20) through nuclear localization. *Scientific reports*, 8(1), 1-14.

Thuault, S., Comunale, F., Hasna, J., Fortier, M., Planchon, D., Elarouci, N., ... & Gauthier-Rouvière, C. (2016). The RhoE/ROCK/ARHGAP25 signaling pathway controls cell invasion by inhibition of Rac activity. *Molecular biology of the cell*, 27(17), 2653-2661.

Ting, T. C., Goralski, M., Klein, K., Wang, B., Kim, J., Xie, Y., & Nijhawan, D. (2019). Aryl sulfonamides degrade RBM39 and RBM23 by recruitment to CRL4-DCAF15. *Cell reports*, 29(6), 1499-1510.

Upadhyay, S. K., & Mackereth, C. D. (2020). Structural basis of UCUU RNA motif recognition by splicing factor RBM20. *Nucleic acids research*, 48(8), 4538-4550.

Wampfler, J., Federzoni, E. A., Torbett, B. E., Fey, M. F., & Tschan, M. P. (2016). The RNA binding proteins RBM38 and DND1 are repressed in AML and have a novel function in APL differentiation. *Leukemia research*, 41, 96-102.

Wang, E., Lu, S. X., Pastore, A., Chen, X., Imig, J., Lee, S. C. W., ... & Ki, M. (2019). Targeting an RNA-binding protein network in acute myeloid leukemia. *Cancer Cell*, 35(3), 369-384.

Watanabe, T., Kimura, A., & Kuroyanagi, H. (2018). Alternative splicing regulator RBM20 and cardiomyopathy. *Frontiers in molecular biosciences*, 5.

Weeland, C. J., van den Hoogenhof, M. M., Beqqali, A., & Creemers, E. E. (2015). Insights into alternative splicing of sarcomeric genes in the heart. *Journal of molecular and cellular cardiology*, 81, 107-113.

Witkiewicz, A. K., McMillan, E. A., Balaji, U., Baek, G., Lin, W. C., Mansour, J., ... & Choti, M. A. (2015). Whole-exome sequencing of pancreatic cancer defines genetic diversity and therapeutic targets. *Nature communications*, 6(1), 1-11.

World Health Organization https://www.who.int/cancer/en/

Wu, J., Zhou, X. J., Sun, X., Xia, T. S., Li, X. X., Shi, L., ... & Ding, Q. (2017). RBM38 is involved in TGF-β-induced epithelial-to-mesenchymal transition by stabilising zonula occludens-1 mRNA in breast cancer. *British journal of cancer*, 117(5), 675-684.

Xu, K., Liu, B., & Ma, Y. (2019). The tumor suppressive roles of ARHGAP25 in lung cancer cells. *OncoTargets and therapy*, 12, 6699.

Xue, J. Q., Xia, T. S., Liang, X. Q., Zhou, W., Cheng, L., Shi, L., ... & Ding, Q. (2014). RNA-binding protein RNPC1: acting as a tumor suppressor in breast cancer. *BMC cancer*, 14(1), 322.

Yin, L. L., Wen, X. M., Li, M., Xu, Y. M., Zhao, X. F., Li, J., & Wang, X. W. (2018). A gene mutation in RNA-binding protein 10 is associated with lung adenocarcinoma progression and poor prognosis. *Oncology letters*, 16(5), 6283-6292.

Zhang, H. T., Zhang, Z. W., Xue, J. H., Kong, H. B., Liu, A. J., Li, S. C., ... & Xu, D. G. (2013). Differential expression of the RNA-binding motif protein 3 in human astrocytoma. *Chinese Medical Journal*, 126(10), 1948-1852.

Zhou, J., Chen, Q., Zou, Y., Zheng, S., & Chen, Y. (2019). Stem cells and cellular origins of mammary gland: updates in rationale, controversies, and cancer relevance. *Stem Cells International*, 2019.

Zhu, C., Yin, Z., Ren, J., McCormick, R. J., Ford, S. P., & Guo, W. (2015). RBM20 is an essential factor for thyroid hormone-regulated titin isoform transition. *Journal of molecular cell biology*, 7(1), 88-90.

Zhu, C., Yin, Z., Tan, B., & Guo, W. (2017). Insulin regulates titin pre-mRNA splicing through the PI3K-Akt-mTOR kinase axis in a RBM20-dependent manner. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease*, 1863(9), 2363-2371.