UC Riverside

UC Riverside Electronic Theses and Dissertations

Title

Protein Methylation and Quantifications of Nucleotide-Binding Proteins

Permalink

https://escholarship.org/uc/item/82c4z9m1

Author

Bade, David

Publication Date

2021

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Protein Methylation and Quantifications of Nucleotide-Binding Proteins

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Environmental Toxicology

by

David L. Bade

March 2021

Dissertation Committee:

Dr. Yinsheng Wang, Chairperson
Dr. David Volz
Dr. Xuan Liu

Copyright by
David L. Bade
2021

The Dissertation of David L. Bade is approved:

_____

_____

_____

Committee Chairperson

University of California, Riverside

Acknowledgements

The last five years have been among the most difficult yet rewarding of my life. Research can be trying at times, though I can say for certain I would not have made it through without the support of my principal investigator, Dr. Yinsheng Wang. He provided me with the opportunity to continue my education past my undergraduate degree, and supported me not just financially, but more importantly he taught me how to think critically about my research projects. When I was a new member of his lab, I found I just completed the research that was in front of me. Overtime however, Dr. Wang taught me that through careful study of publications, I could mold my project to make it more successful. He has my eternal gratitude for this.

I would also like to acknowledge Dr. David Volz and Dr. Xuan Liu for serving on my dissertation committee, as well as Dr. Jikui Song for donating his time as a member of my guidance committee. In addition, I would like to thank Dr. Li Fan and Dr. Xuemei Chen for their service on my qualifying committee. I would like to thank Dr. Joey Genereux, Dr. David Carter, Dr. Matthew Collins, and Clay Clark for suggestions in my research and facilities access.

I would like to thank Dr. David Eastmond, Dr. Jason Cheng, and Dr. Wenwan Zhong, as well as Antonio Knox and Dawn Loyola for their guidance and assistance in moving through the ETOX program. I would also like to acknowledge Dr. Kevin Simpson for his advice on improving my teaching, in particularly after my hip surgery. I would also like to thank Dr. Katie Burnette for not only her assistance in improving my teaching, but for also being my friend.

I would not have made it through my degree without the friendship and support of many lab members and friends at UCR. I would like to thank Dr. Lin Li and Dr. Xiaoxia Dai for instructing me in most of the research techniques that I employed in my research. I would also like to particularly thank Dr. Nathan Price, for his advice and assistance in not only improving my writing, but also for such things as his driving and companionship for picking up groceries at the store. I would like to thank Dr. Preston Williams, Dr. Weili Miao, Dr. Yuxiang Cui, and Dr. Ming Huang for their friendship and guidance on Mass Spectrometry instrumentation, analysis, and optimization. I would like to thank Gwendolyn Gonzales, Yen-Yu Yang, and Dr. Rong Cai for their tremendous help in experiments, but also for their support as friends. I would like to thank Dr. Xiaochuan Liu for sharing her knowledge in western blot and stable knockdown optimization. I would like to thank Dr. Kailin Yu for her assistance in creating knockout cell lines. I would like to thank Dr. Hua Du and Dr. Xiaomei He for sharing their experience with many types of experiments. Beyond this, I would like to thank the other current and former members of Wang Research Group, including Dr. Shuo Liu, Dr. Pengcheng Wang, Dr. Nicole Williams, Dr. Yang Yu, Dr. Nicholas Amato, Dr. Jun Wu, Dr. Yuxiang Sun, Dr. Quanqing Zhang, Dr. Jiabin Wu, Dr. Lok Ming Tam, Dr. Andrew Kellum, Xuejiao Dong, Ying Tan, Su Guo, Jiekai Yin, Jun Yuan, Zi Gao, Yinan Wang, Feng Tan, Garrit Clabaugh, Xin Wang, and Xingyuan Chen for their support and friendship.

I would like to thank the other members of the ETOX department, including Dr. Marissa Giroux, Dr. Sara Vliet, Dr. Linfeng Gao, and Dr. Zachary Cryder. I would also like to thank my friends from other departments, including Dr. Bobby Schafer, Craig

Douglass, Sheri Tao, Tim Gburek, Jedediah Kistner-Morris, Julian Morales and Kenny Nelson for their friendship. I would also like to thank my friends Jake Farley, Emily Thoreson, Elizabeth Costales, Paul Costales, and Daniel Runyon, among so many others for their constant support.

I would like to thank my parents, Jeffrey and Sylvia Bade, as well as by brother and sister-in-law Robert and Kristie Bade, for their constant and everlasting support. I would not be here if not for them.

Lastly, I would like to thank my fiancée, and one of the most supportive, caring, and intelligent people I have ever met, Tianyu Qi. Few things these past years have been stable. One thing that has been is her constant love and support these last years. I look forward to our life together.

Copyright Acknowledgements

The text and figures in Chapter 2, in part or full, are a reprint of the material as it appears in Biochemical and Biophysical Research Communications 2021 Feb 6;546:54-58. The coauthor (Dr. Yinsheng Wang) listed in that publication directed and supervised the research that forms the basis of this chapter.

Dedication

To my parents Jeffrey and Sylvia,

who supported me in every endeavor,

with unconditional love and support.

To my fiancée Tianyu Qi,

who inspires me every day to be a better person.

ABSTRACT OF THE DISSERTATION

Protein Methylation and Quantifications of Nucleotide-Binding Proteins

by

David L. Bade

Doctor of Philosophy, Graduate Program in Environmental Toxicology
University of California, Riverside, March 2021
Dr. Yinsheng Wang, Chairperson

Post-translational modifications (PTMs) constitute an evolutionally conserved and ubiquitous mechanism for regulating protein structure and function. The first part of this dissertation focuses on the discovery of an epitranscriptomic mechanism of regulation for N-terminal methyltransferase 1 (NTMT1) expression, and the identification of novel substrates of α-N-methylation and histidine methylation. The goal of the second half of this dissertation is to discover nucleotide-binding proteins (i.e. small GTPases and kinases) as drivers and suppressors for cancer metastasis and acquired radioresistance.

In Chapter Two, we discovered that NTMT1 expression is regulated by reader/writer/eraser proteins of N^6-methyladenosine (m^6A). We further identified Mortality Factor 4 Like 1 (MRG15) as a novel α-N-methylated protein. We subsequently demonstrated that NTMT1 is responsible for α-N-methylation of MRG15, and this methylation is also regulated through an m^6A-based epitranscriptomic mechanism.

In Chapter Three, we interrogated a series of publicly available mass spectrometry datasets with MaxQuant to identify novel protein substrates for α-N-methylation. Our

results uncovered 219 instances of N-terminal methylation in 196 proteins. We then identified, through affinity purification and mass spectrometry analysis, that Vesicle-associated membrane protein 4 (VAMP4) is α-N-methylated. We subsequently confirmed that NTMT1 is the primary enzyme responsible for the N-terminal methylation of VAMP4.

In Chapter Four, we analyzed publicly available mass spectrometry datasets to identify novel protein substrates for histidine methylation. Using this method, we uncovered 33 instances of histidine methylation among 26 proteins. We subsequently used affinity purification and mass spectrometry analysis to confirm the histidine methylation of Pre-mRNA-splicing factor RBM22 (RBM22).

In Chapter Five, we employed stable isotopic labeling by amino acids in cell culture (SILAC) and parallel-reaction monitoring (PRM) to monitor the differential expression of kinases in MCF-7 and MDA-MB-231 breast cancer cells and their corresponding radioresistant C6 and C5 clones. Using this method, we were able to identify and quantify the relative expression levels of 300 and 281 kinases in C5/MDA-MB-231 and C6/MCF-7 pairs of breast cancer cells, respectively. We further identified transcription initiation factor TFIID subunit 9 (TAF9) as a driver of radioresistance in breast cancer cell lines.

In Chapter Six, we employed a proteomic method, based on multiple-reaction monitoring (MRM), for quantitative profiling of GTP-binding proteins using isotope-coded GTP probes. After probe labeling, tryptic digestion, and affinity enrichment of labeled peptides, the expression level differences of GTP-binding proteins in two

matched pairs of primary/metastatic melanoma cell lines (WM-115/WM-266-4 and IGR-39/IGR-37) were investigated with LC-MRM analysis. Among the most upregulated proteins, adenylate kinase 4 (AK4) was identified as a potential driver of melanoma metastasis. We further demonstrated that AK4 expression is necessary for melanoma cell invasion and migration.

Table of Contents

List of Tables

List of Figures

Chapter 1: Overview

1 Introduction

Among the most significant scientific accomplishments in the late 1990s was the sequencing of the human genome (1). This milestone heralded the era of epigenetics, where we learned that modifications other than just the canonical DNA sequence can affect gene expression (2). Given the flow of genetic information, wherein DNA is transcribed into mRNA, which is then translated into protein; it is not surprising that these modifications can occur at many positions within the cell, are written dynamically, and affect genetic events that occur downstream (3). As we know today, epigenetic modifications can occur through several mechanisms, including DNA methylations, RNA post-transcriptional modifications, and post-translation modifications (PTMs) on histone proteins (2). The discovery of PTMs contributes to the much larger field of proteomic studies.

One important method for studying the proteome is to compare and then differentiate the amount of specific proteins between two samples, for which 'Western blot' was developed (4). Western blotting techniques were expanded for identifying peptides with a specific PTM by using antibodies designed to bind with the modified peptide (5,6). Although Western blotting allows for identifying and quantifying specific proteins and peptides, it is time intensive, prohibiting its use for high-throughput studies to quantify multiple proteins at the same time.

The disadvantage of Western blotting has led to the application of many techniques, predominantly utilizing LC-MS/MS analysis, for comparative proteomic studies. The high-throughput capabilities of mass spectrometry have allowed for identifying protein biomarkers of many diseases, including cancer (7,8). Additionally, the high sensitivity and selectivity of LC-MS/MS permits the identification of PTMs (9). In this chapter, I will first introduce common mass spectrometric methods for proteomic analysis, including discovery-based and targeted proteomics. I will then discuss two types of protein PTMs, i.e. α-N-methylation and histidine methylation. Finally, I will go over the use of mass spectrometry for quantitative proteomics to compare the expression of kinases and small GTPases, and discuss the screening of these proteins to discover drivers of cancer metastasis and the roles of these nucleotide-binding proteins in acquired radioresistance.

1.1 Mass Spectrometry-Based Characterization of Protein PTMs

After translation, proteins are subjected to many types of PTMs, which can modulate their structures and functions. Until recently, the study of protein PTMs has consisted primarily of immunochemistry and Edman degradation (10). While these methods can target single PTMs, they are inefficient for scanning multiple target locations at the same time and are susceptible to significant signal interference from untargeted proteins. In the past few decades, mass spectrometry has emerged as a powerful tool for identifying PTMs and it overcomes the limitations of these conventional biochemical methods.

1.1.1 Bottom-up Proteomics for PTM Identification

Bottom-up proteomics is a standard method of using mass spectrometry for qualitative and quantitative proteomics. In this approach, proteins from whole-cell lysate are digested with endopeptidase(s) to produce short peptides that are subsequently separated by liquid chromatography and analyzed with a mass spectrometer (11). While many mass spectrometric methods focus on the study of the proteome, highly abundant peptides can suppress the signal of those peptides of low abundance. This is further complicated for experiments designed to identify peptides with specific PTMs. Thus, the formulation of methods designed to enrich peptides of interest from the mixture of the entire proteome has been an ongoing research area for many laboratories. Several decades ago, Hopp et al. (12) created a hydrophilic sequence of eight amino acids (FLAG tag), which is tagged to either the N or C-terminus of a protein of interest. This tag can then be bound to an antibody targeting the FLAG tag, allowing for the subsequent removal of non-tagged proteins, thus leading to the enrichment of the protein of interest (Figure 1.1). After this, the cysteines on the protein are reduced and alkylated through the use of dithiothreitol and iodoacetamide, respectively. The protein is subsequently digested with endopeptidases, to produce smaller peptide fragments capable of being analyzed on a mass spectrometer. The most common endopeptidase used for this purpose is trypsin, primarily due to its stability and robustness. Trypsin is capable of cleaving amide bonds on the C-terminal side of lysine and less commonly arginine residues. However, due to the randomness that these residues appear on proteins, the use of trypsin alone cannot produce a full map of proteins in the proteome. Many other endopeptidases

are also used to varying degrees, including Glu-C which is capable of cleaving amide bonds C-terminal to a glutamic acid. Given the somewhat random nature by which different residues appear in a peptide, the use of one endopeptidase may produce peptides that are too short or too long for mass spectrometric analysis. Thus, by employing several different endopeptidases, sometimes sequentially, to produce short enough peptides, one can optimize protein sequence coverage.

1.2 PTMS

Protein PTMs constitute a ubiquitous and evolutionarily conserved mechanism of gene regulation. As the name implies, PTMs occur after the protein has been translated from mRNA. Up to now over 200 distinct types of PTMs have been discovered, existing across many organisms (13,14). Notably, only a few types of PTMs make up the majority of the actual PTMs found within proteins, and these modifications include phosphorylation, ubiquitination, acetylation, glycosylation, and methylation (15).

Protein methylation is the transfer of one or more methyl groups to an amino acid. The methyl groups are donated by S-adenosyl-L-methionine (SAM) (16). Several decades ago, ε-N-methylation of lysine on the flagellar protein of *Salmonella typhimurium* was discovered (17). A few years later, this modification was also identified on histones (18). The methylation was later to found to also occur on arginine, cysteine, histidine, and at the α-N-terminus of a protein (15). Methylations of lysine and arginine are among the most studied, as the current data suggest that this is significantly more common than other types (19). The biological functions of α-N-methylation and histidine methylation, however, remain under-investigated.

1.2.1 α-N-Methylation

Protein α-N-Methylation was first discovered in *Escherichia coli* in the 1970s, on ribosomal subunits L16 and L33 (20,21). Since these early discoveries, α-N-Methylation has also been identified in proteins from several other organisms, including histone H2B from *Asterias rubens* as well as ribosomal protein L11 and chemotaxi-flagellar apparatus in *E. coli* (22–25).

A decade ago, α-N-terminal protein methyltransferase 1 (TAE1) was discovered in *Saccharomyces cerevisiae*, and it is able to α-N-methylate 40S ribosomal protein S25-A (RPS25A) and 40S ribosomal protein S25-B (RPS25B) (26). These targets all possessed a conserved ProProLys motif. Investigations into this motif led to the discovery of human homolog of TAE1, N-terminal methyltransferase 1 (NTMT1). NTMT1 can methylate several proteins including regulator of chromosome condensation (RCC1), SET translocation (SET), retinoblastoma (Rb), ribosomal protein L23a (RPL23A), myosin light polypeptide 2 (MYL2), and myosin light polypeptide 3 (MYL3) (27). Since the discovery of NTMT1, several other protein substrates have also been identified, including centromere protein A (CENP-A), centromere protein B (CENP-B), damaged DNA binding protein 2 (DDB2), and Obg Like ATPase 1 (OLA1) (28–31). A comparative profiling of these substrates shows a binding motif of XPK/R, wherein ‘X’ represents alanine, proline, serine, or glycine. Since then, further studies into the substrate preference of NTMT1 have led to a significant expansion of the XPK motif (‘X’ can be any residue type except Asp and Glu) (32,33). These discoveries further increased the pool of potential α-N-methylated proteins, suggesting most protein substrates for NTMT1

have yet to be discovered. N-terminal methyltransferase 2 (NTMT2) was later discovered to monomethylate several known targets of NTMT1 (34). Further research demonstrated that NTMT2 could form a heterotrimer with NTMT1, priming its function (Figure 1.6) (34–36).

METTL13 was the second N-terminal methyltransferase identified in humans. This enzyme is unique because it possesses two domains, each capable of methylating protein substrates in different ways. The C-terminal domain was discovered to methylate EEF1A1 at K55, whereas the N-terminal domain of METTL13 can methylate EEF1A1 at the N-terminus (37,38). Peptide screening later demonstrated that EEF1A1 acts upon an N-terminal motif that is quite distinct from that of NTMT1/2, increasing even further the pool of potential α-N-methylated proteins (37).

1.2.2 Functional Importance of α-N-methylation

Considerable evidence has been found to suggest that α-N-methylation is involved in protein binding. For CENP-B, α-N-methylation can enhance its ability to bind with the CENP-B Box and satellite DNA (29). In the same vein, loss of α-N-methylation in DDB2 was discovered to depress its ability to localize to the nucleus and bind with UV damaged DNA (30). α-N-methylation of RCC1 is necessary for proper localization of the protein to chromatin during mitosis (39). N-terminal methylated RCC1 is also necessary for the switch from GDP-bound RAN to GTP-bound RAN. RAN is the primary protein responsible for transporting many proteins and RNA to and out of the nucleus, and RAN's binding to GTP and GDP is vital for its movement across the nuclear membrane (40). Through this α-N-methylation, RCC1 can regulate nucleocytoplasmic transport (41)

Additionally, because GTP-bound RAN production is necessary for proper mitotic spindle assembly and function, α-N-methylation also plays an important role (39,42).

Aberrant NTMT1 expression has been discovered in several types of cancer (43). In breast cancer, downregulation was shown to promote breast cancer cells' growth and metastasis (44). In colorectal cancer cells, NTMT1 is upregulated and may possess a role as an oncogene (45). In mice, loss of NTMT1 expression was found to impair DNA repair and cause premature aging (46). METTL13 is upregulated in pancreatic and lung cancer, whereby N-terminally methylated EEF1A1 increases translational output and tumorigenesis (38). METTL13 has also been demonstrated to drive tumor growth in hepatocellular carcinoma (47,48). Together, these results show a significant role for α-N-methylation in regulating cancer progression and prognosis.

1.2.3 Histidine Methylation

Among the different types of protein methylation, histidine methylation remains one of the least studied. This is surprising, as the first discovery of histidine methylation occurred nearly 100 years ago when β-alanyl-Nπ-methyl-L-histidine (Anserine) was isolated from goose eggs (49). In 1951, 1-methylhistidine was isolated from cat urine and then from human urine as 3-methylhistidine three years later (50,51). Actin soon became the first protein known to have histidine methylation, which was later expanded to include myosin (52,53). Both of these methylated histidine proteins were confirmed to exist across several species of vertebrates and invertebrates, suggesting that they are evolutionarily conserved (52). To date, there have been several other proteins found to harbor methylated

histidine, including histones, myosin light chain kinase 2 (MLCK2), S100A9, methyl-coenzyme m reductase (MCR), and large subunit ribosomal protein 3 (RPL3) (54–57).

Actin-histidine N-methyltransferase (SETD3) is the first protein currently known as a histidine methyltransferase, methylating histidine residues on actin (58,59). Recent studies in *S. cerevisiae* identified histidine protein methyltransferase 1 (HPM1) as a histidine methyltransferase, suggesting its human ortholog Methyltransferase Like 18 (METTL18) may also be a histidine methyltransferase (55). HPM1 has been shown to methylate histidine on RPL3 in yeast (55). Recent research has also identified Methyltransferase Like 9 (METTL9) as a writer of 1-methyl-histidine (1MH) (60). Subsequent experiments demonstrated that METTL9 is capable of methylating histidine on several proteins, including S100A9 and DnaJ Heat Shock Protein Family (Hsp40) Member B12 (DNAJB12) (60).

1.2.4 Functional Importance of Histidine Methylation

In yeast, actin histidine methylation at H73 is associated with an increased rate of ATP cleavage and affects actin polymerization. This suggests that histidine methylation on actin may affect cell motility (61). Eom et al. (62) demonstrated that SETD3 targets gene promoters and methylates histone H3K4 and H3K36, promoting the transcription of *MYOG* gene. This is important because there is significant evidence that SETD3 can perturb biological processes. In one instance, knockdown of SETD3 in C2C12 cells was found to suppress skeletal muscle differentiation (63). Moreover, hSETD3 has been identified as a PCNA-binding protein, suggesting a role in DNA replication/repair (64).

SETD3 has also been shown to be involved in many types of cancer, including promoting radiosensitivity in cervical cancer cells by suppressing the expression of kinesin light chain 4 (KLC4) (65). SETD3 can also promote DNA damage-associated apoptosis in colon cancer cells (66). SETD3 has also been discovered as a potential biomarker of breast cancer prognosis (67). Conversely, SETD3 has been shown to regulate hepatocellular carcinoma by promoting cell proliferation while inhibiting metastasis (68). Together, these results revealed the crucial roles histidine methylation has on regulating biological processes and cancer prognosis, despite the limited research conducted into histidine methylation.

1.3 Epitranscriptomic Regulation

With well over 170 known types of covalent modifications on many types of RNA, including messenger RNA (mRNA), ribosomal RNA (rRNA), and transfer RNA (tRNA), post-transcriptional modifications are a ubiquitous and evolutionarily conserved mechanism for controlling gene expression (69–71). Together, these RNA modifications are members of the epitranscriptome. One of the most common types of mRNA modifications, N^6-methyladenosine (m^6A), was first discovered in 1974 in mouse L cells (72). In 1978, researchers discovered that HeLa cells modified with m^6A showed instability over time (73). Further discoveries linked methyltransferase like 3 (METTL3) as the primary enzyme responsible for methylating adenosine to m^6A (74). It was soon discovered that METTL3 accomplishes this primarily by forming a complex with METTL14 and Pre-mRNA-splicing regulator WTAP (WTAP) (Figure 1.3) (75,76).

FTO was the first m^6A demethylase discovered, and it also exhibits limited demethylation activity towards 3-methyluridine (m^3U) (77,78). Since these discoveries, evidence has shown that FTO more specifically demethylates N^6,2′-*O*-dimethyladenosine (m^6A_m), exclusively found in cap regions of mRNAs (79). ALKBH5 was the second m^6A demethylase discovered (Figure 1.3), and it was also shown to have no demethylase activity toward m^6A_m (80).

Earlier studies suggested that m^6A has a role in controlling the stability of mRNAs before they are translated into proteins (81). Research has also suggested that YTH domain family protein (YTHDF1) and YTHDF2 bind to m^6A to stabilize and destabilize the mRNA, respectively (81). YTHDF3 can promote mRNA stability by recognizing m^6A or assist YTHDF2 in destabilizing mRNA (Figure 1.3) (82). Further research, however, suggested that YTHDF1/2/3 may all act redundantly to destabilize mRNA and promote cellular differentiation (83).

1.3.1 The Role of Epitranscriptomics in Cancer

Aberrant expression of genes encoding m^6A writer/reader/eraser proteins has been associated with various types of cancer. METTL3 plays roles in promoting radioresistance in glioblastoma stem cells through methylation of the mRNA of SRY-box transcription factor 2 (SOX2), promoting stability and increasing the expression of SOX2 (84). SOX2 has been implicated in promoting radioresistance in glioblastoma stem cells through enhancing DNA repair (84). m^6A was also suppressed in approximately 70% of endometrial tumors, leading to proliferation and tumorigenesis (85,86). Pancreatic cancer is one of the deadliest types of cancer, leading to a fatality rate of near 100% (85,86).

Recent research showed that depletion of METTL3 sensitized pancreatic cancer cells to gemcitabine, 5-fluorouracil, and cisplatin (85).

1.4 Mass Spectrometry-based Quantitative Proteomics

Mass spectrometry has been adapted for use in quantitative proteomics experiments, due to the benefits of high throughput and sensitive analysis. The two main types are discovery proteomics and targeted proteomics.

1.4.1 Discovery Proteomics

Discovery proteomics, also referred to as "shotgun" proteomics, is a standard method of bottom-up proteomics (11). There are two main types of shogun proteomics, data-dependent acquisition (DDA) and data-independent acquisition (DIA). In the DDA mode, mass spectrometry typically performs a MS1 survey scan followed by the fragmentation of the top N most abundant ions from the survey scan where MS/MS data are acquired (Figure 1.4). In the DIA mode, rather than focusing on the most abundant ions, parent ions are isolated sequentially and fragmented in specific time windows (87). There are unique benefits of either method, where DDA is much faster, as opposed to a more thorough fragmentation and scanning of lower abundance ions in DIA (88).

Discovery-based proteomics does not require any specific knowledge of the parent ions in a sample for analysis. This means that the data can be reanalyzed after data collection at later time point for different research projects, for the purpose of identifying novel peptides and PTMs, for instance. This has led to the discovery of over 10,000 protein groups and 30,000 phosphorylation sites from humans and mice (89). The possibility of further identifying unique PTM sites has led to the development of online file-sharing

databases to share publicly available datasets. MassIVE repository (MassIVE) and the Proteomics Identification Database (PRIDE) are examples of such online databases for the sharing of datasets (90,91). These datasets can then be mined for the identification of novel peptides with PTMs. One of the most common tools for this purpose is MaxQuant quantitative software (92). Using MaxQuant, individual MS/MS scans can be compared against the theoretical fragmentation patterns of all tryptic digested peptides from all proteins in an organism, detailed in a FASTA file. Individual parameters in MaxQuant can be altered depending on what PTMs of interest to identify, providing value in reexamining discovery datasets for different research projects.

Discovery proteomics can be further coupled with the use of desthiobiotin-linked probes for the targeted enrichment of specific subfamilies of proteins (93). For instance, one can link GTP to desthiobiotin, wherein desthiobiotin can be bound and enrichment with streptavidin beads from a sample, allowing for the subsequent enrichment of GTP-binding proteins (94). With this type of technique, you can increase the ease with which GTP-binding protein can be observed using discovery-based proteomic experiments by removing those peptides that one is not attempting to study.

1.4.2 Targeted Proteomics

Because mass spectrometers collect MS/MS on precursor ions based on their relative abundances in discovery proteomics, this method can be ineffective in monitoring specific peptides of interest. This has led to the growing popularity of targeted proteomics in recent years. There are two main types of targeted proteomics, multiple-reaction monitoring (MRM) and parallel-reaction monitoring (PRM) (95,96). The most common

form of targeted proteomics method, MRM, sometimes also referred to as selective ion monitoring (SRM), is typically conducted on a triple-quadrupole mass spectrometer, where the first and third quadrupoles serve as mass filters, and the second quadrupole works as a collision cell. When a mass spectrometer is set to run in the MRM scanning mode, the sample is typically scanned to identify predefined particular precursor-product ion pairs based on these ions' expected m/z (Figure 1.4) (97). Typically three precursor-product ion pairs are monitored per peptide, where each pair is often predetermined through the analysis of datasets acquired from discovery proteomics. However, there may be too many peptides to monitor within a given sample so the method can be refined by scheduling specific scans only at specific retention time (RT) windows.

To do this, one can calculate the normalized RT (iRT) value for each peptide on the target list (98). The iRT value is an empirically determined indicator of retention time. One can compare the iRT of peptides of interest to the iRT value of reference peptides. For this purpose, one can use tryptic peptides of bovine serum albumin. The iRT values of the earliest and latest reference peptides are assigned as 0 and 100, respectively, whilst those of other peptides assigned based on when they elute in the full-MS scan using a linear regression curve. Prior to running samples, the reference peptides are run to determine at which point during the full-MS scan they elute. This information can then be compared to the other peptides being targeted to assess when they in turn will appear during the scan. Using scheduled RT windows, the number of ions being scanned at any particular time is reduced. The use of MRM analysis for quantitative proteomics is superior to DDA method

for its ability to scan particular peptides of interest reproducibly and with high sensitivity (99).

PRM is a relatively new method in targeted proteomics, where a precursor ion is isolated with a quadrupole mass filter, fragmented through collisional activation, then its product ions analyzed with a high-resolution Orbitrap or time-of-flight mass analyzer (Figure 1.4) (100). A major benefit of PRM over MRM is the ability to scan all potential product ions from the fragmentation of the precursor ion instead of only a few predefined one (101). The Orbitrap analyzer also allows for high-resolution mass measurement of the product ions, contributing to the confidence in peptide identification (102). PRM can also be improved through scheduled RT windows, allowing for much more peptide precursor ions to be identified in a single LC run. Results from these runs can further be analyzed through the use of Skyline software, which provides an interface for visualizing the MS^2 signal (103). Together, PRM provides highly accurate, reproducible, and reliable peptide quantification.

1.4.3 SILAC and quantitative proteomics

Stable isotope labeling by amino acids in cell culture (SILAC) is a standard method of differentiating between two different cell lysates in a sample. Classical SILAC involves labeling particular amino acids in cell culture with either light or heavy versions of an amino acid (Figure 1.5) (104). As trypsin is the most common endopeptidase employed for sample preparation, light and heavy variants of lysine and arginine are often preferred. This ensures that the light or heavy variant amino acids occur in all peptide except the C-terminal end peptides of some proteins produced through proteolytic digestion. To begin

this process, culture media is prepared wherein heavy isotope-labelled amino acids, such as ^{13}C- and ^{15}N-labelled lysine and arginine, are added in lieu of their unlabeled counterparts. The cells are then cultured in the SILAC medium over a minimum of five cell doubling times to ensure ample time for complete labeling of the proteome. After cell lysis, the two labeled samples are mixed at a 1:1 ratio by mass. Following reduction, alkylation, and digestion, the samples are then subjected to LC-MS and MS/MS analysis. From the acquired MS spectra, each labeled precursor ion of a peptide appears as a two sets of peaks with a fixed *m/z* shift depended on the difference in *m/z* of the light and heavy amino acids. The signal of peaks can then be quantified, and a ratio can be obtained between the two cell populations.

1.5 Human Kinome

Kinases are among the most important types of enzymes that are involved in a number of biological processes, including cellular proliferation and migration (Figure 1.6) (105). The primary function of kinome is the phosphorylation of small molecules and other proteins through the use of adenosine triphosphate (ATP) as the phosphate donor, affecting processes such as signal transduction and cellular homeostasis (93). Aberrant kinase expression is a hallmark of many types of cancer, particularly through modulating the effectiveness of chemotherapy and radiation treatment. Breast cancer is the most common type of cancer diagnosed in women (106). There are several kinases that have been associated with radioresistance in breast cancer cells. Maternal embryonic leucine zipper kinase (MELK) is a serine/threonine kinase that is known to be enriched in triple-negative breast cancer (TNBC), a particular type of cancer resistant to radiation treatment. MELK

is upregulated in radioresistance breast cancer and is believed to be associated with the repair of DNA double strand breaks (DSBs) (107). Ionizing radiation is effective against cancer cells commonly through inducing this type of damage, so the acquired ability of cancerous cells to counter DNA damage could allow for increased resistance to such a treatment. Ataxia-telangiectasia mutated (ATM) is another protein that has been linked to radioresistance in breast cancer cell lines (108). ATM is a sensor for the detection of DSBs, which is important for initiation of repair mechanisms (109). Increase in the expression of this gene supports the repair of cancer cells treated with ionizing radiation (108).

1.6 GTP-binding Proteins

Guanosine triphosphate (GTP) is commonly associated with providing energy for protein synthesis and signal transduction cascade in cells (109). Aberrant expression of GTP-binding proteins is associated with promoting and suppressing cellular metastasis, invasion, and migration in several different types of cancer. Ras-Related Protein Rab-27A (RAB27A) is overexpressed in several advanced cancers. In pancreatic cancer, RAB27A expression can promote metastasis at distant organs while suppressing the invasive capability of cancer cells (110). RAB27A has also been discovered to promote invasion and migration of breast cancer cells through promoting the secretion of insulin-like growth factor-II (IGF-II), which regulates the expression of several proteins, prominently including matrix metallopeptidase 9 (MMP9), an important enzyme in the degradation of the extracellular matrix (ECM) (111). Ras Homolog Family Member C (RhoC) has also been identified as a driver of melanoma metastasis, where expression of RHOC promotes

invasion and metastasis. This phenotype can also be reversed through downregulation of RhoC (112).

The search for novel drivers of metastasis in cancer cells can be applied for the treatment of patients suffering from these maladies with inhibitors. Rac Family Small GTPase 1 (RAC) and Cell Division Cycle 42 (CDC42) are two GTPases that can promote epithelial-to-mesenchymal transition and tumor metastasis in several types of cancer. Several inhibitors have since been developed for the targeting of these GTP-binding proteins(113).

1.7 Scope of the Dissertation

The research covered in this dissertation focuses on two main lines of work. First is the discovery of the epitranscriptomic regulation of NTMT1 expression through the dynamic addition of m^6A. Next, I sought to identify novel instances of α-N-methylation and histidine methylation in proteins. This study aimed to improve our understanding of the underlying mechanisms and functions behind two poorly understood forms of PTMs that have been found to affect many cellular processes and diseases. The second line of research focuses on discovering novel protein drivers of metastasis in melanoma cell lines and acquired radioresistance in breast cancer cells.

In Chapter two, I sought to determine how α-N-methylation is regulated in cells. I discovered that NTMT1 is dynamically regulated through the epitranscriptome by the knockout of the m^6A writer protein METTL3, the reader proteins YTHDF1/2/3, and the eraser protein FTO. Using LC-MS/MS, I subsequently identified that, in HEK293T cells, following the removal of the initial methionine residue, MRG15 can be mono-, di-, and tri-

methylated at the α-N-terminus. Given that MRG15 possesses the conserved XPK motif, which is common among targets of NTMT1, I subsequently sought to determine if NTMT1 is the enzyme responsible for this MRG15 α-N-methylation. Previous studies have demonstrated the importance of the 4th lysine of α-N-methylated proteins exhibiting the XPK motif in allowing NTMT1 to bind (31,32,44). Consistent with previous findings, upon mutation of the 4th residue of MRG15 to glutamine, we observed a complete loss of α-N-methylation. We also observed the loss of N-terminal methylation of MRG15 in cells treated with siRNA targeting NTMT1, further confirming that NTMT1 is the primary enzyme responsible for MRG15 α-N-methylation. I then sought to determine whether the N-terminal methylation of MRG15 is regulated through the epitranscriptomic method previously identified. Through LC-MS/MS analysis and the enrichment of ectopically expression MRG15 in HEK293T cells and isogenic cells lacking expression of METTL3, YTHDF1/2/3, and FTO, I found that MRG15 α-N-methylation is regulated consistent with our previous data. Together, our study demonstrated the first known epitranscriptomic mechanism for regulating α-N-methylation and its effect on the novel α-N-methylated gene, MRG15.

In Chapter three, I sought to discover novel protein substrates exhibiting α-N-methylation. For this project, I reanalyzed publicly available discovery proteomics files with MaxQuant to identify N-terminal peptides that are mono-, di-, or tri-methylated. Using this method, I identified 219 unique instances of N-terminal methylation on 196 individual proteins. Among these proteins, VAMP4 is a novel protein exhibiting the XPK motif that NTMT1 is known to target. Through LC-MS/MS analysis and the enrichment of

ectopically expression VAMP4 in HEK293T cells, I identified the mono- and di-methylation of the α-N-terminus of VAMP4. I next sought to determine whether NTMT1 is the enzyme response of methylating VAMP4. Through the mutation of the 4th lysine to glutamine, I observed a significant loss of α-N-methylation. The data were further substantiated by expressing VAMP4 in isogenic cells lacking expression of NTMT1, revealing a nearly complete loss of α-N-methylation of VAMP4. Together, our data expanded significantly the current pool of α-N-methylated genes and identified VAMP4 as a novel substrate of NTMT1.

In Chapter four, I reanalyzed publicly available datasets to identify novel histidine methylated proteins. I subsequently identified 33 instances of histidine methylation across 26 individual proteins. I further sought to confirm the histidine methylation of RBM22 via LC-MS/MS analysis. Following the ectopic expression of RBM22, I observed the presence of monomethylation of histidine H183 in RBM22. Collectively I identified many more potential histidine methylated proteins, and in particular, RBM22.

In Chapter five, we employed a PRM method for analyzing the expression of >75% of the human kinome. We employed this method to measure the difference in kinase expression between parental breast cancer cell lines (MDA-MB-231 and MCF-7) and their radioresistant counterparts (C5 and C6). We quantified the relative protein expression levels of 300 and 281 kinases in C5/MDA-MB-231 and C6/MCF7 paired cells, respectively. I further identified TAF9 as a driver of acquired radioresistance in these cancer lines through clonogenic survival assays.

In Chapter six, we sought to measure the differential expression of GTPase binding proteins between primary melanoma cell lines (WM115 and IGR39) and their metastatic counterparts (WM266-4 and IGR37). From these data, we identified adenylate kinase 4 (AK4) as a GTPase showing a significant increase in expression in metastatic cell lines. To demonstrate the effect of AK4 on melanoma metastasis, we constructed several stable knockdown cell lines of AK4 in WM266-4 and IGR37 cells and stable overexpression lines of AK4 in WM115 and IGR39 cells. When comparing these stable lines to cells expressing either a control shRNA or vector sequence, I observed through migration/invasion assays that AK4 expression consistently promotes the migration and invasion of melanoma cells. I further confirmed this increase in migration through the use of wound healing scratch assays, wherein cells with greater AK4 expression migrated faster than those melanoma cells with lower AK4 expression. Collectively, these results identify AK4 as a novel driver for melanoma metastasis.

References

1. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, et al. Initial sequencing and analysis of the human genome. Nature. 2001 Feb 15;409(6822):860–921.
2. Allis CD, Jenuwein T. The molecular hallmarks of epigenetic control. Nat Rev Genet. 2016 Jun 27;17(8):487–500.
3. Kenyon GL, DeMarini DM, Fuchs E, Galas DJ, Kirsch JF, Leyh TS, et al. Defining the mandate of proteomics in the post-genomics era: workshop report. Mol Cell Proteomics. 2002 Oct;1(10):763–780.
4. Towbin H, Staehelin T, Gordon J. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc Natl Acad Sci USA. 1979 Sep;76(9):4350–4354.
5. Stoevesandt O, Taussig MJ. Phospho-specific antibodies by design. Nat Biotechnol. 2013 Oct;31(10):889–891.
6. Fuchs SM, Strahl BD. Antibody recognition of histone post-translational modifications: emerging issues and future prospects. Epigenomics. 2011 Jun;3(3):247–249.
7. Miao W, Fan M, Huang M, Li JJ, Wang Y. Targeted profiling of heat shock proteome in radioresistant breast cancer cells. Chem Res Toxicol. 2019 Feb 18;32(2):326–332.
8. Miao W, Wang Y. Targeted quantitative kinome analysis identifies PRPS2 as a promoter for colorectal cancer metastasis. J Proteome Res. 2019 May 3;18(5):2279–2286.
9. Villén J, Gygi SP. The SCX/IMAC enrichment approach for global phosphorylation analysis by mass spectrometry. Nat Protoc. 2008;3(10):1630–1638.
10. Larsen MR, Trelle MB, Thingholm TE, Jensen ON. Analysis of posttranslational modifications of proteins by tandem mass spectrometry. BioTechniques. 2006 Jun;40(6):790–798.
11. Zhang Y, Fonslow BR, Shan B, Baek M-C, Yates JR. Protein analysis by shotgun/bottom-up proteomics. Chem Rev. 2013 Apr 10;113(4):2343–2394.
12. Hopp TP, Prickett KS, Price VL, Libby RT, March CJ, Pat Cerretti D, et al. A short polypeptide marker sequence useful for recombinant protein identification and purification. Nat Biotechnol. 1988 Oct;6(10):1204–1210.
13. Walsh CT, Garneau-Tsodikova S, Gatto GJ. Protein posttranslational modifications: the chemistry of proteome diversifications. Angew Chem Int Ed Engl. 2005 Dec 1;44(45):7342–7372.
14. Prabakaran S, Lippens G, Steen H, Gunawardena J. Post-translational modification: nature's escape from genetic imprisonment and the basis for dynamic information encoding. Wiley Interdiscip Rev Syst Biol Med. 2012 Dec;4(6):565–583.
15. Duan G, Walther D. The roles of post-translational modifications in the context of protein interaction networks. PLoS Comput Biol. 2015 Feb 18;11(2):e1004049.
16. KIM S, PAIK W. Studies on the Origin of e-N-Methyl-L-lysine in Protein*.
17. Ambler RP, Rees MW. Epsilon-N-Methyl-lysine in bacterial flagellar protein. Nature. 1959 Jul 4;184:56–57.
18. Murray K. The Occurrence of iε-N-Methyl Lysine in Histones. Biochemistry. 1964 Jan;3(1):10–15.
19. Paik WK, Paik DC, Kim S. Historical review: the field of protein methylation. Trends Biochem Sci. 2007 Mar;32(3):146–152.

20. Wittmann-Liebold B, Pannenbecker R. Primary structure of protein L33 from the large subunit of the Escherichia coli ribosome. FEBS Lett. 1976 Sep 15;68(1):115–118.

21. Brosius J, Chen R. The primary structure of protein L16 located at the peptidyltransferase center of Escherichia coli ribosomes. FEBS Lett. 1976 Sep 15;68(1):105–109.

22. Martinage A, Briand G, Van Dorsselaer A, Turner CH, Sautiere P. Primary structure of histone H2B from gonads of the starfish Asterias rubens. Identification of an N-dimethylproline residue at the amino-terminal. Eur J Biochem. 1985 Mar 1;147(2):351–359.

23. Dognin MJ, Wittmann-Liebold B. The primary structure of L11, the most heavily methylated protein from Escherichia coli ribosomes. FEBS Lett. 1977 Dec 15;84(2):342–346.

24. Alix JH, Hayes D, Lontie JF, Colson C, Glatigny A, Lederer F. Methylated amino acids in ribosomal proteins from Escherichia coli treated with ethionine and from a mutant lacking methylation of protein L11. Biochimie. 1979;61(5-6):671–679.

25. Stock A, Schaeffer E, Koshland DE, Stock J. A second type of protein methylation reaction in bacterial chemotaxis. J Biol Chem. 1987 Jun 15;262(17):8011–8014.

26. Webb KJ, Lipson RS, Al-Hadid Q, Whitelegge JP, Clarke SG. Identification of protein N-terminal methyltransferases in yeast and humans. Biochemistry. 2010 Jun 29;49(25):5225–5235.

27. Tooley CES, Petkowski JJ, Muratore-Schroeder TL, Balsbaugh JL, Shabanowitz J, Sabat M, et al. NRMT is an alpha-N-methyltransferase that methylates RCC1 and retinoblastoma protein. Nature. 2010 Aug 26;466(7310):1125–1128.

28. Jia K, Huang G, Wu W, Shrestha R, Wu B, Xiong Y, et al. In vivo methylation of OLA1 revealed by activity-based target profiling of NTMT1. Chem Sci. 2019 Sep 21;10(35):8094–8099.

29. Dai X, Otake K, You C, Cai Q, Wang Z, Masumoto H, et al. Identification of novel α-n-methylation of CENP-B that regulates its binding to the centromeric DNA. J Proteome Res. 2013 Sep 6;12(9):4167–4175.

30. Cai Q, Fu L, Wang Z, Gan N, Dai X, Wang Y. α-N-methylation of damaged DNA-binding protein 2 (DDB2) and its function in nucleotide excision repair. J Biol Chem. 2014 Jun 6;289(23):16046–16056.

31. Sathyan KM, Fachinetti D, Foltz DR. α-amino trimethylation of CENP-A by NRMT is required for full recruitment of the centromere. Nat Commun. 2017 Mar 7;8:14678.

32. Dong C, Mao Y, Tempel W, Qin S, Li L, Loppnau P, et al. Structural basis for substrate recognition by the human N-terminal methyltransferase 1. Genes Dev. 2015 Nov 15;29(22):2343–2348.

33. Petkowski JJ, Schaner Tooley CE, Anderson LC, Shumilin IA, Balsbaugh JL, Shabanowitz J, et al. Substrate specificity of mammalian N-terminal α-amino methyltransferase NRMT. Biochemistry. 2012 Jul 31;51(30):5942–5950.

34. Petkowski JJ, Bonsignore LA, Tooley JG, Wilkey DW, Merchant ML, Macara IG, et al. NRMT2 is an N-terminal monomethylase that primes for its homologue NRMT1. Biochem J. 2013 Dec 15;456(3):453–462.

35. Mackie BD, Chen D, Dong G, Dong C, Parker H, Schaner Tooley CE, et al. Selective peptidomimetic inhibitors of NTMT1/2: rational design, synthesis, characterization, and crystallographic studies. J Med Chem. 2020 Sep 10;63(17):9512–9522.

36. Faughn JD, Dean WL, Schaner Tooley CE. The N-terminal methyltransferase homologs NRMT1 and NRMT2 exhibit novel regulation of activity through heterotrimer formation. Protein Sci. 2018 Sep 24;27(9):1585–1599.

37. Jakobsson ME, Małecki JM, Halabelian L, Nilges BS, Pinto R, Kudithipudi S, et al. The dual methyltransferase METTL13 targets N terminus and Lys55 of eEF1A and modulates codon-specific translation rates. Nat Commun. 2018 Aug 24;9(1):3411.

38. Liu S, Hausmann S, Carlson SM, Fuentes ME, Francis JW, Pillai R, et al. METTL13 Methylation of eEF1A Increases Translational Output to Promote Tumorigenesis. Cell. 2019 Jan 24;176(3):491–504.e21.

39. Chen T, Muratore TL, Schaner-Tooley CE, Shabanowitz J, Hunt DF, Macara IG. N-terminal alpha-methylation of RCC1 is necessary for stable chromatin association and normal mitosis. Nat Cell Biol. 2007 May;9(5):596–603.

40. Kuersten S, Ohno M, Mattaj IW. Nucleocytoplasmic transport: Ran, beta and beyond. Trends Cell Biol. 2001 Dec;11(12):497–503.

41. Renault L, Kuhlmann J, Henkel A, Wittinghofer A. Structural basis for guanine nucleotide exchange on Ran by the regulator of chromosome condensation (RCC1). Cell. 2001 Apr 20;105(2):245–255.

42. Bischoff FR, Ponstingl H. Catalysis of guanine nucleotide exchange on Ran by the mitotic regulator RCC1. Nature. 1991 Nov 7;354(6348):80–82.

43. Tooley JG, Schaner Tooley CE. New roles for old modifications: emerging roles of N-terminal post-translational modifications in development and disease. Protein Sci. 2014 Dec;23(12):1641–1649.

44. Bonsignore LA, Butler JS, Klinge CM, Schaner Tooley CE. Loss of the N-terminal methyltransferase NRMT1 increases sensitivity to DNA damage and promotes mammary oncogenesis. Oncotarget. 2015 May 20;6(14):12248–12263.

45. Shields KM, Tooley JG, Petkowski JJ, Wilkey DW, Garbett NC, Merchant ML, et al. Select human cancer mutants of NRMT1 alter its catalytic activity and decrease N-terminal trimethylation. Protein Sci. 2017 Aug;26(8):1639–1652.

46. Bonsignore LA, Tooley JG, Van Hoose PM, Wang E, Cheng A, Cole MP, et al. NRMT1 knockout mice exhibit phenotypes associated with impaired DNA repair and premature aging. Mech Ageing Dev. 2015 Mar;146-148:42–52.

47. Elsemman IE, Mardinoglu A, Shoaie S, Soliman TH, Nielsen J. Systems biology analysis of hepatitis C virus infection reveals the role of copy number increases in regions of chromosome 1q in hepatocellular carcinoma metabolism. Mol Biosyst. 2016 Apr 26;12(5):1496–1506.

48. Li L, Zheng Y-L, Jiang C, Fang S, Zeng T-T, Zhu Y-H, et al. HN1L-mediated transcriptional axis AP-2γ/METTL13/TCF3-ZEB1 drives tumor growth and metastasis in hepatocellular carcinoma. Cell Death Differ. 2019 Nov;26(11):2268–2283.

49. Ackermann D, Timpe O, Poller K. Über das Anserin, einen neuen Bestandteil der Vogelmuskulatur. Hoppe-Seyler´s Zeitschrift für physiologische Chemie. 1929 Jan;183(1-2):1–10.

50. Searle JM, Westall RG. The occurrence of free methylhistidine in urine. Biochem J. 1951 Apr;48(4):1.

51. Tallan HH, Stein WH, Moore S. 3-Methylhistidine, a new amino acid from human urine. J Biol Chem. 1954 Feb;206(2):825–834.

52. Johnson P, Perry SV. Biological activity and the 3-methylhistidine content of actin and myosin. Biochem J. 1970 Sep;119(2):293–298.

53. Asatoor AM, Armstrong MD. 3-Methylhistidine, a component of actin. Biochem Biophys Res Commun. 1967 Jan;26(2):168–174.

54. Meyer HE, Mayr GW. N pi-methylhistidine in myosin-light-chain kinase. Biol Chem Hoppe Seyler. 1987 Dec;368(12):1607–1611.

55. Webb KJ, Zurita-Lopez CI, Al-Hadid Q, Laganowsky A, Young BD, Lipson RS, et al. A novel 3-methylhistidine modification of yeast ribosomal protein Rpl3 is dependent upon the YIL110W methyltransferase. J Biol Chem. 2010 Nov 26;285(48):37598–37606.

56. Grabarse W, Mahlert F, Shima S, Thauer RK, Ermler U. Comparison of three methyl-coenzyme M reductases from phylogenetically distant organisms: unusual amino acid modification, conservation and adaptation. J Mol Biol. 2000 Oct 20;303(2):329–344.

57. Raftery MJ, Harrison CA, Alewood P, Jones A, Geczy CL. Isolation of the murine S100 protein MRP14 (14 kDa migration-inhibitory-factor-related protein) from activated spleen cells: characterization of post-translational modifications and zinc binding. Biochem J. 1996 May 15;316 (Pt 1):285–293.

58. Kwiatkowski S, Seliga AK, Vertommen D, Terreri M, Ishikawa T, Grabowska I, et al. SETD3 protein is the actin-specific histidine N-methyltransferase. Elife. 2018 Dec 11;7.

59. Wilkinson AW, Diep J, Dai S, Liu S, Ooi YS, Song D, et al. SETD3 is an actin histidine methyltransferase that prevents primary dystocia. Nature. 2019;565(7739):372–376.

60. Davydova E, Shimazu T, Schuhmacher MK, Jakobsson ME, Willemen HLDM, Liu T, et al. The methyltransferase METTL9 mediates pervasive 1-methylhistidine modification in mammalian proteomes. Nat Commun. 2021 Feb 9;12(1):891.

61. Nyman T, Schüler H, Korenbaum E, Schutt CE, Karlsson R, Lindberg U. The role of MeH73 in actin polymerization and ATP hydrolysis. J Mol Biol. 2002 Apr 5;317(4):577–589.

62. Eom GH, Kim K-B, Kim JH, Kim J-Y, Kim J-R, Kee HJ, et al. Histone methyltransferase SETD3 regulates muscle differentiation. J Biol Chem. 2011 Oct 7;286(40):34733–34742.

63. Zhao M-J, Xie J, Shu W-J, Wang H-Y, Bi J, Jiang W, et al. MiR-15b and miR-322 inhibit SETD3 expression to repress muscle cell differentiation. Cell Death Dis. 2019 Feb 22;10(3):183.

64. Cooper SE, Hodimont E, Green CM. A fluorescent bimolecular complementation screen reveals MAF1, RNF7 and SETD3 as PCNA-associated proteins in human cells. Cell Cycle. 2015 Aug 3;14(15):2509–2519.

65. Li Q, Zhang Y, Jiang Q. SETD3 reduces KLC4 expression to improve the sensitization of cervical cancer cell to radiotherapy. Biochem Biophys Res Commun. 2019 Aug 27;516(3):619–625.

66. Abaev-Schneiderman E, Admoni-Elisha L, Levy D. SETD3 is a positive regulator of DNA-damage-induced apoptosis. Cell Death Dis. 2019 Jan 25;10(2):74.

67. Hassan N, Rutsch N, Győrffy B, Espinoza-Sánchez NA, Götte M. SETD3 acts as a prognostic marker in breast cancer patients and modulates the viability and invasion of breast cancer cells. Sci Rep. 2020 Feb 10;10(1):2262.

68. Xu L, Wang P, Feng X, Tang J, Li L, Zheng X, et al. SETD3 is regulated by a couple of microRNAs and plays opposing roles in proliferation and metastasis of hepatocellular carcinoma. Clin Sci. 2019 Oct 30;133(20):2085–2105.

69. Pan T. Modifications and functional genomics of human transfer RNA. Cell Res. 2018 Apr;28(4):395–404.

70. Sharma S, Lafontaine DLJ. View From A Bridge ': A New Perspective on Eukaryotic rRNA Base Modification. Trends Biochem Sci. 2015 Oct;40(10):560–575.

71. Dominissini D, Moshitch-Moshkovitz S, Schwartz S, Salmon-Divon M, Ungar L, Osenberg S, et al. Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. Nature. 2012 May 10;485(7397):201–206.

72. Perry RP, Kelley DE. Existence of methylated messenger RNA in mouse L cells. Cell. 1974 Jan;1(1):37–42.

73. Sommer S, Lavi U, Darnell JE. The absolute frequency of labeled N-6-methyladenosine in HeLa cell messenger RNA decreases with label time. J Mol Biol. 1978 Sep 25;124(3):487–499.

74. Bokar JA, Shambaugh ME, Polayes D, Matera AG, Rottman FM. Purification and cDNA cloning of the AdoMet-binding subunit of the human mRNA (N6-adenosine)-methyltransferase. RNA. 1997 Nov;3(11):1233–1247.

75. Ping X-L, Sun B-F, Wang L, Xiao W, Yang X, Wang W-J, et al. Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. Cell Res. 2014 Feb;24(2):177–189.

76. Wang P, Doxtader KA, Nam Y. Structural basis for cooperative function of METTL3 and METTL14 methyltransferases. Mol Cell. 2016 Jul 21;63(2):306–317.

77. Jia G, Fu Y, Zhao X, Dai Q, Zheng G, Yang Y, et al. N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. Nat Chem Biol. 2011 Oct 16;7(12):885–887.

78. Jia G, Yang C-G, Yang S, Jian X, Yi C, Zhou Z, et al. Oxidative demethylation of 3-methylthymine and 3-methyluracil in single-stranded DNA and RNA by mouse and human FTO. FEBS Lett. 2008 Oct 15;582(23-24):3313–3319.

79. Mauer J, Luo X, Blanjoie A, Jiao X, Grozhik AV, Patil DP, et al. Reversible methylation of m6Am in the 5' cap controls mRNA stability. Nature. 2017 Jan 19;541(7637):371–375.

80. Zheng G, Dahl JA, Niu Y, Fedorcsak P, Huang C-M, Li CJ, et al. ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. Mol Cell. 2013 Jan 10;49(1):18–29.

81. Wang X, Zhao BS, Roundtree IA, Lu Z, Han D, Ma H, et al. N(6)-methyladenosine Modulates Messenger RNA Translation Efficiency. Cell. 2015 Jun 4;161(6):1388–1399.

82. Shi H, Wang X, Lu Z, Zhao BS, Ma H, Hsu PJ, et al. YTHDF3 facilitates translation and decay of N6-methyladenosine-modified RNA. Cell Res. 2017 Mar;27(3):315–328.

83. Zaccara S, Jaffrey SR. A Unified Model for the Function of YTHDF Proteins in Regulating m6A-Modified mRNA. Cell. 2020 Jun 25;181(7):1582–1595.e18.

84. Visvanathan A, Patil V, Arora A, Hegde AS, Arivazhagan A, Santosh V, et al. Essential role of METTL3-mediated m6A modification in glioma stem-like cells maintenance and radioresistance. Oncogene. 2018 Jan 25;37(4):522–533.

85. Taketo K, Konno M, Asai A, Koseki J, Toratani M, Satoh T, et al. The epitranscriptome m6A writer METTL3 promotes chemo- and radioresistance in pancreatic cancer cells. Int J Oncol. 2018 Feb;52(2):621–629.

86. Liu J, Eckert MA, Harada BT, Liu S-M, Lu Z, Yu K, et al. m6A mRNA methylation regulates AKT activity to promote the proliferation and tumorigenicity of endometrial cancer. Nat Cell Biol. 2018 Aug 27;20(9):1074–1083.

87. Hu A, Noble WS, Wolf-Yadlin A. Technical advances in proteomics: new developments in data-independent acquisition. [version 1; peer review: 3 approved]. F1000Res. 2016 Mar 31;5.

88. Venable JD, Dong M-Q, Wohlschlegel J, Dillin A, Yates JR. Automated approach for quantitative analysis of complex peptide mixtures from tandem mass spectra. Nat Methods. 2004 Oct;1(1):39–45.

89. Huttlin EL, Jedrychowski MP, Elias JE, Goswami T, Rad R, Beausoleil SA, et al. A tissue-specific atlas of mouse protein phosphorylation and expression. Cell. 2010 Dec 23;143(7):1174–1189.

90. Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic Acids Res. 2019 Jan 8;47(D1):D442–D450.

91. Choi M, Carver J, Chiva C, Tzouros M, Huang T, Tsai T-H, et al. MassIVE.quant: a community resource of quantitative mass spectrometry-based proteomics datasets. Nat Methods. 2020 Sep 14;17(10):981–984.

92. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol. 2008 Dec;26(12):1367–1372.

93. Patricelli MP, Szardenings AK, Liyanage M, Nomanbhoy TK, Wu M, Weissig H, et al. Functional interrogation of the kinome using nucleotide acyl phosphates. Biochemistry. 2007 Jan 16;46(2):350–358.

94. Cai R, Huang M, Wang Y. Targeted Quantitative Profiling of GTP-Binding Proteins in Cancer Cells Using Isotope-Coded GTP Probes. Anal Chem. 2018 Dec 18;90(24):14339–14346.

95. Doerr A. Targeting with PRM. Nat Methods. 2012 Oct;9(10):950.

96. Marx V. Targeted proteomics. Nat Methods. 2012 Dec 27;10(1):19–22.

97. Lange V, Picotti P, Domon B, Aebersold R. Selected reaction monitoring for quantitative proteomics: a tutorial. Mol Syst Biol. 2008 Oct 14;4:222.

98. Escher C, Reiter L, MacLean B, Ossola R, Herzog F, Chilton J, et al. Using iRT, a normalized retention time for more targeted measurement of peptides. Proteomics. 2012 Apr;12(8):1111–1121.

99. Gianazza E, Tremoli E, Banfi C. The selected reaction monitoring/multiple reaction monitoring-based mass spectrometry approach for the accurate quantitation of proteins: clinical applications in the cardiovascular diseases. Expert Rev Proteomics. 2014 Dec;11(6):771–788.

100. Bourmaud A, Gallien S, Domon B. Parallel reaction monitoring using quadrupole-Orbitrap mass spectrometer: Principle and applications. Proteomics. 2016 May 27;16(15-16):2146–2159.

101. Peterson AC, Russell JD, Bailey DJ, Westphall MS, Coon JJ. Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics. Mol Cell Proteomics. 2012 Nov;11(11):1475–1488.

102. Haag AM. Mass analyzers and mass spectrometers. Adv Exp Med Biol. 2016;919:157–169.

103. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, et al. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics. 2010 Apr 1;26(7):966–968.

104. Ong S-E, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, et al. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics. 2002 May;1(5):376–386.

105. Fleuren EDG, Zhang L, Wu J, Daly RJ. The kinome “at large” in cancer. Nat Rev Cancer. 2016 Feb;16(2):83–98.

106. Harbeck N, Penault-Llorca F, Cortes J, Gnant M, Houssami N, Poortmans P, et al. Breast cancer. Nat Rev Dis Primers. 2019 Sep 23;5(1):66.

107. Speers C, Zhao SG, Kothari V, Santola A, Liu M, Wilder-Romans K, et al. Maternal embryonic leucine zipper kinase (MELK) as a novel mediator and biomarker of radioresistance in human breast cancer. Clin Cancer Res. 2016 Dec 1;22(23):5864–5875.

108. Bian L, Meng Y, Zhang M, Guo Z, Liu F, Zhang W, et al. ATM Expression Is Elevated in Established Radiation-Resistant Breast Cancer Cells and Improves DNA Repair Efficiency. Int J Biol Sci. 2020 Feb 4;16(7):1096–1106.

109. Lavin MF, Kozlov S, Gatei M, Kijas AW. ATM-Dependent Phosphorylation of All Three Members of the MRN Complex: From Sensor to Adaptor. Biomolecules. 2015 Oct 23;5(4):2877–2902.

110. Kren N, Michaud D, Bagchi S, Greene K, Pylayeva-Gupta Y. Rab27a plays a dual role in metastatic propensity of pancreatic cancer. Sci Rep. 2020 Apr 30;10(1):7390.

111. Wang J-S, Wang F-B, Zhang Q-G, Shen Z-Z, Shao Z-M. Enhanced expression of Rab27A gene by breast cancer cells promoting invasiveness and the metastasis potential by secretion of insulin-like growth factor-II. Mol Cancer Res. 2008 Mar;6(3):372–382.

112. Boone B, Van Gele M, Lambert J, Haspeslagh M, Brochez L. The role of RhoC in growth and metastatic capacity of melanoma. J Cutan Pathol. 2009 Jun;36(6):629–636.

113. Maldonado MDM, Dharmawardhane S. Targeting rac and cdc42 gtpases in cancer. Cancer Res. 2018 Jun 15;78(12):3101–3111.

Figures

Figure 1. 1 - Method for enrichment of proteins of interest for PTM identification

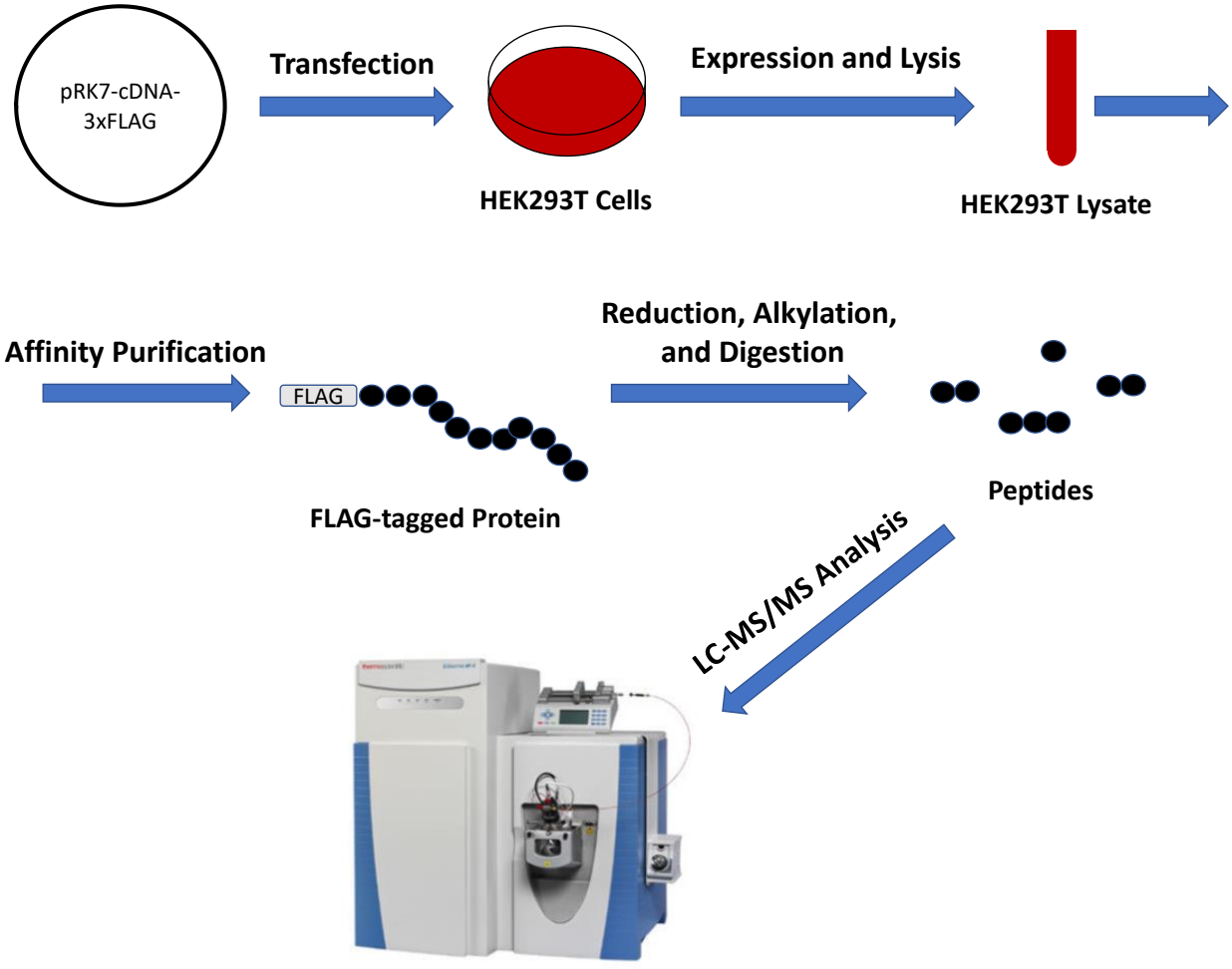


Figure 1. 2- Distributive α-N-methylation of RCC1 by NTMT1 and NTMT2

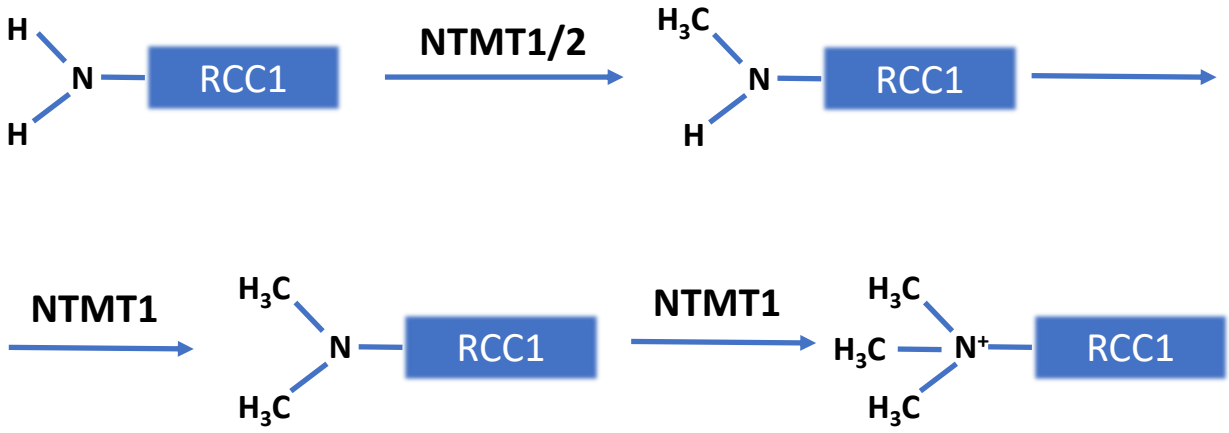


Figure 1. 3- m^{6}A Regulation Diagram

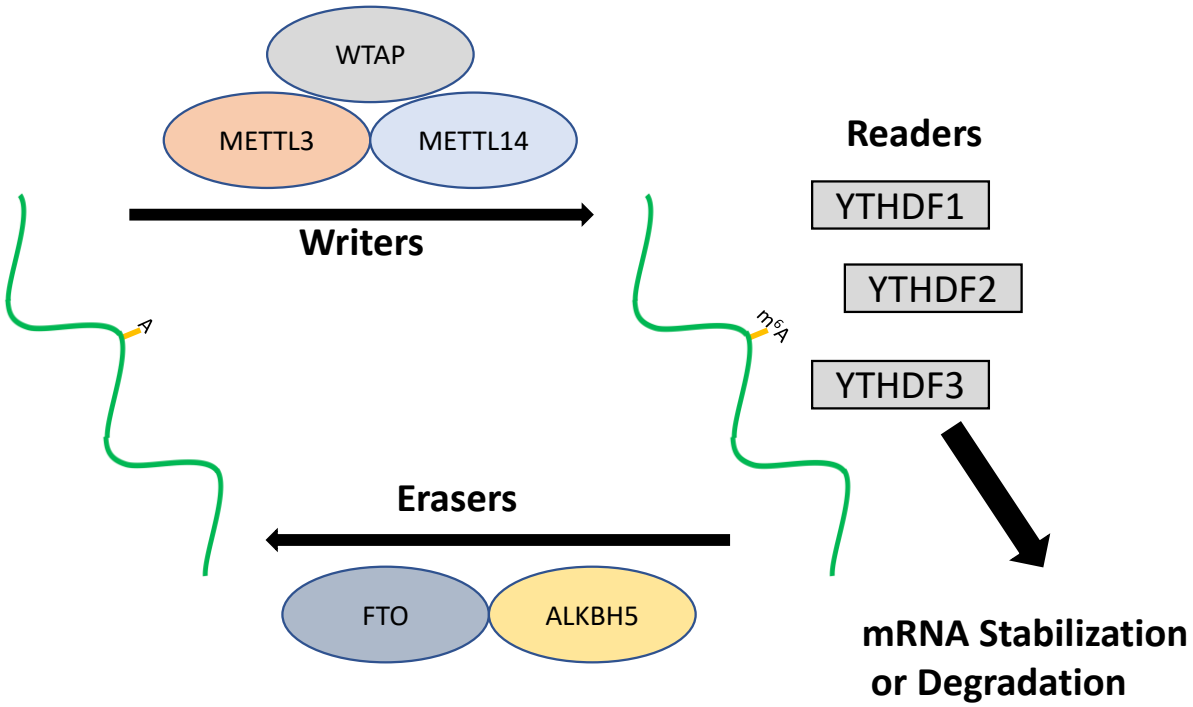


Figure 1. 4- Schematic of Bottom-up Mass Spectrometry Scanning Methods

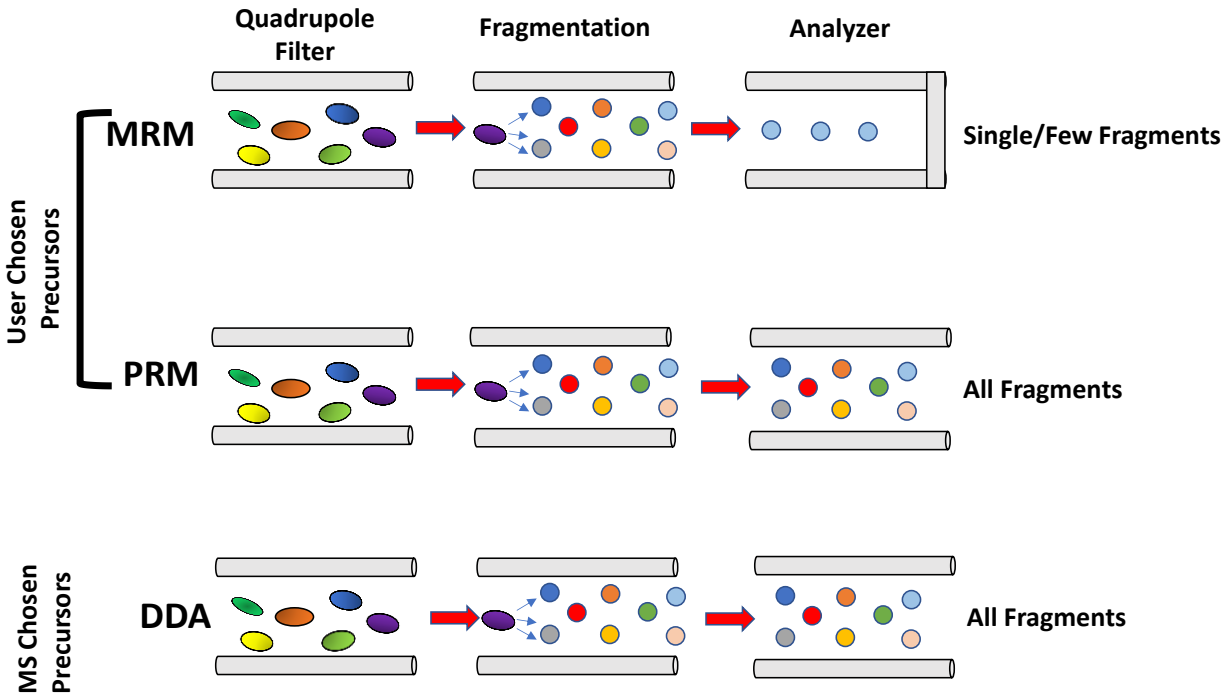


Figure 1. 5- Diagram of SILAC-based Labelling

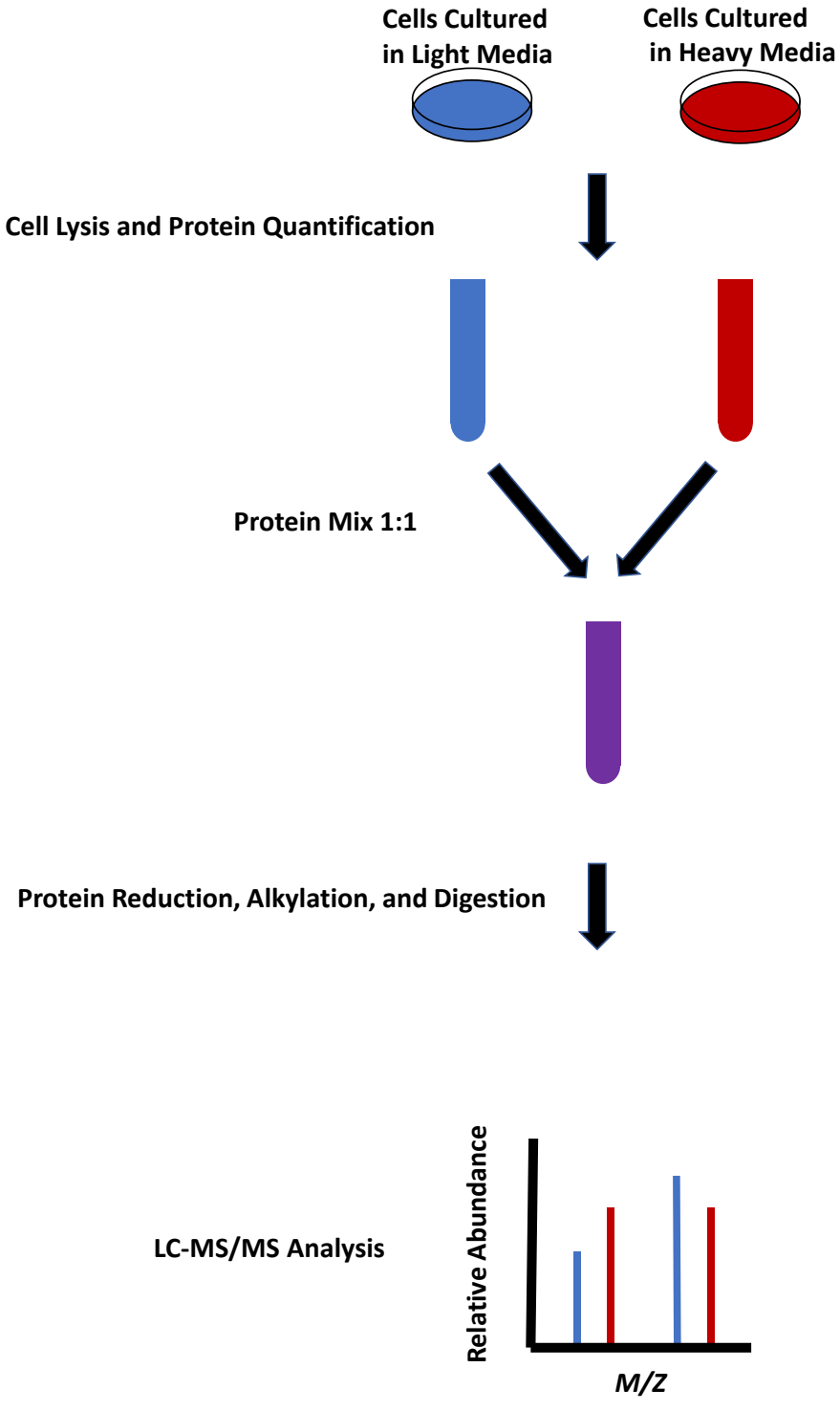
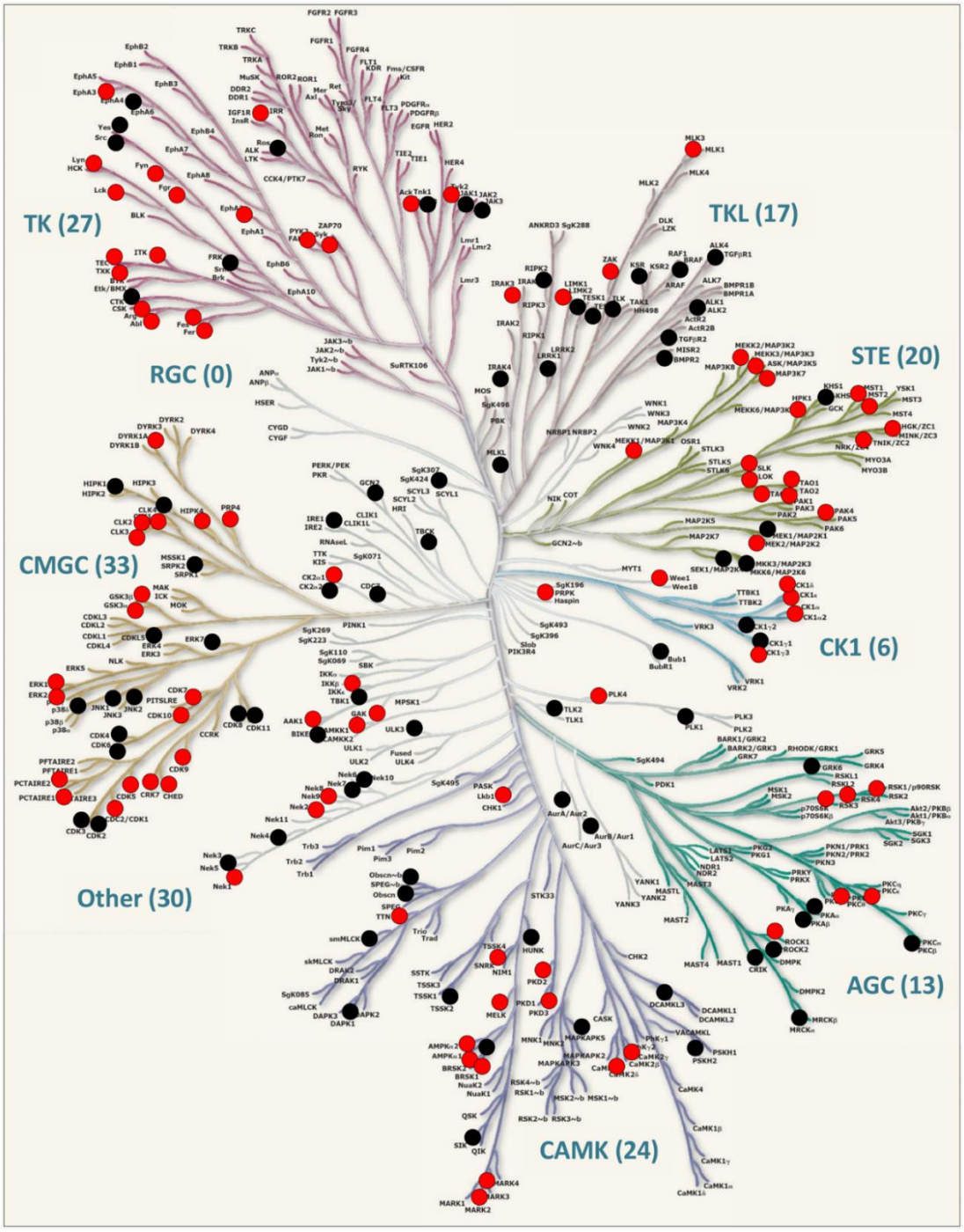


Figure 1. 6- The dendrogram of the human kinome (adopted from König S, Nimtz M, Scheiter M, Ljunggren HG, Bryceson YT, et al. (2012) Kinome Analysis of Receptor-Induced Phosphorylation in Human Natural Killer Cells. PLOS ONE 7(1): e29672.).



Chapter 2: Modulation of N-terminal Methyltransferase 1 by an N^6-methyladenosine-based Epitranscriptomic Mechanism

1. Introduction

Post-transcriptional modifications of mRNAs and post-translational modifications (PTMs) of proteins constitute evolutionarily conserved mechanisms of gene regulation. α-N-methylation of proteins was initially observed several decades ago [1, 2]. N-terminal methyltransferase 1 (NTMT1) was the first discovered enzyme for protein α-N-methylation [3], and a number of substrate proteins have been identified for NTMT1, including regulator of chromatin condensation 1 (RCC1), centromere proteins A and B (CENP-A and CENP-B), damage DNA-binding protein 2 (DDB2), and poly(ADP-ribose) polymerase 3 [3-8]. In this vein, N-terminal methylations of CENP-A and CENP-B promote the recruitment of kinetochore and binding of CENP-B to α-satellite DNA, respectively [6, 8, 9], and loss of function of NTMT1 could result in diminished DNA repair and elevated sensitivity to DNA damaging agents [5, 10, 11]. In addition, NTMT2 forms a heterodimer with NTMT1 and enhances its function in protein α-N-methylation [12].

NTMT1 is down-regulated in breast cancer tissues [10] and mutations disrupting the catalytic functions of NTMT1 were found in different types of cancer [13], suggesting its potential as a therapeutic target. Moreover, loss of function of NTMT1 was shown to elicit premature aging in mice [11]. Despite these discoveries, N-terminal methylation is still a

poorly understood form of PTM, where little is known about the regulatory mechanisms of this modification.

Recent studies revealed that the N^6 position of adenosine in RNA can be dynamically methylated and this methylation assumes important functions in normal development [14]; dysregulation of this methylation is implicated in many human diseases, including cancer and neurological disorders [15]. Methyltransferase-like 3 (METTL3), together with other regulatory subunits, including METTL14, RBM15, RBM15B, WTAP, VIRMA, and ZC3H13, were shown to constitute the major methyltransferase complex for m^6A generation [14, 16, 17]. m^6A in RNA can be recognized by a number of proteins, including the YTH domain-containing family proteins 1/2/3 (YTHDF1/2/3) [18-21]. In addition, m^6A in mRNA can be demethylated by FTO and ALKBH5 [22, 23].

In this study, we found that NTMT1 protein expression is subjected to regulation by m^6A reader, writer and eraser proteins. We also showed that these m^6A regulators could modulate the N-terminal methylation of MRG15 protein.

2. Materials and Methods

2.1. Plasmid preparation

The coding sequence of human *MRG15* gene was amplified from a cDNA library prepared from mRNAs isolated from HEK293T cells and cloned into the pRK7 plasmid between the XbaI and BamHI restriction sites, in which three tandem repeats of Flag epitope tag (DYKDDDDK) were fused at the carboxyl terminus. MRG15-K4Q mutant

plasmid was amplified from the MRG15 plasmid with primers designed to contain the mutation.

2.2. Cell Culture

HEK293T human embryonic kidney cells (ATCC), and the isogenic *METTL3*$^{-/-}$, *FTO*$^{-/-}$, *ALKBH5*$^{-/-}$, *YTHDF1*$^{-/-}$, *YTHDF2*$^{-/-}$, and *YTHDF3*$^{-/-}$ cells, which were previously generated by CRISPR-Cas9 [24, 25], were cultured in Dulbecco's Modified Eagle Medium (DMEM; Fisher Scientific, Hampton, NH). All culture media were supplemented with 10% fetal bovine serum (FBS; Fisher Scientific, Hampton, NH), and 100 IU/mL penicillin/streptomycin (GE Healthcare, Chicago, IL). The cells were maintained at 37°C in a humidified environment containing 5% CO_2.

2.3. Preparation of Flag-tagged MRG15 protein and subsequent LC-MS/MS analysis

HEK293T cells and the aforementioned knockout cells were cultured in antibiotic-free media at a density of 5×10^6 cells/well. Plasmid for ectopic expression of Flag-tagged MRG15 (1 μg) was transfected into cells cultured in 6-well plates using Lipofectamine 2000 (Life Technologies, Carlsbad, CA). After a 24-hr incubation, the cells were harvested and lysed in CelLytic M (Sigma-Aldrich, St. Louis, MO), supplemented with a protease inhibitor cocktail (Sigma-Aldrich). C-terminally Flag-tagged MRG15 protein was isolated from the resultant lysates by affinity purification with anti-FLAG M2 beads (Sigma-Aldrich), and digested with Glu-C (NEB) at a protein/enzyme ratio of 10:1.

The Glu-C-produced peptides were analyzed by LC-MS/MS on an LTQ XL linear ion trap mass spectrometer equipped with a nanoelectrospray ionization source (Thermo

Scientific, San Jose, CA) and with an Easy-nLC II. Separation was conducted by using a home-made trapping column (150 μm × 50 mm) and a separation column (75 μm × 120 mm), both of which were packed with ReproSil-Pur C18-AQ resin (3 μm in diameter, Dr. Maisch HPLC GmbH, Ammerbuch-Entringen, Germany). Peptide samples were initially loaded onto the trapping column with a mixture of 0.1% formic acid in CH_3CN/H_2O (2:98, v/v) at a flow rate of 5.0 μl/min. The peptides were separated using at 40-min linear gradient of 2-40% acetonitrile in 0.1% formic acid, and the flow rate was 300 μl/min. The mass spectrometer was operated in the positive-ion mode with a spray voltage of 1.8 kV. The full-scan mass spectra were acquired in the *m/z* range of 150-1000. MS/MS were recorded in a selected-ion monitoring mode, where the doubly protonated ions of unmodified, mono-, di-, and tri-methylated forms of N-terminal peptide of MRG15 were chosen for fragmentation. All data were analyzed manually.

2.4. siRNA knockdown of NTMT1

The control siRNA and human NTMT1 SMARTpool siRNA were obtained from Thermo Scientific. Sequences of NTMT1 SMARTpool siRNA were GCGAGGUGAUAGAAGACGA, AGGUGGAUAUGGUCGACAU, UGAGGGAAGGCCCGAACAA and GGACUGUGGAGCUGGCAUU. HEK293T cells were cultured in 6-well plates in antibiotic-free medium at a density of 5×10^5 cells per well for 24 hr, and each well of cells were transfected with 100 pmol siRNA using Lipofectamine 2000 (Invitrogen).

2.5. Western Blot Analysis

HEK293T cells and the isogenic CRISPR knockout cells were lysed at 50-80% confluency with CelLytic M, supplemented with a protease inhibitor cocktail (Sigma-Aldrich). The protein concentrations of the resultant lysates were determined with Bradford Reagent (Bio-Rad, Hercules, CA), and the whole cell lysate (12 μg) was denatured by boiling in Laemmli loading buffer (Bio-Rad). The lysate was resolved with SDS-PAGE and transferred to a nitrocellulose membrane (Bio-Rad) at 4°C overnight. The resultant membrane was blocked with PBS-T (PBS with 0.1% Tween 20) supplemented with 5% powdered milk (Bio-Rad) at room temperature for 45 min, and subsequently incubated with primary antibody at room temperature for 2 hr, and then with secondary antibody at room temperature for 1 hr. The HRP signal from Amersham ECL Select western blotting detection reagent was then recorded (GE Healthcare, Chicago, IL). Antibodies recognizing human NTMT1 (Abcam, Cambridge, United Kingdom, ab72660, 1:2,500 dilution) and GAPDH (Santa Cruz Biotechnology, Dallas, TX, sc-32233, 1:5,000) were used as primary antibodies for Western blot analyses. Anti-rabbit IgG peroxidase antibody (Sigma-Aldrich, A0545, 1:10,000 dilution), and anti-mouse IgG kappa binding peroxidase antibody (Santa Cruz Biotechnology, sc-516102, 1:5,000 dilution) were employed as secondary antibodies.

3. Results and Discussion

3.1. The expression of NTMT1 protein is subjected to regulation by m^{6}A reader, writer and eraser proteins

We set out to investigate if the expression of NTMT1 protein can be altered upon genetic depletion of m^{6}A reader, writer, and eraser proteins. We first analyzed the expression level of NTMT1 protein in HEK293T cells or the isogenic cells with genetic ablation of *METTL3*, which encodes the catalytic subunit of the major m^{6}A writer complex

[16], or either of the two m^6A eraser proteins, FTO and ALKBH5 [22, 23]. Our Western blot results revealed that NTMT1 protein level was increased in HEK293T cells upon genetic depletion of METTL3, but decreased upon genetic depletion of FTO (Figure 2.1). Genetic depletion of ALKBH5, nevertheless, did not induce any appreciable changes in the expression level of NTMT1 protein (Figure 2.1).

We also assessed how genetic knockout of YTHDF1/2/3 proteins affects the expression level of NTMT1 protein in HEK293T cells. We observed elevated expression of NTMT1 protein in HEK293T cells upon genetic ablation of YTHDF1 or YTHDF2, whereas depletion of YTHDF3 resulted in diminished expression of NTMT1 protein (Figure 2.1). Together, the above results demonstrated that NTMT1 protein expression can be dynamically regulated by m^6A reader/writer/eraser proteins.

3.2. MRG15 is N-terminally methylated by NTMT1

Previous studies revealed a number of substrates for NTMT1, and many of them carry an N-terminal XPK motif ('X' represents any amino acid) after the removal of the initial methionine [3]. The N-terminus of MRG15 possesses such a motif. Hence, we next asked whether MRG15 is α-N-methylated. To this end, we constructed a plasmid for ectopic expression of MRG15 protein, which is fused with 3 tandem repeats of Flag epitope tag on its C-terminus, in HEK293T cells. We subsequently isolated the Flag-tagged protein from the whole cell lysate with anti-Flag M2 affinity gel, digested the protein with Glu-C, and analyzed the ensuing mixture with LC-MS and MS/MS. Our results revealed that MRG15 is methylated on the N-terminus (Figure 2.2a-b).

Next we sought to determine whether NTMT1 is responsible for the N-terminal methylation of MRG15. Our results revealed that siRNA-mediated knockdown of NTMT1 led to nearly complete loss of α-N-methylation of MRG15 (Figure 2.2c-d). Previous studies showed that the 4th lysine is important for the NTMT1-mediated protein α-N-methylation [3]. We found that the α-N-methylation of MRG15 is entirely abolished in its variant with the 4th lysine residue being mutated to a glutamine (MRG15-K4Q) (Figure 2.2a). These results support that MRG15 is a substrate for NTMT1.

3.3. N-terminal methylation on MRG15 is modulated through an m^{6}A-based epitranscriptomic mechanism

Our above results revealed that NTMT1 protein is subjected to regulation by reader, writer and eraser proteins of m^{6}A. We subsequently asked whether α-N-methylation of MRG15 is also subjected to regulation by these m^{6}A modulators. In line with the findings made from NTMT1 protein, we observed a notable increase in N-terminal trimethylation of MRG15, which is accompanied with a concomitant decrease in unmethylated N-terminal peptide upon genetic ablation of METTL3, whereas an opposite trend was found for MRG15 isolated from the isogenic HEK293T cells depleted of FTO (Figure 2.3).

We also examined how genetic ablation of YTHDF1/2/3 affects N-terminal methylation of MRG15. We found that individual depletion of YTHDF1 or YTHDF2 led to a significant increase in trimethylation, which is associated with a concomitant decrease in monomethylation of the N-terminus of MRG15 (Figure 2.3C & E). Conversely, genetic depletion of YTHDF3 led to a significant decrease in α-N-methylation of MRG15 (Figure 2.3D).

4. Conclusions

To summarize, we found that NTMT1, the major methyltransferase for protein α-N-methylation, could be subjected to regulation by m^6A reader, writer, and eraser proteins. To our knowledge, this is the first report about the regulatory mechanism for this evolutionarily conserved type of post-translational modification. We also uncovered that MRG15 is N-terminally methylated by NTMT1 and demonstrated that this methylation is modulated in a similar way as NTMT1 by the m^6A reader, writer, and eraser proteins.

References

1. A. Stock, S. Clarke, C. Clarke, J. Stock, N-terminal methylation of proteins: structure, function and specificity, FEBS Lett., 220 (1987) 8-14.
2. J.G. Tooley, C.E. Schaner Tooley, New roles for old modifications: emerging roles of N-terminal post-translational modifications in development and disease, Protein Sci., 23 (2014) 1641-1649.
3. C.E. Tooley, J.J. Petkowski, T.L. Muratore-Schroeder, J.L. Balsbaugh, J. Shabanowitz, M. Sabat, W. Minor, D.F. Hunt, I.G. Macara, NRMT is an a-N-methyltransferase that methylates RCC1 and retinoblastoma protein, Nature, 466 (2010) 1125-1128.
4. T. Chen, T.L. Muratore, C.E. Schaner-Tooley, J. Shabanowitz, D.F. Hunt, I.G. Macara, N-terminal alpha-methylation of RCC1 is necessary for stable chromatin association and normal mitosis, Nat. Cell Biol., 9 (2007) 596-603.
5. Q. Cai, L. Fu, Z. Wang, N. Gan, X. Dai, Y. Wang, a-*N*-Methylation of damaged DNA-binding protein 2 (DDB2) and its function in nucleotide excision repair, J. Biol. Chem., 289 (2014) 16046-16056.
6. X. Dai, K. Otake, C. You, Q. Cai, Z. Wang, H. Masumoto, Y. Wang, Identification of novel alpha-*N*-methylation of CENP-B that regulates its binding to the centromeric DNA, J. Proteome Res., 12 (2013) 4167-4175.
7. X. Dai, S.L. Rulten, C. You, K.W. Caldecott, Y. Wang, Identification and functional characterizations of N-terminal a-N-methylation and phosphorylation of serine 461 in human poly(ADP-ribose) polymerase 3, J. Proteome Res., 14 (2015) 2575-2582.
8. K.M. Sathyan, D. Fachinetti, D.R. Foltz, a-amino trimethylation of CENP-A by NRMT is required for full recruitment of the centromere, Nat. Commun., 8 (2017) 14678.
9. A.O. Bailey, T. Panchenko, K.M. Sathyan, J.J. Petkowski, P.J. Pai, D.L. Bai, D.H. Russell, I.G. Macara, J. Shabanowitz, D.F. Hunt, B.E. Black, D.R. Foltz, Posttranslational modification of CENP-A influences the conformation of centromeric chromatin, Proc. Natl. Acad. Sci. USA, 110 (2013) 11827-11832.
10. L.A. Bonsignore, J.S. Butler, C.M. Klinge, C.E. Schaner Tooley, Loss of the N-terminal methyltransferase NRMT1 increases sensitivity to DNA damage and promotes mammary oncogenesis, Oncotarget, 6 (2015) 12248-12263.
11. L.A. Bonsignore, J.G. Tooley, P.M. Van Hoose, E. Wang, A. Cheng, M.P. Cole, C.E. Schaner Tooley, NRMT1 knockout mice exhibit phenotypes associated with impaired DNA repair and premature aging, Mech. Ageing Dev., 146-148 (2015) 42-52.
12. J.D. Faughn, W.L. Dean, C.E. Schaner Tooley, The N-terminal methyltransferase homologs NRMT1 and NRMT2 exhibit novel regulation of activity through heterotrimer formation, Protein Sci., 27 (2018) 1585-1599.
13. K.M. Shields, J.G. Tooley, J.J. Petkowski, D.W. Wilkey, N.C. Garbett, M.L. Merchant, A. Cheng, C.E. Schaner Tooley, Select human cancer mutants of NRMT1 alter its catalytic activity and decrease N-terminal trimethylation, Protein Sci., 26 (2017) 1639-1652.
14. M. Frye, B.T. Harada, M. Behm, C. He, RNA modifications modulate gene expression during development, Science, 361 (2018) 1346-1349.
15. H. Huang, H. Weng, J. Chen, m^6A modification in coding and non-coding RNAs: Roles and therapeutic implications in cancer, Cancer Cell, 37 (2020) 270-288.

16. J.A. Bokar, M.E. Shambaugh, D. Polayes, A.G. Matera, F.M. Rottman, Purification and cDNA cloning of the AdoMet-binding subunit of the human mRNA (N6-adenosine)-methyltransferase, RNA, 3 (1997) 1233-1247.

17. J. Liu, Y. Yue, D. Han, X. Wang, Y. Fu, L. Zhang, G. Jia, M. Yu, Z. Lu, X. Deng, Q. Dai, W. Chen, C. He, A METTL3-METTL14 complex mediates mammalian nuclear RNA N^6-adenosine methylation, Nat. Chem. Biol., 10 (2014) 93-95.

18. P.J. Hsu, Y. Zhu, H. Ma, Y. Guo, X. Shi, Y. Liu, M. Qi, Z. Lu, H. Shi, J. Wang, Y. Cheng, G. Luo, Q. Dai, M. Liu, X. Guo, J. Sha, B. Shen, C. He, Ythdc2 is an N^6-methyladenosine binding protein that regulates mammalian spermatogenesis, Cell Res., 27 (2017) 1115-1127.

19. A. Li, Y.S. Chen, X.L. Ping, X. Yang, W. Xiao, Y. Yang, H.Y. Sun, Q. Zhu, P. Baidya, X. Wang, D.P. Bhattarai, Y.L. Zhao, B.F. Sun, Y.G. Yang, Cytoplasmic m^6A reader YTHDF3 promotes mRNA translation, Cell Res., 27 (2017) 444-447.

20. X. Wang, B.S. Zhao, I.A. Roundtree, Z. Lu, D. Han, H. Ma, X. Weng, K. Chen, H. Shi, C. He, N^6-methyladenosine modulates messenger RNA translation efficiency, Cell, 161 (2015) 1388-1399.

21. S. Zaccara, S.R. Jaffrey, A unified model for the function of YTHDF proteins in regulating m^6A-modified mRNA, Cell, (2020) doi: 10.1016/j.cell.2020.1005.1012.

22. G. Jia, Y. Fu, X. Zhao, Q. Dai, G. Zheng, Y. Yang, C. Yi, T. Lindahl, T. Pan, Y.G. Yang, C. He, N^6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO, Nat. Chem. Biol., 7 (2011) 885-887.

23. G. Zheng, J.A. Dahl, Y. Niu, P. Fedorcsak, C.M. Huang, C.J. Li, C.B. Vagbo, Y. Shi, W.L. Wang, S.H. Song, Z. Lu, R.P. Bosmans, Q. Dai, Y.J. Hao, X. Yang, W.M. Zhao, W.M. Tong, X.J. Wang, F. Bogdan, K. Furu, Y. Fu, G. Jia, X. Zhao, J. Liu, H.E. Krokan, A. Klungland, Y.G. Yang, C. He, ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility, Mol. Cell, 49 (2013) 18-29.

24. W. Miao, L. Li, Y. Zhao, X. Dai, X. Chen, Y. Wang, HSP90 inhibitors stimulate DNAJB4 protein expression through a mechanism involving N^6-methyladenosine, Nat Commun, 10 (2019) 3613.

25. Y.Y. Yang, K. Yu, L. Li, M. Huang, Y. Wang, Proteome-wide interrogation of small GTPases regulated by N^6-methyladenosine modulators, Anal. Chem., 92 (2020) 10145-10152.

Figure 2. 1. The alterations in expression levels of NTMT1 protein following knockout of m6A reader/writer/eraser genes.

(a-c) Western blot images showing the relative levels of NTMT1 protein in HEK293T cells and the isogenic cells with m^6A erasers (a), writer (b) and readers (c) being genetically depleted by CRISPR-Cas9. (d) Quantification results for the relative expression levels of NTMT1 protein in HEK293T cells vs. the isogenic cells with ALKBH5, FTO, METTL3 and YTHDF1/2/3 genes being individually depleted. "ns", $p > 0.05$; "∗", $0.01 \leq p < 0.05$; "∗∗∗", $p < 0.001$. The data represent the mean ± S.D. of results from 3 or 4 independent experiments. All p values were calculated using the unpaired, two-tailed t-test.

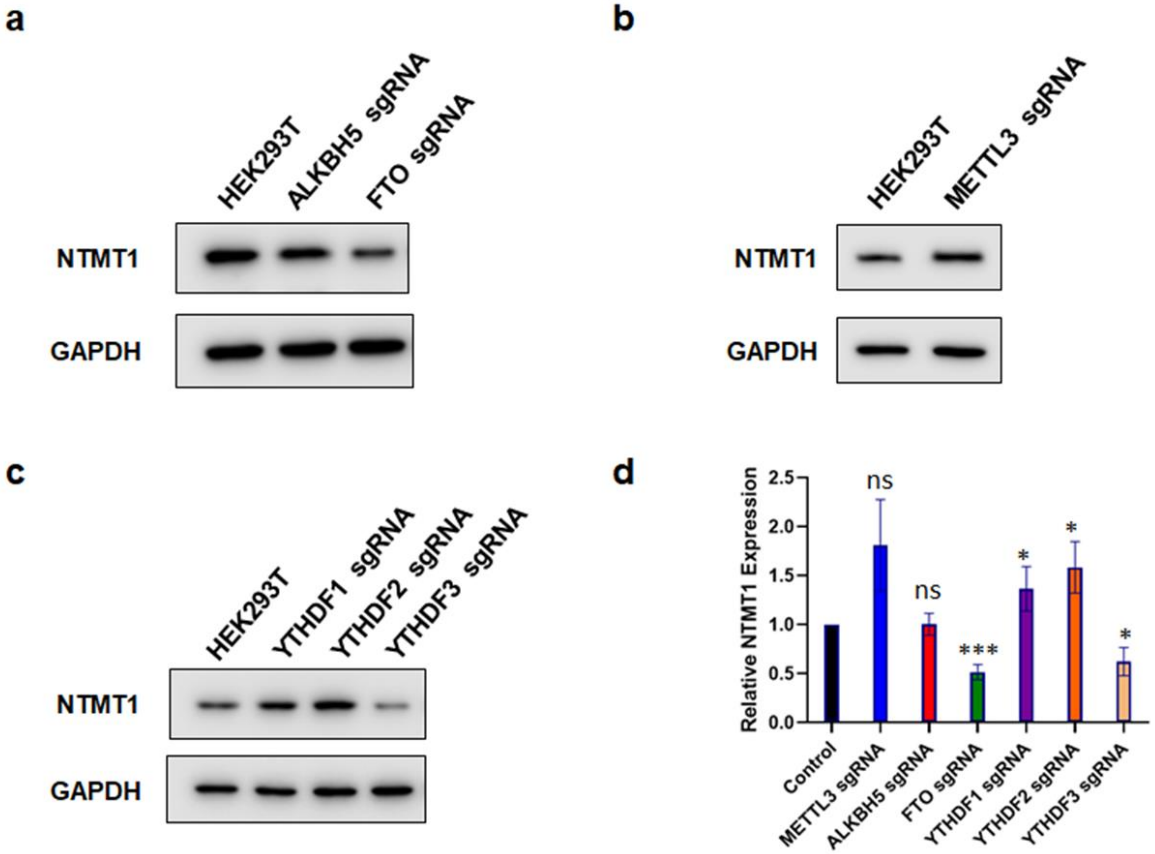


Figure 2. 2. Identification and characterization of α-N-terminal methylation of MRG15.

(a) 'Ultra-zoom' scan ESI-MS showing the $[M+2H]^{2+}$ ions of the Glu-C-produced N-terminal peptide of MRG15 and its K4Q mutant, i.e. APKQDPKPKFQE and APQQPKPKFQE, isolated from HEK293T cells. (b) MS/MS of tri-methylated N-terminal peptide of MRG15. Displayed in the inset is a scheme summarizing the observed fragment ions. (c) RT-qPCR and Western blot showing the efficient knockdown of *NTMT1* gene in HEK293T cells. (d) Knockdown of NTMT1 led to marked attenuation of α-N-methylation of MRG15. The methylation levels were quantified based on the relative abundances of the $[M+2H]^{2+}$ ions of the unmethylated and the mono-, di- and tri-methylated forms of the N-terminal peptide of MRG15. The *p* values were calculated using the unpaired, two-tailed *t*-test.

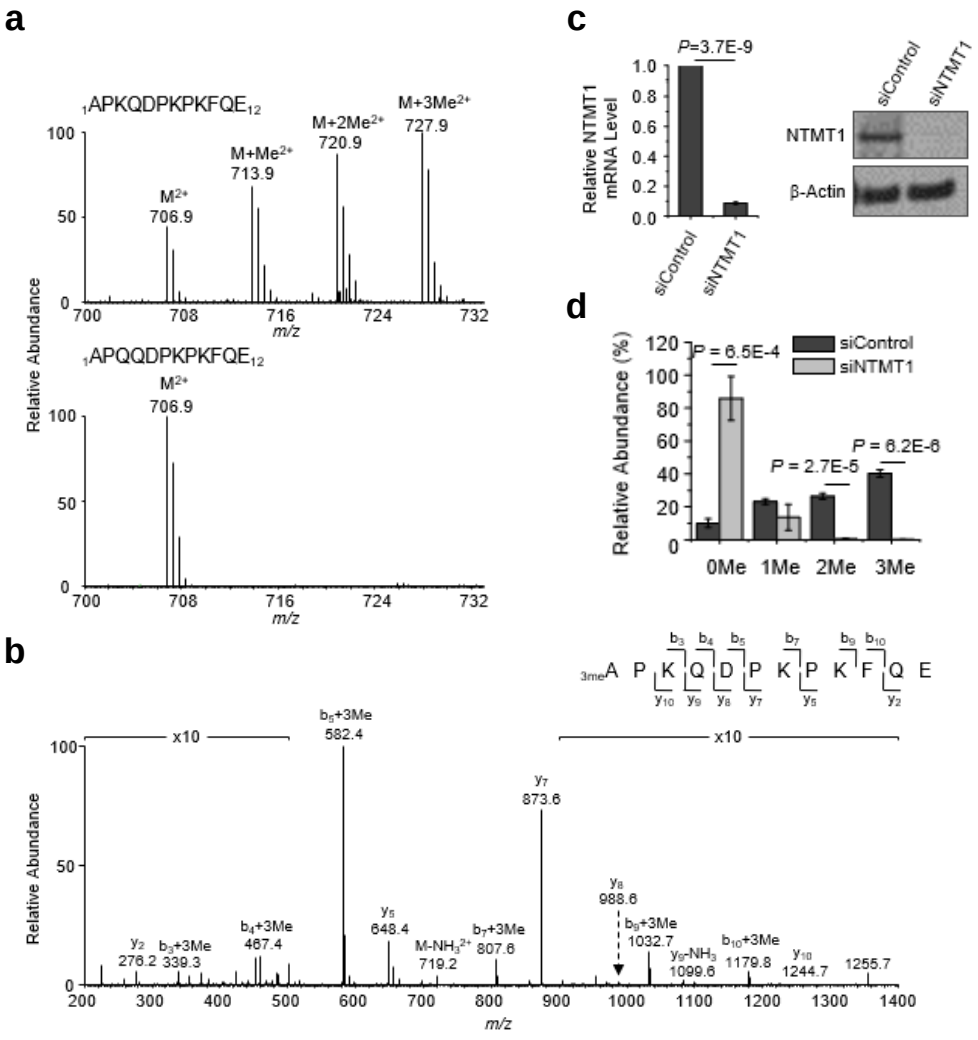
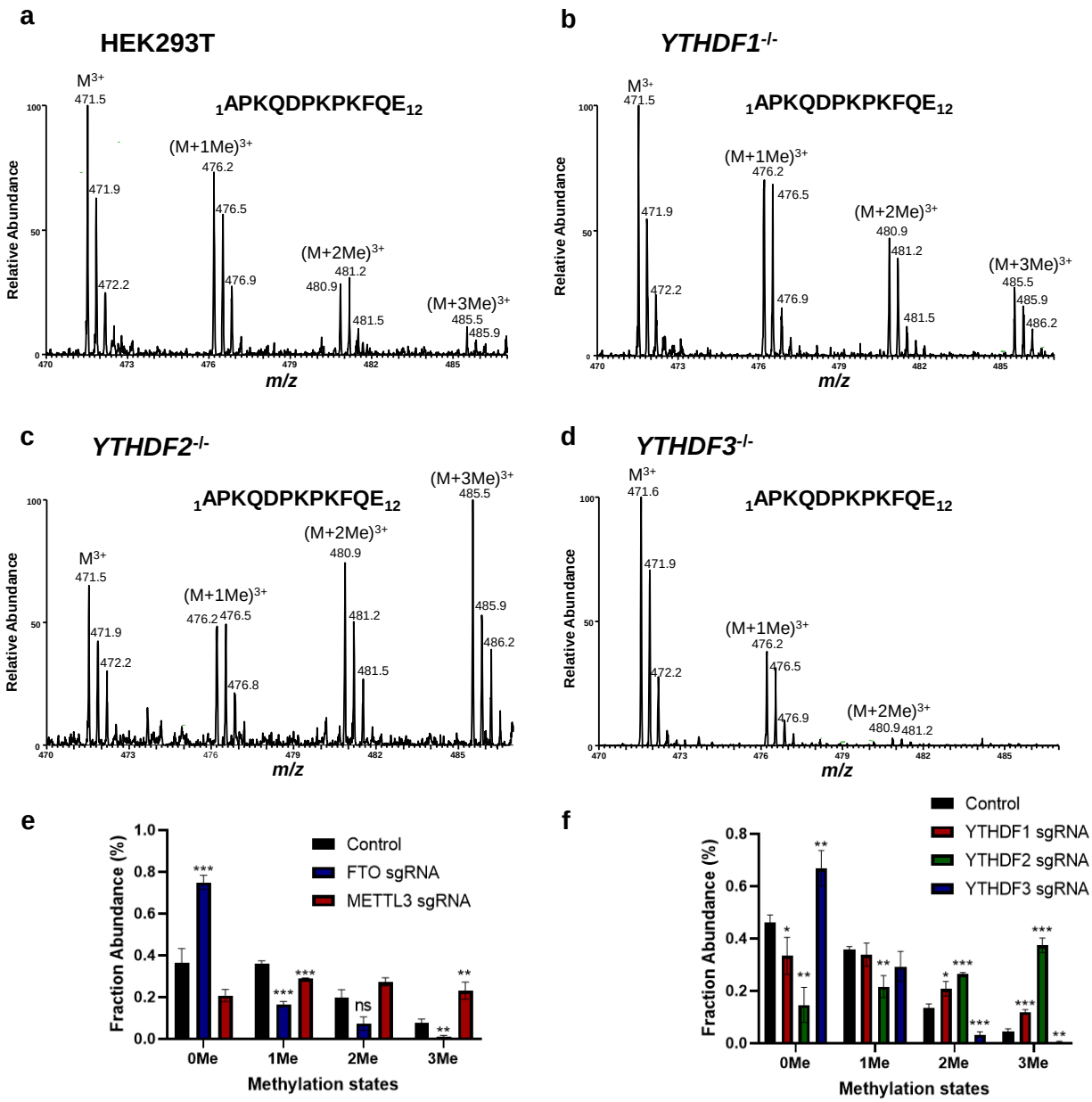


Figure 2. 3. N-terminal methylation levels of MRG15 are modulated by m6A reader/writer/eraser genes.

(a-d) 'Ultra-zoom' scan ESI-MS showing the $[M+3H]^{3+}$ ions of the Glu-C-produced N-terminal peptide of C-terminally Flag-tagged MRG15 isolated from HEK293T cells or the isogenic cells with YTHDF1/2/3 genes being individually ablated by CRISPR-Cas9. (e-f) Genetic ablations of FTO, METTL3, YTHDF1, YTHDF2, and YTHDF3 alter the α-N-methylation of MRG15. The methylation levels were quantified based on the relative abundances of the $[M+3H]^{3+}$ ions of the unmethylated and the mono-, di- and tri-methylated forms of the N-terminal peptide of MRG15. The *p* values were calculated using unpaired, two-tailed *t*-test: "ns", $p > 0.05$; "∗", $0.01 \leq p < 0.05$; "**", $0.001 \leq p < 0.01$; "∗∗∗", $p < 0.001$.


a
HEK293T
$_1$APKQDPKPKFQE$_{12}$
M3+ 471.5
(M+1Me)3+ 476.2
(M+2Me)3+ 480.9
(M+3Me)3+ 485.5
Relative Abundance
m/z
b
YTHDF1-/-
$_1$APKQDPKPKFQE$_{12}$
M3+ 471.5
(M+1Me)3+ 476.2
(M+2Me)3+ 480.9
(M+3Me)3+ 485.5
Relative Abundance
m/z
c
YTHDF2-/-
$_1$APKQDPKPKFQE$_{12}$
M3+ 471.5
(M+1Me)3+ 476.2
(M+2Me)3+ 480.9
(M+3Me)3+ 485.5
Relative Abundance
m/z
d
YTHDF3-/-
$_1$APKQDPKPKFQE$_{12}$
M3+ 471.6
(M+1Me)3+ 476.2
(M+2Me)3+ 480.9
Relative Abundance
m/z
e
Control
FTO sgRNA
METTL3 sgRNA
Fraction Abundance (%)
0Me
1Me
2Me
3Me
Methylation states
f
Control
YTHDF1 sgRNA
YTHDF2 sgRNA
YTHDF3 sgRNA
Fraction Abundance (%)
0Me
1Me
2Me
3Me
Methylation states

Chapter 3: Analysis of Publicly Available Datasets Reveals Novel α-N-Methylated Protein Substrates

1. Introduction

Post-translational modifications (PTMs) of proteins constitute a ubiquitous and evolutionarily conserved mechanism of gene regulation. Among these, α-N-methylation was first discovered decades ago in *Escherichia coli* (110). N-terminal methyltransferase 1 (NTMT1) was the first discovered enzyme that catalyzes protein α-N-methylation (29). Since its discovery, several NTMT1 substrate proteins have been identified, including regulator of chromatin condensation 1 (RCC1), damage DNA-binding protein 2 (DDB2), poly(ADP-ribose) polymerase 3 (PARP3), and Obg Like ATPase 1 (OLA1) (29,32,33,111). Although these targets all possess a common XPK motif (wherein X is alanine, proline, or serine), further *in vitro* and structural studies support that this consensus motif may be more flexible than the currently known substrates suggest (36). Interestingly, eEF1A lysine and N-terminal methyltransferase (EEF1AKNMT) was recently discovered to methylate EEF1A1 at the α-N-terminus (34), but it possessed an N-terminal sequence distinct from the XPK motif. This discovery enlarges the potential pool of N-terminal methylated proteins, suggesting the presence of many unidentified substrates for α-N-methylation.

Protein α-N-methylation appears to play several important roles in cancer. Knockdown of NTMT1 can increase the effectiveness of etoposide treatment and γ-irradiation in breast cancer cell lines while promoting tamoxifen sensitivity (46). Likewise, EEF1AKNMT is down-regulated in bladder carcinoma, as increased expression suppresses cell

proliferation, migration, and invasion (112). Conversely, EEF1AKNMT has also been shown to suppress apoptosis (113). These discoveries demonstrate a vital role of α-N-methylation in affecting cancer progression and prognosis of patients.

Mass spectrometry-based techniques have been widely used in high-throughput proteomic analysis and have contributed a significant source of data on potential α-N-methylation targets. Spectra produced through shotgun proteomic analysis, such as those acquired from data-dependent acquisition (DDA), allow for the repeated interrogation of data for different projects. In the past couple of decades, several databases have been created to coordinate the uploading of publicly available raw datasets, enabling the global exchange of mass spectrometry metadata (92,93,114). These data can then be compared against the human proteome with MaxQuant, to identify peptides containing PTMs (94).

In this study, we utilized these publicly available data sets and analyzed them with MaxQuant. We discovered 219 unique instances of N-terminal methylation on 196 individual proteins. We also showed that the α-N-methylation of Vesicle-associated membrane protein 4 (VAMP4) was due predominantly to the enzyme NTMT1.

2. Materials and Methods

2.1. Plasmid Preparation

The coding sequence of the human *VAMP4* was amplified from a cDNA library prepared from HEK293T cells and cloned into the pRK7 plasmid between the XbaI and BamHI restriction sites, in which three tandem repeats of Flag epitope tag (DYKDDDDK)

were fused at the carboxyl terminus. VAMP4-K4Q mutant plasmid was amplified from the VAMP4 plasmid with primers designed to contain the mutation.

2.2. Preparation of NTMT1 Knockout Cells using CRISPR-Cas9

Genome editing by the CRISPR/Cas9 system was conducted following the previously reported protocols (115), where the single guide RNAs (sgRNAs) were designed by using the online sgRNA tool (http://www.broadinstitute.org/rnai/public/analysis-tools/sgrna-design). ODNs corresponding to target sequences were obtained from IDT and inserted into the hSpCas9 plasmid pX330 (Addgene). The constructed plasmids were then transfected into HEK293T cells using Lipofectamine 2000 (Invitrogen) in a 6-well plate, and individual cells were cultured for further analysis. Genomic DNA was extracted from distinct clonal cell lines, and specific DNA regions surrounding the targeted sites were screened by PCR, followed by agarose gel electrophoresis to assess the modification efficiency and by Sanger sequencing to identify the deletion loci. A set of clones where both alleles were cleaved by Cas9 was isolated, and the successful deletion of NTMT1 was validated by Western blot analysis. The guide sequence was ACGACGTGATCTGGATCCAG **TGG** (letters in bold indicate the PAM motif).

2.3. Cell Culture

HEK293T cells and isogenic *NTMT1*$^{-/-}$ cells were cultured in Dulbecco's modified eagle medium (DMEM) supplemented with 10% fetal bovine serum (Invitrogen, Carlsbad, CA) and penicillin (100 IU/mL). The cells were maintained at 37°C in a humidified environment containing 5% CO_2.

2.4. Tryptic digestion and Mass Spectrometry Analysis of Whole-cell Protein Lysates

The whole-cell lysates prepared from HEK293T cells were combined at 1:1 ratio (by mass, determined by Bradford assay), and 30 µg protein lysate were loaded onto a 10% SDS-PAGE gel. After electrophoresis, the gel lanes were cut into 11 slices according to apparent molecular weight ranges of proteins (< 20, 20-25, 25-30, 30-37, 37-42, 42-50, 50-62, 62-75, 75-100, 100-150, >150 kDa), reduced in-gel with dithiothreitol, and alkylated with iodoacetamide. The processed proteins were subsequently digested in-gel with Trypsin (Thermo Fisher) at an enzyme/substrate ratio of 1:100 in 50 mM NH_4HCO_3 (pH 8.5) at 37 °C overnight. Subsequently, peptides were recovered from gels with a solution containing 5% acetic acid in H_2O and then with a solution containing 2.5% acetic acid in an equal-volume mixture of CH_3CN and H_2O.

Samples were automatically loaded at 3 µL/min onto a precolumn (150 µm i.d. and 3.5 cm in length) packed with ReproSil-Pur 120 C18-AQ stationary-phase material (5 µm in particle size, 120 Å in pore size, Dr. Maisch). The precolumn was connected to a 20-cm fused-silica analytical column (PicoTip Emitter, New Objective, 75 µm i.d.) packed with 3 µm C18 beads (ReproSil-Pur 120 C18-AQ, Dr. Maisch). The peptides were then resolved using a 180-min gradient of 2-45% acetonitrile in 0.1% formic acid, and the flow rate was maintained at 300 nL/min.

The mass spectrometer was operated in a data-dependent acquisition mode. Full-scan mass spectra were acquired in the range of *m/z* 350-1500 using the Orbitrap analyzer at a

resolution of 70,000 at *m/z* 200. Up to 25 most abundant ions found in MS with a charge state of 2 or above were sequentially isolated and collisionally activated in the HCD cell with a normalized collision energy of 28 to yield MS/MS.

2.5. DDA Data File Processing with MaxQuant Analysis

We first downloaded previously published raw LC-MS/MS data (116–118) and searched them, along with the in-house proteomic data generated from the proteomic samples of HEK293T cells files, against the *Homo Sapiens* IPI protein database (version 3.68) or the Uniprot protein FASTA (UP000005640) through MaxQuant quantitative software (94). The maximum number of miss-cleavages for trypsin was two per peptide. Cysteine carbamidomethylation was set as a fixed modification, while monomethylation, dimethylation, and trimethylation of the α-N-terminal peptide were set as variable modifications. The tolerances in mass accuracy were 25 ppm and 0.6 Da for MS and MS/MS, respectively. Identified peptides under a posterior error probability score (PEP) of 0.01 were considered significant. Unique proteins were further subjected to Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis.

2.6. LC-MS/MS Identification of Novel α-N-methylated Proteins

HEK293T cells and isogenic *NTMT1*$^{-/-}$ cells were cultured in 6-well plates containing antibiotic-free DMEM media at a density of 5×10^6 cells/well. Plasmid for ectopic expression of FLAG-tagged VAMP4 (1.5 μg) was transfected into cells using

Lipofectamine 2000 (Life Technologies, Carlsbad, CA). For inhibitor experiments, the cells were treated with 10 μM Y2-6, a small-molecule inhibitor of NTMT1 developed by Prof. Rong Huang at Purdue University. After a 24-hr incubation, the cells were harvested and lysed in CelLytic M (Sigma-Aldrich, St. Louis, MO), supplemented with a protease inhibitor cocktail (Sigma-Aldrich). C-terminally Flag-tagged wild-type and K4Q mutant VAMP4 protein were isolated from the resultant lysates by affinity purification with anti-FLAG M2 beads (Sigma-Aldrich), and digested with Glu-C (NEB) at a protein/enzyme ratio of 10:1.

The resultant peptides from Glu-C-digestion were analyzed by LC-MS/MS on an LTQ XL linear ion trap mass spectrometer equipped with a nanoelectrospray ionization source (Thermo Scientific, San Jose, CA) and an Easy-nLC II nano-HPLC system. The separation was conducted by using a trapping column (150 μm × 50 mm) and a separation column (75 μm × 120 mm), both of which were packed in-laboratory with ReproSil-Pur C18-AQ resin (3 μm in diameter, Dr. Maisch HPLC GmbH, Ammerbuch-Entringen, Germany). Peptide samples were first loaded onto the trapping column with a mixture of 0.1% formic acid in CH_3CN/H_2O (2:98, v/v) at a flow rate of 5.0 μl/min. The peptides were separated using a 40-min linear gradient of 2-40% acetonitrile in 0.1% formic acid, and the flow rate was 300 μl/min. The mass spectrometer was operated in the positive-ion mode with a spray voltage of 1.8 kV. The full-scan mass spectra were acquired in the *m/z* range of 150-1000, while MS/MS were recorded in a selected-ion monitoring mode. Quadruply protonated ions of unmodified, mono-, and di- methylated forms of the VAMP4 N-terminal peptide were selected for fragmentation. All data were analyzed manually.

3. Results and Discussion

3.1. Proteome-wide Identification of Novel α-N-terminal Methylated Proteins

By employing the peptide identification strategy with MaxQuant detailed in the Materials and Methods section, we identified 219 unique instances of α-N-methylation on a total of 196 total proteins. Each of these identifications exhibited a PEP value under 0.01. The MS/MS of the N-terminal peptides of VAMP4, RAN, and DDX39B are shown in Figure 1, which support their α-N- methylation (Figure 3.1a-c).

Gene Ontology and KEGG pathway analysis suggest that the identified N-terminally methylated proteins are significantly associated with RNA binding, particularly associated with Poly(A) and spliceosome. In addition, most of the proteins identified can bind other proteins, particularly with translation at the ribosome (Figure 3.1d-e). This suggests a significantly broader list of substrates than previously indicated, most with N-terminal sequences that have hitherto been unassociated with proteins that exhibit α-N-methylation.

3.2. Identification of α-N-methylation of VAMP4

From our list of potential α-N-terminal methylated genes, we chose to confirm the methylation of VAMP4. VAMP4 possess a canonical N-terminal XPK motif, consistent with previously discovered substrates for NTMT1. We constructed plasmids wherein the coding sequence of VAMP4 is c-terminally fused with three consecutive repeats of the FLAG epitope tag. The plasmids were expressed in *HEK293T* cells, after which the FLAG-tagged protein was enriched from the cell lysate through affinity purification with anti-

FLAG M2 beads digested with Glu-C, and the resulting peptides were analyzed with LC-MS and MS/MS. Our results confirm that VAMP4 is α-N-methylated (Figure 3.2a-c).

3.3. NTMT1 catalyzes the α-N-methylation of VAMP4

Recognizing that NTMT1 is known to methylate proteins with a similar N-terminal sequence as VAMP4, we sought to determine whether NTMT1 is responsible for this α-N-methylation. To that end, we compared relative VAMP4 α-N-methylation between *HEK293T* cells and isogenic cells with the genetic ablation of NTMT1. In cells lacking NTMT1 expression, we observed a loss of VAMP4 N-terminal methylation (Figure 3.3a-b, 3.4a).

Recent research has documented the discovery of inhibitors targeting the N-terminal methyltransferase activity of NTMT1 (119,120). To further confirm that NTMT1 is responsible for the α-N-methylation of VAMP4, we treated *HEK293T* cells with a small molecule inhibitor of NTMT1. It turned out that, inhibition of NTMT1 by a small molecule also gave rise to marked decreases in α-N-methylation of VAMP4 (Figure 3.4b).

To investigate this further, we produced a VAMP4-K4Q mutant plasmid. Previous research has shown that this mutation blocks the ability of NTMT1 to methylate its substrate protein (31,32). Consistent with previous studies (31,32), we found that replacement of the 4th lysine in the N-terminal sequence of VAMP4 with a glutamine led to a nearly complete loss α-N-methylation (Figure 3.3c, 3.4a).

4. Conclusions

In this study, we analyzed publicly available mass spectrometry DDA files with MaxQuant to discover novel α-N-terminal methylate proteins. With this method, we identified 219 unique instances of N-terminal methylation among 196 individual proteins. Among those proteins, we confirmed the N-terminal methylation of VAMP4 through the enrichment of FLAG-tagged protein. We also uncovered that NTMT1 is the primary protein responsible for the N-terminal methylation of VAMP4.

RAN is essential for regulating the nucleocytoplasmic transport of primarily proteins and RNA. This is regulated by the dynamic binding of GDP or GTP, wherein RAN is bound to GDP in the cytoplasm, and GTP in the nucleus (42). Uniquely, RAN's function is regulated by RCC1, the first protein known to be α-N-methylated by NTMT1 (29). RCC1 promotes both the transfer of GDP bound RAN to GTP bound RAN, thereby regulating nucleocytoplasmic transport (43,44). In addition, RCC1 promotes the binding of GTP bound RAN to chromatin, which is necessary for proper spindle assembly (41,121).

VAMP4 is a known component of the SNARE complex, with several established roles in regulating intracellular transport. This binding is believed to occur in a 'zipper' mode, wherein protein attachment begins at the N-terminus and moves to the C-terminus, suggesting an essential role of the N-terminus in initial binding (122). In neurons, VAMP4 regulates asynchronous transport of neurotransmitters (123,124). Beyond this, VAMP4 can form a complex with SNAP23, Syntaxin-5, and α-SNAP, which promotes lipid droplet fusion (125). It will be important to examine how α-N-methylation modulates these biological functions of VAMP4.

References

1. Stock A, Clarke S, Clarke C, Stock J. N-terminal methylation of proteins: structure, function and specificity. FEBS Lett. 1987 Aug 10;220(1):8–14.
2. Tooley CES, Petkowski JJ, Muratore-Schroeder TL, Balsbaugh JL, Shabanowitz J, Sabat M, et al. NRMT is an alpha-N-methyltransferase that methylates RCC1 and retinoblastoma protein. Nature. 2010 Aug 26;466(7310):1125–1128.
3. Cai Q, Fu L, Wang Z, Gan N, Dai X, Wang Y. α-N-methylation of damaged DNA-binding protein 2 (DDB2) and its function in nucleotide excision repair. J Biol Chem. 2014 Jun 6;289(23):16046–16056.
4. Dai X, Rulten SL, You C, Caldecott KW, Wang Y. Identification and Functional Characterizations of N-Terminal α-N-Methylation and Phosphorylation of Serine 461 in Human Poly(ADP-ribose) Polymerase 3. J Proteome Res. 2015 Jun 5;14(6):2575–2582.
5. Jia K, Huang G, Wu W, Shrestha R, Wu B, Xiong Y, et al. In vivo methylation of OLA1 revealed by activity-based target profiling of NTMT1. Chem Sci. 2019 Sep 21;10(35):8094–8099.
6. Dong C, Mao Y, Tempel W, Qin S, Li L, Loppnau P, et al. Structural basis for substrate recognition by the human N-terminal methyltransferase 1. Genes Dev. 2015 Nov 15;29(22):2343–2348.
7. Jakobsson ME, Małecki JM, Halabelian L, Nilges BS, Pinto R, Kudithipudi S, et al. The dual methyltransferase METTL13 targets N terminus and Lys55 of eEF1A and modulates codon-specific translation rates. Nat Commun. 2018 Aug 24;9(1):3411.
8. Bonsignore LA, Butler JS, Klinge CM, Schaner Tooley CE. Loss of the N-terminal methyltransferase NRMT1 increases sensitivity to DNA damage and promotes mammary oncogenesis. Oncotarget. 2015 May 20;6(14):12248–12263.
9. Zhang Z, Zhang G, Kong C, Zhan B, Dong X, Man X. METTL13 is downregulated in bladder carcinoma and suppresses cell proliferation, migration and invasion. Sci Rep. 2016 Jan 14;6:19261.
10. Liang H, Fu Z, Jiang X, Wang N, Wang F, Wang X, et al. miR-16 promotes the apoptosis of human cancer cells by targeting FEAT. BMC Cancer. 2015 Jun 2;15:448.
11. Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic Acids Res. 2019 Jan 8;47(D1):D442–D450.
12. Desiere F, Deutsch EW, King NL, Nesvizhskii AI, Mallick P, Eng J, et al. The PeptideAtlas project. Nucleic Acids Res. 2006 Jan 1;34(Database issue):D655–8.
13. Choi M, Carver J, Chiva C, Tzouros M, Huang T, Tsai T-H, et al. MassIVE.quant: a community resource of quantitative mass spectrometry-based proteomics datasets. Nat Methods. 2020 Sep 14;17(10):981–984.
14. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol. 2008 Dec;26(12):1367–1372.
15. Sakuma T, Nishikawa A, Kume S, Chayama K, Yamamoto T. Multiplex genome engineering in human cells using all-in-one CRISPR/Cas9 vector system. Sci Rep. 2014 Jun 23;4:5400.

16. Nagaraj N, Wisniewski JR, Geiger T, Cox J, Kircher M, Kelso J, et al. Deep proteome and transcriptome mapping of a human cancer cell line. Mol Syst Biol. 2011 Nov 8;7:548.

17. Yeom J, Ju S, Choi Y, Paek E, Lee C. Comprehensive analysis of human protein N-termini enables assessment of various protein forms. Sci Rep. 2017 Jul 26;7(1):6599.

18. Qi TF, Guo L, Huang M, Li L, Miao W, Wang Y. Discovery of TBC1D7 as a potential driver for melanoma cell invasion. Proteomics. 2020 Jul 2;20(14):e1900347.

19. Mackie BD, Chen D, Dong G, Dong C, Parker H, Schaner Tooley CE, et al. Selective peptidomimetic inhibitors of NTMT1/2: rational design, synthesis, characterization, and crystallographic studies. J Med Chem. 2020 Sep 10;63(17):9512–9522.

20. Dong G, Yasgar A, Peterson DL, Zakharov A, Talley D, Cheng KC-C, et al. Optimization of High-Throughput Methyltransferase Assays for the Discovery of Small Molecule Inhibitors. ACS Comb Sci. 2020 Aug 10;22(8):422–432.

21. Dai X, Otake K, You C, Cai Q, Wang Z, Masumoto H, et al. Identification of novel α-n-methylation of CENP-B that regulates its binding to the centromeric DNA. J Proteome Res. 2013 Sep 6;12(9):4167–4175.

22. Kuersten S, Ohno M, Mattaj IW. Nucleocytoplasmic transport: Ran, beta and beyond. Trends Cell Biol. 2001 Dec;11(12):497–503.

23. Bischoff FR, Ponstingl H. Catalysis of guanine nucleotide exchange on Ran by the mitotic regulator RCC1. Nature. 1991 Nov 7;354(6348):80–82.

24. Renault L, Kuhlmann J, Henkel A, Wittinghofer A. Structural basis for guanine nucleotide exchange on Ran by the regulator of chromosome condensation (RCC1). Cell. 2001 Apr 20;105(2):245–255.

25. Halpin D, Kalab P, Wang J, Weis K, Heald R. Mitotic spindle assembly around RCC1-coated beads in Xenopus egg extracts. PLoS Biol. 2011 Dec 1;9(12):e1001225.

26. Chen T, Muratore TL, Schaner-Tooley CE, Shabanowitz J, Hunt DF, Macara IG. N-terminal alpha-methylation of RCC1 is necessary for stable chromatin association and normal mitosis. Nat Cell Biol. 2007 May;9(5):596–603.

27. Pobbati AV, Stein A, Fasshauer D. N- to C-terminal SNARE complex assembly promotes rapid membrane fusion. Science. 2006 Aug 4;313(5787):673–676.

28. Raingo J, Khvotchev M, Liu P, Darios F, Li YC, Ramirez DMO, et al. VAMP4 directs synaptic vesicles to a pool that selectively maintains asynchronous neurotransmission. Nat Neurosci. 2012 May 1;15(5):738–745.

29. Lin P-Y, Chanaday NL, Horvath PM, Ramirez DMO, Monteggia LM, Kavalali ET. VAMP4 Maintains a Ca2+-Sensitive Pool of Spontaneously Recycling Synaptic Vesicles. J Neurosci. 2020 Jul 8;40(28):5389–5401.

30. Boström P, Andersson L, Rutberg M, Perman J, Lidberg U, Johansson BR, et al. SNARE proteins mediate fusion between cytosolic lipid droplets and are implicated in insulin sensitivity. Nat Cell Biol. 2007 Nov;9(11):1286–1293.

Figure 3. 1. ESI-MS/MS of the N-terminal peptides of representative α-N-methylated proteins. and Gene Ontology and KEGG pathway analysis of Novel α-N-methylated proteins.

MS/MS of the N-terminal peptides of VAMP4 (a), RAN (b), and DDX39b (c). Gene Ontology (d) and KEGG pathway (e) analysis of novel α-N-methylated proteins.

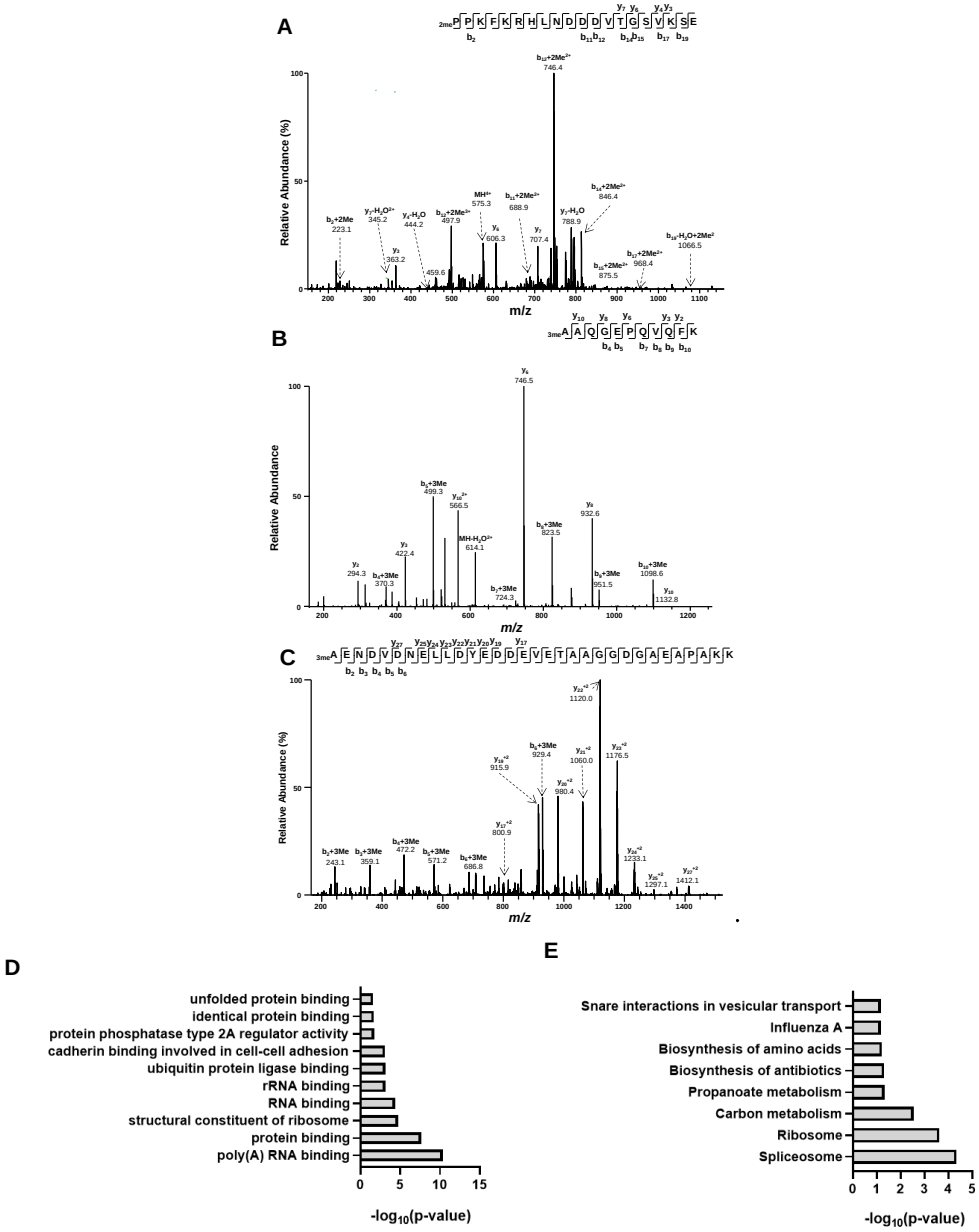


Figure 3. 2. Identification of α-N-Methylation of VAMP4.

MS/MS of the $[M + 3H]^{3+}$ ions of unmethylated (a), monomethylated (b), and dimethylated N-terminal peptide of VAMP4.

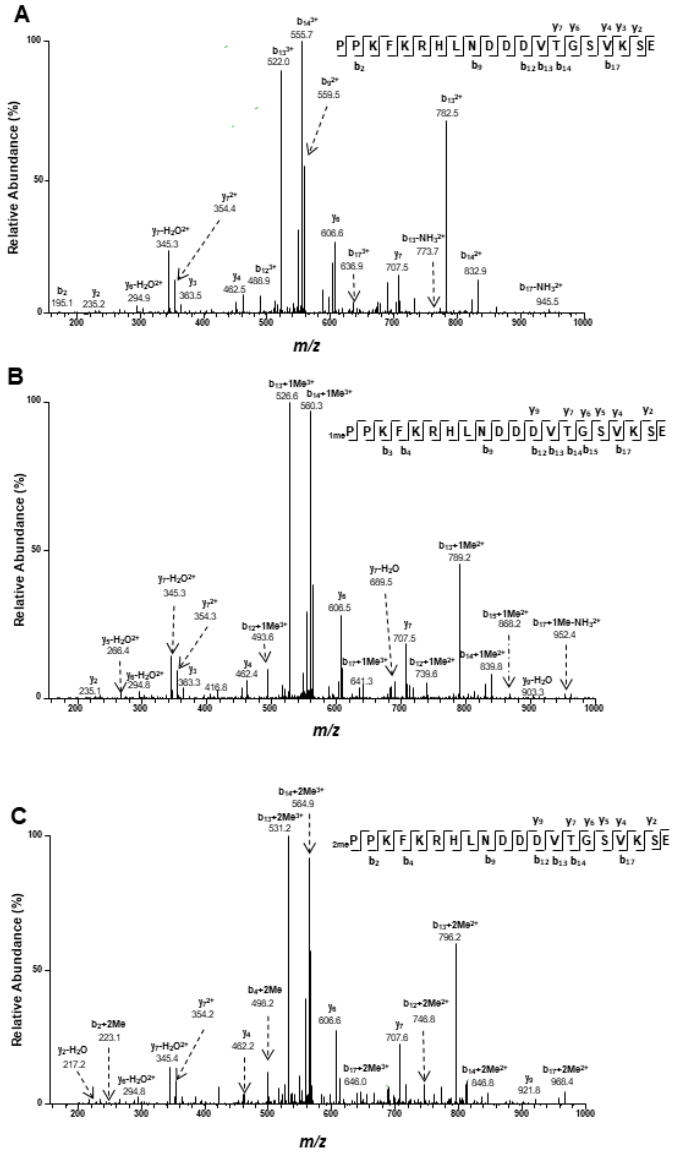


Figure 3. 3. NTMT1 is the methyltransferase responsible for VAMP4 α-N-terminal methylation.

'Ultra-zoom' scan ESI-MS showing the $[M+4H]^{4+}$ ions of the Glu-C-produced N-terminal peptide of VAMP4 isolated from HEK293T cells (a) and the isogenic *NTMT1*$^{-/-}$ cells (b), and VAMP-K4Q mutant isolated from HEK293T cells (c).

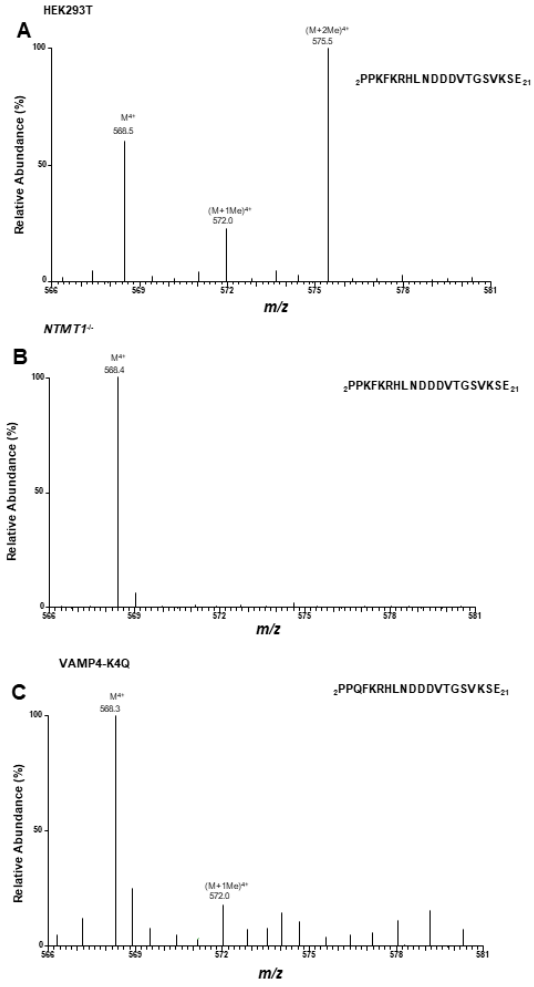
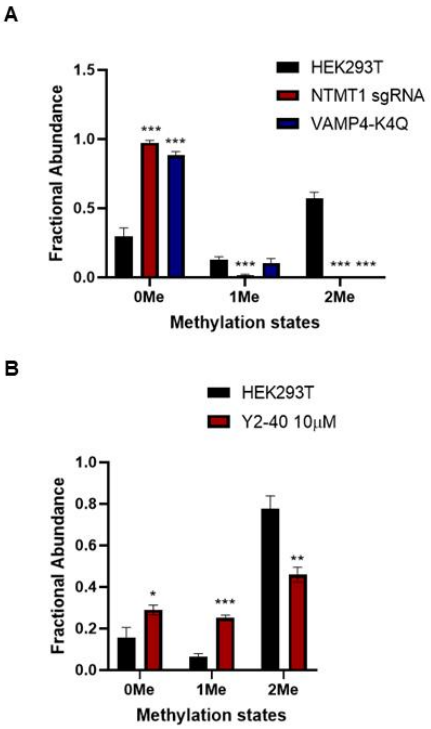


Figure 3. 4. VAMP4-K4Q mutant, genetic ablation of NTMT1, and NTMT1 inhibition decrease or abolish N-terminal Methylation of VAMP4.

Quantification of the effect of genetic ablation of NTMT1 and mutation of 4th lysine of on α-N-methylation of VAMP4 (a). Effect of NTMT1 small molecular inhibitor treatment on α-N-methylation of VAMP4. The methylation levels were quantified based on the relative abundances of the $[M+3H]^{3+}$ ions of the unmethylated and the mono-, di- and tri-methylated forms of the N-terminal peptide of MRG15. The *p* values were calculated using unpaired, two-tailed *t*-test: "ns", $p > 0.05$; "∗", $0.01 \leq p < 0.05$; "∗∗", $0.001 \leq p < 0.01$; "∗∗∗", $p < 0.001$.



Chapter 4: Discovery of Novel Histidine Methylated Proteins through Analysis of Publicly Available Datasets

1. Introduction

Post-translational modifications (PTMs) constitute an evolutionarily conserved mechanism for regulating the structure and functions of proteins. First discovered decades ago, histidine methylation represents one of several types of protein methylation (1). SETD3 was the first protein discovered to methylate histidine residues in actin (2,3). In addition, HPM1 was recently identified as a histidine methyltransferase in *Saccharomyces cerevisiae*, suggesting that its human ortholog METTL18 may also be a histidine methyltransferase (4). Since the identification of these methyltransferases, several unique protein substrates have also been discovered, including myosin heavy chain, myosin light chain kinase 2 (MLCK2), and 60S ribosomal protein L3 (RPL3) (4–6). Recent research have also identified METTL9 as the enzyme responsible for 1-methyl-histidine on several substrates, including S100A9 and DnaJ Heat Shock Protein Family (Hsp40) Member B12 (DNAJB12) (7).Although several targets have been revealed, recent studies suggest that the number of histidine methylated proteins may be far more expansive than what we currently have evidence for (8).

The SETD3-mediated histidine methylation has been implicated in affecting a variety of physiological systems and cellular mechanisms. For instance, inhibition of SETD3 expression has been shown to suppress muscle cell differentiation in mice (9). hSETD3 has also been identified as a PCNA-binding protein, which suggests a role of this methyltransferase in DNA replication/repair (10). Beyond this, recent research has illustrated SETD3's potential as a biomarker of prognosis for triple-negative and p53 mutant breast cancer patients (11). These results suggest that histidine methylation impacts many critical cellular functions and the discovery of additional substrates may

provide a better understanding about the biological functions of this important type of post-translational modification.

One source of discovery has arrived by high-throughput mass spectrometry-based techniques, which have been employed for proteomic analysis in recent years and contributed to publicly available datasets. These datasets acquired from discovery-based proteomic techniques, such as data-dependent acquisition (DDA), can be reexamined for many different projects, allowing for optimized research efficiency. This has led to the development of databases designed to store and curate individual datasets for the global exchange of metadata (12,13).

In this study, we analyzed these publicly available data sets using MaxQuant. We discovered 33 unique instances of histidine methylation in 26 unique proteins. We further confirmed the methylation of histidine 183 in RBM22.

2. Materials and Methods

2.1 Plasmid Preparation

The coding sequence of the human *RBM22* was amplified from a cDNA library prepared from HEK293T cells and cloned into the pRK7 plasmid between the XbaI and BamHI restriction sites. The cDNA was fused with three tandem repeats of Flag epitope tag (DYKDDDDK) at the carboxyl terminus.

2.2 LC-MS/MS Identification of Novel Histidine Methylated Proteins

HEK293T cells were cultured in Dulbecco's modified eagle medium (DMEM) supplemented with 10% fetal bovine serum (Invitrogen, Carlsbad, CA) and penicillin (100 IU/mL). The cells were maintained at 37°C in a humidified environment containing 5% CO_2. These cells were cultured in 6-well plates containing antibiotic-free DMEM media at a density of 5×10^6 cells/well. Plasmid

for ectopic expression of FLAG-tagged RBM22 (1 μg) was transfected into cells using Lipofectamine 2000 (Life Technologies, Carlsbad, CA). After a 48-hr incubation, the cells were harvested and lysed in CelLytic M lysis reagent (Sigma-Aldrich, St. Louis, MO), supplemented with a protease inhibitor cocktail (Sigma-Aldrich). C-terminally Flag-tagged RBM22 was isolated from the resultant lysate by affinity purification using anti-FLAG M2 beads (Sigma-Aldrich) and digested with trypsin (NEB) at a protein/enzyme ratio of 50:1.

The resulting tryptic digestion mixture was analyzed on a Q Exactive Plus quadrupole-Orbitrap mass spectrometer coupled with an UltiMate UHPLC system (Thermo Scientific, San Jose, CA). Samples were automatically loaded at 3 μL/min onto a pre-column (150 μm i.d. and 3.5 cm in length) packed with ReproSil-Pur 120 C18-AQ stationary-phase material (5 μm in particle size, 120 Å in pore size, Dr. Maisch). The precolumn was connected to a 20-cm fused-silica analytical column (PicoTip Emitter, New Objective, 75 μm i.d.) packed with 3 μm C18 beads (ReproSil-Pur 120 C18-AQ, Dr. Maisch). The peptides were then resolved using a 180-min gradient of 2-45% acetonitrile in 0.1% formic acid, and the flow rate was maintained at 300 nL/min.

The mass spectrometer was operated in a data-dependent acquisition mode. Full-scan mass spectra were acquired in the range of *m/z* 350-1500 using the Orbitrap analyzer at a resolution of 70,000 at *m/z* 200. Up to 25 most abundant ions found in MS with a charge state of 2 or above were sequentially isolated and collisionally activated in the HCD cell with a normalized collision energy of 28 to yield MS/MS.

2.3 MaxQuant Analysis

We first downloaded the LC-MS/MS data that were acquired in the DDA mode and made publicly available by other researchers (14–16). These data, along with the in-house generated DDA data for the whole-cell protein lysate of HEK293T cells, were searched against the *Homo*

Sapiens IPI protein database (version 3.68) or the Uniprot protein FASTA (UP000005640) using MaxQuant (17). The maximum number of miss-cleavages for trypsin and GluC was two per peptide. Cysteine carbamidomethylation was set as a fixed modification, while mono- and di-methylation of histidine were set as variable modifications. The tolerances in mass accuracy were 25 ppm and 0.6 Da for MS and MS/MS, respectively. Identified peptides with posterior error probability (PEP) scores being less than 0.01 were considered significant. Unique proteins were further subjected to Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis. All data were analyzed manually.

3. Results and Discussion

3.1 High-throughput Identification of Novel Histidine Methylated Proteins

By searching publicly available and in-house generated proteomic data with MaxQuant, we discovered 33 instances of histidine methylation from 26 distinct proteins. All targets met a PEP score threshold of less than 0.01, with a mass deviation of less than 20 ppm. Figure 1 shows the MS/MS of a tryptic peptide derived from RBM22 with H183 being mono-methylated, and a tryptic peptide from ASPH with H725 being mono- and di-methylated (Figure 4.1a-c). The histidine methylation of RBM22 is unique, as the methylated histidine is part of a conserved zinc finger and involved in Zn^{2+} binding (18). RBM22 is central to the U2–U6 snRNA complex's catalytic core, promoting the active conformation of this structure (19,20). Further, we have identified Zinc Finger CCCH-Type Containing 8 (ZC3H8) and Splicing factor U2AF 35 kDa subunit-like protein (U2AF1L5) to exhibit histidine methylation on their zinc finger motif's, further suggestion a role for this PTM in regulating zinc finger function. In addition, ASPH, is an ER membrane-anchored 2-oxoglutarate oxygenase (2OG) whose C-terminus contains the primary catalytic domain for hydroxylating asparaginyl and aspartyl residues on epidermal growth factor-like domains (EGFDs) (21). ASPH binding with iron as a cofactor in a unique mechanism involving only two ligand-

binding sites (His679/His725) (22). What is more, ASPH has been shown to promote the Notch signaling pathway, through binding several EGFD containing ligands. In breast cancer, the promotion of the Notch pathway initiated MMP/ADAM-mediated exosomal synthesis and release, contributing to the degradation and subsequent breakdown of the extracellular matrix, leading to infiltration and metastasis (23). Further, ASPH is overexpressed in 70-90% of solid tumor cancers, with documented roles of regulating cancer cell proliferation and metastasis in many types of cancer (24). Previous data has shown that histidine methylation can likely modulate binding of cofactors (25). Hence, it will be important to examine how this methylation modulates the iron binding and enzymatic activity of ASPH. Along this line, it is worth noting that H73 methylation in actin impedes ATP nucleotide exchange, likely through affecting ATP's ability in binding to actin (3).

Further analysis of the gene ontology of the 26 distinct methylated proteins revealed significant involvement in poly(a) RNA binding, as well as U6 snRNA binding (Figure 4.1d) (26). This is further supported by KEGG pathway analysis, revealing the role of several of these proteins in regulating and contributing to the spliceosome (Figure 4.1e) (27).

3.2 Identification of Histidine Methylation in RBM22

Noting that several of our novel histidine methylated proteins have a role in the spliceosome and the importance of RBM22 in activating the catalytic core of the spliceosome complex, we decided to collect additional evidence for the histidine methylation on this protein. For this purpose, we constructed a plasmid to express the cDNA of RBM22 fused with three tandem repeats of FLAG-tag on its C-terminus. We then enriched this protein from HEK293T cells, digested the lysate with trypsin, and subjected the resultant peptide mixture to LC-MS and LC-MS/MS analysis. We discovered that H183 in RBM22 was predominantly monomethylated, where the signal intensity

for H183-containing monomethylated peptide is much stronger than that for the unmodified counterpart (Figure 4.2a-b).

4. Conclusions

In this study, we analyzed publicly available DDA datasets with MaxQuant to identify novel histidine methylation instances. Using this method, we discovered 33 instances of histidine methylation across 26 distinct proteins. We further confirmed the methylation of histidine 183 in RBM22 through LC-MS and MS/MS analyses of ectopically expressed RBM22 protein isolated from HEK293T cells. Further research will focus on discovering how this methylation modulates the function of RBM22. In addition, we will seek to confirm the histidine methylation in ASPH. Given the importance of iron binding as a cofactor of this protein, methylation of one of the two histidine responsible for this likely affects its ability to function, and thus its role in cancer.

References

1. Laki K, Maruyama K, Kominz DR. Evidence for the interaction between tropomyosin and actin. Arch Biochem Biophys. 1962 Aug;98(2):323–330.

2. Kwiatkowski S, Seliga AK, Vertommen D, Terreri M, Ishikawa T, Grabowska I, et al. SETD3 protein is the actin-specific histidine N-methyltransferase. Elife. 2018 Dec 11;7.

3. Wilkinson AW, Diep J, Dai S, Liu S, Ooi YS, Song D, et al. SETD3 is an actin histidine methyltransferase that prevents primary dystocia. Nature. 2019;565(7739):372–376.

4. Webb KJ, Zurita-Lopez CI, Al-Hadid Q, Laganowsky A, Young BD, Lipson RS, et al. A novel 3-methylhistidine modification of yeast ribosomal protein Rpl3 is dependent upon the YIL110W methyltransferase. J Biol Chem. 2010 Nov 26;285(48):37598–37606.

5. Elzinga M, Collins JH. Amino acid sequence of a myosin fragment that contains SH-1, SH-2, and Ntau-methylhistidine. Proc Natl Acad Sci USA. 1977 Oct;74(10):4281–4284.

6. Meyer HE, Mayr GW. N pi-methylhistidine in myosin-light-chain kinase. Biol Chem Hoppe Seyler. 1987 Dec;368(12):1607–1611.

7. Davydova E, Shimazu T, Schuhmacher MK, Jakobsson ME, Willemen HLDM, Liu T, et al. The methyltransferase METTL9 mediates pervasive 1-methylhistidine modification in mammalian proteomes. Nat Commun. 2021 Feb 9;12(1):891.

8. Ning Z, Star AT, Mierzwa A, Lanouette S, Mayne J, Couture J-F, et al. A charge-suppressing strategy for probing protein methylation. Chem Commun. 2016 Apr 7;52(31):5474–5477.

9. Zhao M-J, Xie J, Shu W-J, Wang H-Y, Bi J, Jiang W, et al. MiR-15b and miR-322 inhibit SETD3 expression to repress muscle cell differentiation. Cell Death Dis. 2019 Feb 22;10(3):183.

10. Cooper SE, Hodimont E, Green CM. A fluorescent bimolecular complementation screen reveals MAF1, RNF7 and SETD3 as PCNA-associated proteins in human cells. Cell Cycle. 2015 Aug 3;14(15):2509–2519.

11. Shu W-J, Du H-N. The methyltransferase SETD3-mediated histidine methylation: Biological functions and potential implications in cancers. Biochim Biophys Acta Rev Cancer. 2021 Jan;1875(1):188465.

12. Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. Nucleic Acids Res. 2019 Jan 8;47(D1):D442–D450.

13. Choi M, Carver J, Chiva C, Tzouros M, Huang T, Tsai T-H, et al. MassIVE.quant: a community resource of quantitative mass spectrometry-based proteomics datasets. Nat Methods. 2020 Sep 14;17(10):981–984.

14. Nagaraj N, Wisniewski JR, Geiger T, Cox J, Kircher M, Kelso J, et al. Deep proteome and transcriptome mapping of a human cancer cell line. Mol Syst Biol. 2011 Nov 8;7:548.

15. Yeom J, Ju S, Choi Y, Paek E, Lee C. Comprehensive analysis of human protein N-termini enables assessment of various protein forms. Sci Rep. 2017 Jul 26;7(1):6599.

16. Qi TF, Guo L, Huang M, Li L, Miao W, Wang Y. Discovery of TBC1D7 as a potential driver for melanoma cell invasion. Proteomics. 2020 Jul 2;20(14):e1900347.

17. Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol. 2008 Dec;26(12):1367–1372.

18. Montaville P, Dai Y, Cheung CY, Giller K, Becker S, Michalak M, et al. Nuclear translocation of the calcium-binding protein ALG-2 induced by the RNA-binding protein RBM22. Biochim Biophys Acta. 2006 Nov;1763(11):1335–1343.

19. Rasche N, Dybkov O, Schmitzová J, Akyildiz B, Fabrizio P, Lührmann R. Cwc2 and its human homologue RBM22 promote an active conformation of the spliceosome catalytic centre. EMBO J. 2012 Mar 21;31(6):1591–1604.

20. Chu H, Perea W, Greenbaum NL. Role of the central junction in folding topology of the protein-free human U2-U6 snRNA complex. RNA. 2020 Mar 27;26(7):836–850.

21. Stenflo J, Holme E, Lindstedt S, Chandramouli N, Huang LH, Tam JP, et al. Hydroxylation of aspartic acid in domains homologous to the epidermal growth factor precursor is catalyzed by a 2-oxoglutarate-dependent dioxygenase. Proc Natl Acad Sci USA. 1989 Jan;86(2):444–447.

22. Pfeffer I, Brewitz L, Krojer T, Jensen SA, Kochan GT, Kershaw NJ, et al. Aspartate/asparagine-β-hydroxylase crystal structures reveal an unexpected epidermal growth factor-like domain substrate disulfide pattern. Nat Commun. 2019 Oct 28;10(1):4910.

23. Zheng W, Wang X, Hu J, Bai B, Zhu H. Diverse molecular functions of aspartate β-hydroxylase in cancer (Review). Oncol Rep. 2020 Dec;44(6):2364–2372.

24. Kanwal M, Smahel M, Olsen M, Smahelova J, Tachezy R. Aspartate β-hydroxylase as a target for cancer therapy. J Exp Clin Cancer Res. 2020 Aug 18;39(1):163.

25. Nyman T, Schüler H, Korenbaum E, Schutt CE, Karlsson R, Lindberg U. The role of MeH73 in actin polymerization and ATP hydrolysis. J Mol Biol. 2002 Apr 5;317(4):577–589.

26. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene Ontology: tool for the unification of biology. Nat Genet. 2000 May;25(1):25–29.

27. Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic Acids Res. 2014 Jan;42(Database issue):D199–205.

Table 3.1. List of Histidine Methylated Proteins

Gene	Acc #	Sequence	Modifications	Mass Erro	PEP
ACTB	P60709	EHPVLLTEAPLNPKANRE	Dimethyl (H)	-0.0382	5.07E-05
ADGRE5	P48960	LNSPILFAFSHLE	Dimethyl (H)	0.022592	2.96E-05
ASPH	Q12797	GKVLIFDDSFEHE	Dimethyl (H)	-0.52749	1.49E-06
ASPH	Q12797	GKVLIFDDSFEHE	Methyl(H)	1.7484	7.91E-06
CDC42BPB	Q9Y5S2	QGGRGAGATLEHQQEISK	Methyl(H)	-0.2963	2.02E-22
CPS1	P31327	LSLERILDIYHQE	Dimethyl (H)	0.43552	4.27E-06
DDX39B	Q13838	LLRAIVDCGFEHPSEVQHE	Dimethyl (H)	1.2564	0.000678
DHRS7	Q9Y394	KDILVLPLDLTDTGSHE	Dimethyl (H)	0.91884	3.28E-07
GAPVD1	Q14C86	SLLAMFDPLSSHE	Dimethyl (H)	1.768	1.13E-05
HNRNPF	P52597	NDIYNFFSPLNPVRVHIE	Methyl(H)	-0.00545	9.21E-05
HNRNPF	P52597	NDIYNFFSPLNPVRVHIE	Dimethyl (H)	1.7377	0.000102
HSD17B4	P51659	YVAPLVLWLCHE	Dimethyl (H)	0.83701	0.000319
IPO9	Q96P70	LIQQFLQATVSGLHE	Dimethyl (H)	0.31508	0.000424
ITPR3	Q14573	IKCTSLLPLEDVVSVVTHE	Dimethyl (H)	0.70017	0.000497
KIF11	P52732	LDGFLSILCNNLHE	Dimethyl (H)	-0.55689	0.000486
MEPCE	Q7L2J0	GQHHQQQQAAGGSESHPVPPT	Methyl(H)	-0.20483	2.38E-12
MYH9	P35579	MEDLMSSKDDVGKSVHE	Dimethyl (H)	-0.58605	1.34E-05
NARS1	O43776	DLVCDVVDRILKSPAGSIVHE	Dimethyl (H)	-0.20787	0.000589
POTEI	P0CG38	HGIITNWDDME	Oxidation (M),Me	-1.7418	0.000153
PRDX1	Q06830	TLRLVQAFQFTDKHGE	Dimethyl (H)	-0.23496	4.25E-09
PSMD2	Q13200	LDIMEPKVPDDIYKTHLE	Dimethyl (H)	0.22338	3.35E-07
PSMD2	Q13200	LDIMEPKVPDDIYKTHLE	Oxidation (M),Din	0.27513	0.000659
RBM22	Q9NW64	HEKPTDPDDPLADQNIK	Methyl(H)	0.16502	1.06E-26
SERBP1	Q8NC51	DKRGGSGSHNWGTVKDE	Methyl(H)	-0.00856	3.42E-05
TOP1	P11387	VATFFAKMLDHE	Dimethyl (H)	-0.3462	0.00045
U2AF1L5	P0DN76	LHNKPTFSQTIALLNIYRNPQNS	Methyl(H)	-0.18801	0
VCL	P18206	SILEPVAQQISHLVIMHE	Dimethyl (H)	0.6458	4.3E-05
XRCC5	P13010	SKIQPGSQQADFLDALIVSMDV	Methyl(H)	2.3365	3.82E-15
XRCC5	P13010	SKIQPGSQQADFLDALIVSMDV	Dimethyl (H)	-0.87054	2.9E-06
ZC3H8	Q8N5P1	FSHAPLTPETQELLAK	Methyl(H)	0.24168	1.52E-21
ACTB	P60709	HGIVTNWDDMEKIWHHTFYNI	Oxidation (M),Me	1.1379	0.000115
U2AF1L5	P0DN76	LHNKPTFSQTILIQNIYR	Methyl(H)	0.036835	0
ACTB	P60709	YPIEHGIVTNWDDMEK	Methyl(H)	-0.07499	3.8E-182

Figure 4.1. Representative ESI-MS/MS of Methylated Histidine-containing Peptides, and Results from Gene Ontology and KEGG Pathway Analysis.

MS/MS from publicly available DDA files of the $[M+2H]^{2+}$ ion of a tryptic peptide, i.e. $_{183}$HEKPTDPDDPLADQNIK$_{199}$, derived from RBM22 containing monomethylated H183 (a), and the $[M+3H]^{3+}$ ions of a tryptic peptide, i.e. $_{714}$GKVLIFDDSFEHE$_{726}$, from ASPH with H725 being mono- (b) and di-methylated (c). Shown in (d) and (e) are results from gene ontology (d) and KEGG Pathway (e) analysis of novel histidine methylated proteins.

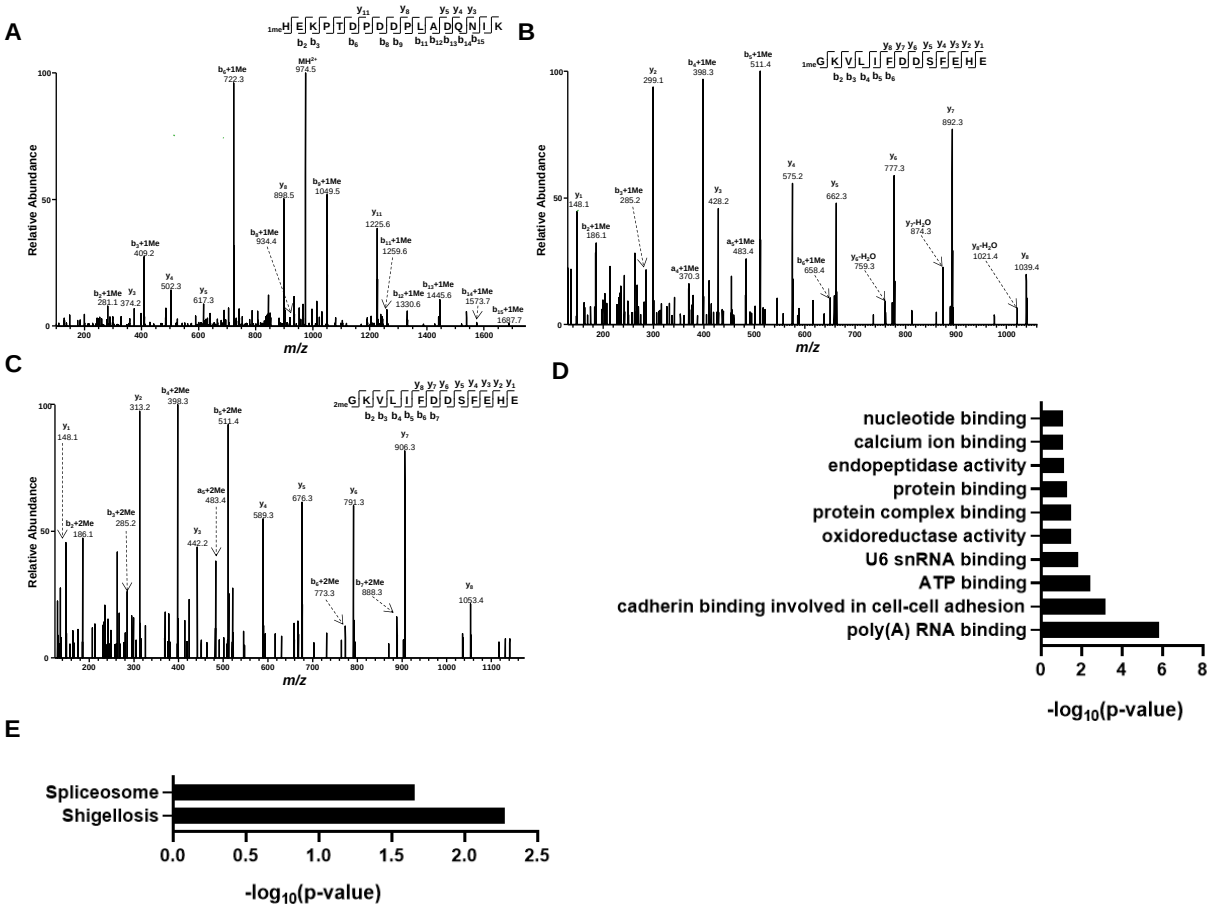
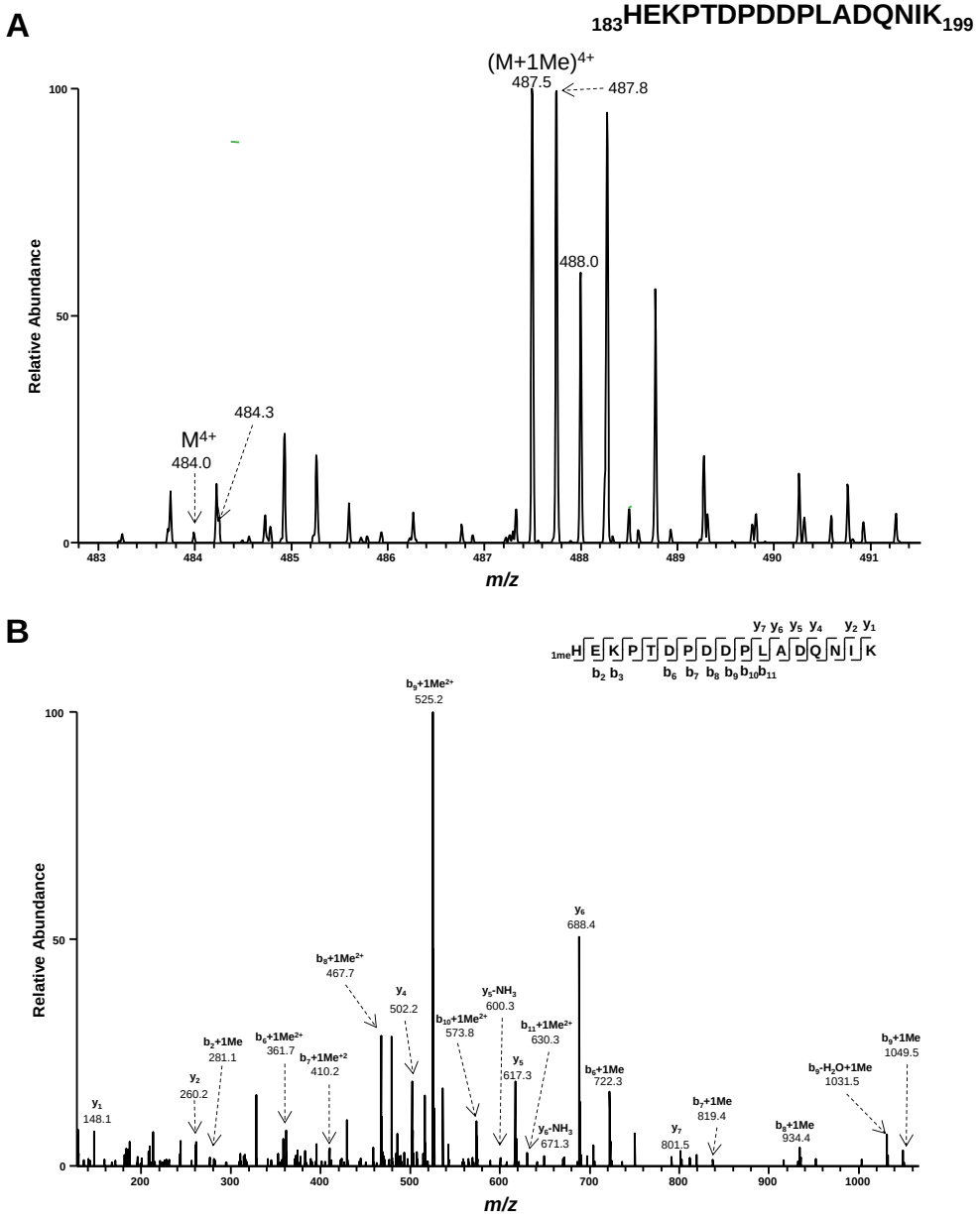


Figure 4.2. H183 in RBM22 is monomethylated in HEK293T cells.

Positive-ion ESI-MS displaying the monomethylation of H183 in RBM22 (a), and MS/MS of the $[M+2H]^{2+}$ ion of the monomethylated H183-containing peptide $_{183}$HEKPTDPDDPLADQNIK$_{199}$ (b).



Chapter 5: Targeted Proteomic Analysis Revealed Kinome Reprogramming during Acquisition of Radioresistance in Breast Cancer Cells

1. Introduction

Breast cancer is the most frequent malignancy in women (1), and the global incidence of breast cancer has been rising at an annual rate of 3.1%, from 641,000 cases in 1980 to over 1.6 million in 2010 (2), and then to almost 2 million cases in 2017 (3). Breast cancer ranks as the fifth most common cause of cancer death in women and it is one of the most expensive malignancies for treatment (4).

Breast cancer treatment is multidisciplinary. The majority of women with early-stage breast cancer are candidates for breast-conserving surgery with radiotherapy or mastectomy (5). Radiotherapy is a treatment method wherein high-energy γ rays introduce DNA damage and kill rapidly proliferating cancer cells. It has been evolving to improve precision in targeting the diseased tissue while minimizing delivery of radiation to surrounding vital organs such as the heart and lungs (6). Despite the advances in radiation therapy, uniform doses of radiation is typically delivered without taking into consideration the differences across breast cancer subtypes (6). As a result, the benefits of this therapy may not be uniform across all patients owing to individual variations in sensitivity toward radiation therapy (6). Moreover, some patients may experience recurrences and develop radiation resistance, which confer poor prognosis and diminished quality of life in these patients (6). Thus, radioresistance presents a major obstacle to current breast cancer radiotherapy (7), and the discovery of novel molecular targets that modulate

radiosensitivity may provide a venue for improving therapeutic efficacy and for developing personalized radiotherapy.

Recent studies showed that aberrant expression and activity of kinases are associated with the acquisition of resistance towards breast cancer radiotherapy. For example, abnormal expression of maternal embryonic leucine zipper kinase (MELK) (8) and ataxia-telangiectasia mutated (ATM) (9) are known to be accompanied with breast cancer that are refractory toward radiation treatment. In addition, Guo et al. (10) reported the application of a targeted proteomic method, relying on the use of desthiobiotin-ATP acyl phosphate probe together with liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis in multiple-reaction monitoring (MRM) mode, to reveal the roles of kinases in the acquisition of resistance toward radiation in breast cancer. However, the results from the ATP-affinity probe enrichment can be affected by both the protein expression level and activity of the kinase, rendering it difficult to differentiate their individual contributions. In addition, the earlier version of the kinome profiling method only allowed for the quantifications of 120 unique kinases (10); thus, some low-abundance kinases associated with radioresistance may not be detected. Therefore, there is a need for in-depth profiling of differential expression of protein kinases that are associated with the development of radioresistance in breast cancer.

In this study, we employed our previously developed parallel-reaction monitoring (PRM)-based LC-MS/MS method (11-16) to examine the differential expression of kinase proteins in two pairs of breast cancer cell lines, i.e. the MDA-MB-231 and MCF-7 breast cancer cells and their corresponding radioresistant sub-clones. Our results led to the

quantifications of the relative expression levels of 300 and 281 kinases in the MDA-MB-231 and MCF-7 pairs of breast cancer cells, respectively. We also revealed the role of one of the differentially expressed kinases, TAF9, in promoting acquired radiation resistance in breast cancer cells.

2. Materials and Methods

2.1.Cell culture

The radioresistant clone (clone 6, referred as C6) of MCF-7 cells were generated by treating the parental cells with a total dose of 60 Gy of γ rays and further exposed to an additional term of 30 fractions of γ rays (17,18). Similarly, MDA-MB-231 cells were exposed to a total dose of 30 Gy of ionizing radiation to generate the radioresistant clone (clone 5, referred as C5) (18,19).

MCF-7 and MDA-MB-231 cells, as well as their paired radioresistant lines were cultured in Dulbecco's modified eagle medium (DMEM). The culture medium was supplemented with 10% fetal bovine serum (Invitrogen, Carlsbad, CA) and penicillin (100 IU/mL). The cells were maintained at 37°C in a humidified atmosphere containing 5% CO_2. Approximately 5×10^6 cells were harvested, washed twice with cold PBS, and lysed by incubating on ice for 30 min with CelLytic M cell lysis reagent (Sigma) containing 1% protease inhibitor cocktail. The cell lysates were centrifuged at 9,000g at 4°C for 20 min, and the resulting supernatants were collected. For SILAC labeling experiments, the cells

were cultured in SILAC medium containing unlabeled lysine and arginine, or [$^{13}C_6$, $^{15}N_2$]-lysine and [$^{13}C_6$]-arginine, for at least five cell doublings.

2.2.Tryptic digestion of whole-cell protein lysates

The whole cell lysates prepared from MCF-7 or MDA-MB-231 breast cancer cells and their radioresistant counterparts were combined at 1:1 ratio (by mass, determined by Bradford assay), incubated with 8 M urea for protein denaturation, and subsequently with dithiothreitol and iodoacetamide for cysteine reduction and alkylation, respectively. The proteins were then digested, at 37°C overnight, with modified MS-grade trypsin (Pierce) at an enzyme/substrate ratio of 1:100 in 50 mM NH_4HCO_3 (pH 8.5). The resulting peptide mixture was dried in a Speed-vac, desalted with OMIX C18 pipette tips (Agilent Technologies), and analyzed by LC-MS/MS in the PRM mode.

2.3.LC-PRM Analysis

All LC-PRM experiments were carried out on a Q Exactive Plus quadrupole-Orbitrap mass spectrometer coupled with an EASY-nLC 1000 system (Thermo Fishier Scientific). The linear predictor of empirical retention time (RT) from normalized RT (iRT) (20) for targeted peptides of kinases was determined by the linear regression of RTs of tryptic peptides of BSA obtained for the current chromatography setup (21-24). This RT-iRT linear relationship was re-defined between every eight LC-MS/MS runs by injecting a tryptic digestion mixture of BSA. The MS/MS for targeted precursor ions were monitored in scheduled PRM mode with an 8-min retention time window. The resulting LC-MS/MS data were processed with Skyline (version 3.5) to plot the extracted-ion chromatograms

and to integrate peak areas (25). The detailed experimental conditions are provided in the Supporting Information.

2.4.Western blot

The detailed sample preparation procedures for Western blot analysis are described in the Supporting Information. Human CHK1 (Cell Signaling Technology, 2360S, 1:2000 dilution) and TAF9 (Proteintech, 10544-1-AP, 1:2500 dilution) antibodies were employed as the primary antibodies. Horseradish peroxidase-conjugated anti-rabbit IgG and IRDye® 680LT Goat anti-Mouse IgG were used as secondary antibodies. Membranes were also probed with anti-Actin antibody (Cell Signaling #4967, 1:10000 dilution) or anti-GAPDH antibody (Santa Cruz Biotechnology, sc-32233, 1:5000 dilution) to verify equal protein loading.

2.5.TCGA, UALCAN and GEPIA data analysis

Clinical data and mRNA expression profiles for the Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) (26) cohort were obtained from cBioPortal (http:// www.cbioportal.org/data_sets.jsp) (27). A total of 1137 patients in the METABRIC cohort with radiotherapy was applied for TCGA analysis. Median mRNA expressions of genes encoding kinases was further applied as cut-off for low/ high expression (568 and 569 patients, respectively). p-values were generated using log-rank (Mantel-Cox) test, and a p-value of less than 0.05 is considered significant.

Box plots for kinase mRNA expression in normal and breast cancer cell lines were generated from the gene expression data for cell lines derived from 114 normal breast

tissues and 1097 breast cancer tissues using UALCAN (http://ualcan.path.uab.edu/index.html) (28). The *t*-test was performed using a PERL script with Comprehensive Perl Archive Network (CPAN) module "Statistics::TTest" (http://search.cpan.org/~yunfang/Statistics-TTest-1.1.0/TTest.pm) (28). Correlation analysis of gene expression was conducted using GEPIA (29), where the expression of targeted genes in 1085 breast cancer tissues was analyzed. Pearson's correlation coefficient was employed to calculate distance metric.

2.6.Plasmid construction and clonogenic survival assay

The coding sequence of human *TAF9* gene was amplified from a cDNA library prepared from mRNAs isolated from HEK293T cells and cloned into the pRK7 plasmid between the XbaI and BamHI restriction sites. The resulting pRK7-TAF9 plasmid was confirmed by Sanger sequencing.

MCF-7 and MDA-MB-231 cells were transfected with a control pRK7 vector or pRK7-TAF9 using Lipofectamine 2000 (Invitrogen) following the manufacturer's protocol. At 24 hr following the transfection, the cells were seeded in triplicate in 6-well plates at a concentration of 900 cells per well and incubated overnight. On the following day, the cells were exposed with 0, 1, 2, 3.5, or 5 Gy of X-rays, delivered by a Rad Source RS-2000 cabinet irradiator (Rad Source Technologies, Buford, GA, United States). Following a 10-day incubation, cell colonies were fixed with a 6% glutaraldehyde solution for 1 hr, and subsequently stained with 0.5% crystal violet for 30 min. All colonies with at least 50 cells were counted.

3. Results and Discussion

3.1. Quantitative assessment of differential expression of kinases in two matched pairs of radioresistant and parental breast cancer cells

To explore the potential functions of kinases in acquired radioresistance, we employed our recently developed PRM-based targeted proteomic method (11), in combination with stable isotope labeling by amino acids in cell culture (SILAC) (30), to examine the differential expression of kinases in two pairs of breast cancer cells, i.e. parental (WT) MCF-7 and MDA-MB-231 cells, as well as their corresponding radioresistant C6 and C5 clones (Figure 5.1a).

The targeted proteomic analysis led to the quantifications of 281 and 300 unique kinases in the MCF-7 WT/C6 and MDA-MB-231 WT/C5 pairs of breast cancer cells, respectively (Figures 5.1b, 5.1c, 5.2 and Table 5.1). All 4-6 PRM transitions selected for each tryptic peptide derived from kinases displayed the same retention time and exhibited a dot product (dotp) value (31) of > 0.7. In addition, more than 90% of all the quantified kinase proteins appeared in both forward and reverse SILAC labeling experiments (Figures 5.1b and 5.1c, Table S1). Furthermore, the ratios of the quantified peptides obtained from forward and reverse SILAC labeling experiments showed a very good linear fit (Figure 5.1b and 5.1c). The results, therefore, revealed the robustness and good reproducibility of the PRM method.

Guo et al. (10) has reported the application of a targeted proteomic method, relying on the use of desthiobiotin-ATP probe together with LC-MS/MS analysis in MRM mode, to

reveal the roles of kinases in the acquisition of resistance toward radiation in MCF-7 breast cancer cells. We next compared the reported MRM and our PRM data to assess the perturbations in protein expression levels and ATP-binding affinities of kinases upon redioresistance acquisition.

The comparison between the two datasets facilitated the expression and ATP-binding affinity of 87 common kinases (Figure 5.3a, Table S2). Kinase ratios (C6/MCF-7) obtained from PRM and MRM analyses lacked apparent correlation (Figure 5.3b), suggesting the perturbed ATP-binding affinity of kinases during radioresistance acquisition. The comparison of the two datasets revealed the perturbations of the kinases such as CHK1 and CDK1 are due to protein expression alone, while the ATP-binding affinities of MAP4K4 and MAPK13 are induced profoundly during radioresistance acquisition (Figure 5.3b). This agrees with a recently research that the activity of MAP4K4 regulates the DNA double-strand breaks (DSBs) repair pathway (32). Along this line, we uncovered enhanced ATP-binding affinities of 10 kinases and diminished affinities of 15 kinases, which could be important in radioresistance acquisition in MCF-7 breast cancer cells (Table S2).

3.2. Kinases are altered upon the development of radiation resistance in breast cancer cells

The PRM method facilitated the quantifications of the relative expressions of 227 common kinases in the two pairs of breast cancer cells, which exhibited similar changes in kinome profiles upon the development of radioresistance (Figure 5.4a and 5.4b). In particular, 45 kinases were commonly altered by at least 1.5-fold in both lines of radioresistant breast cancer cells relative to the corresponding parental lines (Figure 5.4c,

Table 5.1). Moreover, several of these kinases were previously characterized to confer radiation resistance in cancer cells. For example, CHK1, which is involved in regulating DNA-damage-checkpoint responses, is known to be associated with the acquisition of radioresistance in cancer cells (33,34). In addition, up-regulation of HK2 was found to elicit radioresistance through elevating glycolysis rate (35). Moreover, DDR1, a receptor tyrosine kinase, was able to regulate autophagy and ultimately modulate radiosensitivity in glioblastoma (36).

To understand further the molecular mechanisms through which the differentially expressed kinases may contribute to radioresistance, we conducted Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses (Figure 4d). Interestingly, RNA polymerases stand out in both GO and KEGG pathway analyses, while 7 subunits of RNA polymerase II complex were commonly up-regulated in the two radioresistant lines over the corresponding parental lines. Furthermore, analysis of The Cancer Genome Atlas (TCGA) data among those breast cancer patients with radiotherapy revealed that increased expression of one out of seven RNA polymerase II subunits, i.e. POLR2H, was positively correlated with poorer survival of breast cancer patients (Figure 5.5a, Table S3). Therefore, our results suggest that RNA polymerase II may promote the acquisition of radioresistance. However, POLR2H is essential for the survival of most cultured cancer cells (37), rendering it difficult to further validate its role in radioresistance.

3.3.TAF9 drives acquired radioresistance in breast cancer cells

To uncover those differentially expressed kinases that modulate acquired radioresistance in breast cancer cells, we compared our kinome quantification results with the mRNA

levels of kinases in primary breast tumor tissues and normal tissues (Table S3). We found that the differential mRNA expressions of 25 out of 45 perturbed kinases were positively correlated with their protein expression, including CHK1 and TAF9, while the expression of CHEK1 was positively correlated with poorer survival of breast cancer patients (Figure 5.5b). Consistent with the PRM data, our Western blot results revealed increased expression of CHK1 and TAF9 in the radioresistant breast cancer cells over the parental breast cancer cells (Figure 5.5c-e). These results also support the quantification accuracy of the PRM method.

We also observed higher mRNA expression levels of *CHEK1* and *TAF9* genes in the breast tumors compared to normal tissues (Figure 5.5f). Since CHK1 is known to be involved in DNA damage response signaling and correlated with radiosensitivity of cancer cells (33,34), we focused on examining the function of TAF9 in acquired radioresistance in breast cancer cells.

TAF9 is a component of the transcription factor IID (TFIID) complex, the TBP-free TAFII complex (TFTC), the PCAF histone acetyltransferase (HAT) complex and the STAGA transcription coactivator-HAT complex (38). Because TFIID or TFTC are essential for the regulation of RNA polymerase II-mediated transcription, TAF9 may modulate RNA polymerase II-catalyzed transcription. As noted above, the protein expression levels of seven out of eight quantified subunits of RNA polymerase II were elevated in both radioresistant lines of breast cancer cells. Thus, TAF9 may play a central role in the acquisition of radioresistance in breast cancer cells by regulating RNA polymerase II-based transcription.

To assess directly the role of TAF9 in the acquisition of radioresistance, we overexpressed *TAF9* gene in MCF-7 and MDA-MB-231 cells, and assessed its impact on the survival of breast cancer cells following exposure to X rays. Results from our clonogenic survival assay showed decreases in survival of MCF-7 and MDA-MB-231 cells upon ectopic overexpression of TAF9 compared to transfection with empty vector control, with the most pronounced decrease being observed for cells exposed with 2 Gy of X rays (Figure 5.6a and 5.6b).

As TAF9 binds directly to several transcription factors such as p53 (39) and VP16 (40), thereby regulating gene transcription, we also asked if TAF9 regulates the transcription of radioresistant-related genes. For this purpose, we compared the mRNA expression of TAF9 and drivers of radioresistance in 1085 breast cancer tissues. HK2 and DDR1 are two known drivers that we identified to be up-regulated in both radioresistant cell lines, as described above. The mRNA expression levels of these two genes are positively correlated with the mRNA level of TAF9 gene in breast cancer patients, suggesting the potential role of TAF9 in the up-regulation of those drivers for radiation resistance (Figure 6c).

4. Conclusions

In this study, we employed a recently developed PRM-based targeted proteomic method to examine the differential expression of kinase proteins in two pairs of radioresistant breast cancer cells versus the corresponding parental cells. With this method, we quantified 300 and 281 kinases in MDA-MB-231/C5 and MCF-7/C6 pairs of breast cancer cells, respectively. Among these kinases, 45 were commonly altered by at least 1.5-fold in both

pairs of breast cancer cell lines. We also validated that one of the differentially expressed kinases, TAF9, promotes the acquisition of radioresistance in cultured breast cancer cells. Moreover, gene correlation analysis suggested *TAF9*'s role in up-regulating the expression of other genes involved in radioresistance. Overall, our study revealed a new role of TAF9 in promoting the acquisition of radioresistance in breast cancer and uncovered a number of other kinases with potential functions in radioresistance.

References

1. Harbeck N, Penault-Llorca F, Cortes J, Gnant M, Houssami N, Poortmans P, *et al.* Breast cancer. Nat Rev Dis Primers **2019**;5:66
2. Bray F, Ferlay J, Laversanne M, Brewster DH, Gombe Mbalawa C, Kohler B, *et al.* Cancer incidence in five continents: inclusion criteria, highlights from volume X and the global status of cancer registration. Int J Cancer **2015**;137:2060-71
3. Li N, Deng Y, Zhou L, Tian T, Yang S, Wu Y, *et al.* Global burden of breast cancer and attributable risk factors in 195 countries and territories, from 1990 to 2017: results from the global burden of disease study 2017. J Hematol Oncol **2019**;12:140
4. Sullivan R, Peppercorn J, Sikora K, Zalcberg J, Meropol NJ, Amir E, *et al.* Delivering affordable cancer care in high-income countries. Lancet Oncol **2011**;12:933-80
5. Moo T-A, Sanford R, Dang C, Morrow M. Overview of breast cancer therapy. PET Clinics **2018**;13:339-54
6. Choi J, Yoon YN, Kim N, Park CS, Seol H, Park I-C, *et al.* Predicting radiation resistance in breast cancer with expression status of phosphorylated S6K1. Sci Rep **2020**;10:641
7. Yadav P, Shankar BS. Radio resistance in breast cancer cells is mediated through TGF-β signalling, hybrid epithelial-mesenchymal phenotype and cancer stem cells. Biomed Pharmacother **2019**;111:119-30
8. Speers C, Zhao SG, Kothari V, Santola A, Liu M, Wilder-Romans K, *et al.* Maternal embryonic leucine zipper kinase (MELK) as a novel mediator and biomarker of radioresistance in human breast cancer. Clin Cancer Res **2016**;22:5864
9. Bian L, Meng Y, Zhang M, Guo Z, Liu F, Zhang W, *et al.* ATM expression is elevated in established radiation-resistant breast cancer cells and improves DNA repair efficiency. Int J Biol Sci **2020**;16:1096-106
10. Guo L, Xiao Y, Fan M, Li JJ, Wang Y. Profiling global kinome signatures of the radioresistant MCF-7/C6 breast cancer cells using MRM-based targeted proteomics. J Proteome Res **2015**;14:193-201
11. Miao W, Guo L, Wang Y. Imatinib-induced changes in protein expression and ATP-binding affinities of kinases in chronic myelocytic leukemia cells. Anal Chem **2019**;91:3209-14
12. Miao W, Li L, Liu X, Qi TF, Guo L, Huang M, *et al.* A targeted quantitative proteomic method revealed a substantial reprogramming of kinome during melanoma metastasis. Sci Rep **2020**;10:2485
13. Miao W, Li L, Wang Y. High-throughput targeted quantitative analysis of the interaction between HSP90 and kinases. Anal Chem **2019**;91:11507-9
14. Miao W, Wang Y. Targeted quantitative kinome analysis identifies PRPS2 as a promoter for colorectal cancer metastasis. J Proteome Res **2019**;18:2279-86
15. Miao W, Wang Y. Quantitative interrogation of the human kinome perturbed by two BRAF inhibitors. J Proteome Res **2019**;18:2624-31
16. Miao W, Yuan J, Li L, Wang Y. Parallel-reaction monitoring-based proteome-wide profiling of differential kinase protein expression during prostate cancer metastasis in vitro. Anal Chem **2019**;91:9893-900

17. Ahmed KM, Dong S, Fan M, Li JJ. Nuclear factor-κB p65 inhibits mitogen-activated protein kinase signaling pathway in radioresistant breast cancer cells. Mol Cancer Res **2006**;4:945

18. Miao W, Fan M, Huang M, Li JJ, Wang Y. Targeted profiling of heat shock proteome in radioresistant breast cancer cells. Chem Res Toxicol **2019**;32:326-32

19. Cao N, Li S, Wang Z, Ahmed KM, Degnan ME, Fan M, *et al.* NF-κB-mediated HER2 overexpression in radiation-adaptive resistance. Radiat Res **2009**;171:9-21

20. Escher C, Reiter L, MacLean B, Ossola R, Herzog F, Chilton J, *et al.* Using iRT, a normalized retention time for more targeted measurement of peptides. Proteomics **2012**;12:1111-21

21. Miao W, Xiao Y, Guo L, Jiang X, Huang M, Wang Y. A high-throughput targeted proteomic approach for comprehensive profiling of methylglyoxal-induced perturbations of the human kinome. Anal Chem **2016**;88:9773-9

22. Miao W, Li L, Wang Y. A targeted proteomic approach for heat shock proteins reveals DNAJB4 as a suppressor for melanoma metastasis. Anal Chem **2018**;90:6835-42

23. Miao W, Li L, Wang Y. Identification of helicase proteins as clients for HSP90. Anal Chem **2018**;90:11751-5

24. Miao W, Li L, Zhao Y, Dai X, Chen X, Wang Y. HSP90 inhibitors stimulate DNAJB4 protein expression through a mechanism involving N6-methyladenosine. Nat Commun **2019**;10:3613

25. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, *et al.* Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics **2010**;26:966-8

26. Curtis C, Shah SP, Chin S-F, Turashvili G, Rueda OM, Dunning MJ, *et al.* The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature **2012**;486:346

27. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, *et al.* Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Science Signaling **2013**;6:pl1

28. Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi BVSK, *et al.* UALCAN: A portal for facilitating tumor subgroup gene expression and survival analyses. Neoplasia **2017**;19:649-58

29. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res **2017**;45:W98-W102

30. Ong S-E, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, *et al.* Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Molecular & Cellular Proteomics **2002**;1:376-86

31. de Graaf EL, Altelaar AF, van Breukelen B, Mohammed S, Heck AJ. Improving SRM assay development: a global comparison between triple quadrupole, ion trap, and higher energy CID peptide fragmentation spectra. J Proteome Res **2011**;10:4334-41

32. Deng M, Lin J, Nowsheen S, Liu T, Zhao Y, Villalta PW, *et al.* Extracellular matrix stiffness determines DNA repair efficiency and cellular sensitivity to genotoxic agents. Sci Adv **2020**;6:eabb2630

33. Biau J, Chautard E, Verrelle P, Dutreix M. Altering DNA repair to improve radiation therapy: specific and multiple pathway targeting. Front Oncol **2019**;9:1009

34. Wang W-J, Wu S-P, Liu J-B, Shi Y-S, Huang X, Zhang Q-B, *et al.* MYC regulation of CHK1 and CHK2 promotes radioresistance in a stem cell-like population of nasopharyngeal carcinoma cells. Cancer Res **2013**;73:1219

35. Zhong J-T, Zhou S-H. Warburg effect, hexokinase-II, and radioresistance of laryngeal carcinoma. Oncotarget **2017**;8:14133-46

36. Vehlow A, Klapproth E, Jin S, Hannen R, Hauswald M, Bartsch J-W, *et al.* Interaction of discoidin domain receptor 1 with a 14-3-3-beclin-1-Akt1 complex modulates glioblastoma therapy sensitivity. Cell Rep **2019**;26:3672-83.e7

37. Tsherniak A, Vazquez F, Montgomery PG, Weir BA, Kryukov G, Cowley GS, *et al.* Defining a cancer dependency map. Cell **2017**;170:564-76.e16

38. Frontini M, Soutoglou E, Argentini M, Bole-Feysot C, Jost B, Scheer E, *et al.* TAF9b (formerly TAF9L) is a bona fide TAF that has unique and overlapping roles with TAF9. Mol Cell Biol **2005**;25:4638-49

39. Uesugi M, Verdine GL. The α-helical FXXΦΦ motif in p53: TAF interaction and discrimination by MDM2. Proc Natl Acad Sci U S A **1999**;96:14801

40. Uesugi M, Nyanguile O, Lu H, Levine AJ, Verdine GL. Induced α helix in the VP16 activation domain upon binding to a human TAF. Science **1997**;277:1310

Table 5.1. Relative protein expression levels of those kinases that are altered by at least 1.5-fold in both MCF-7 and MDA-MB-231 pairs of radioresistant/parental cells. The values represent the mean ratio ± S.D.

Protein	C5/MDA-MB-231	C6/MCF-7
AKT1	0.46 ± 0.21	0.55 ± 0.14
BCR	2.17	2.05 ± 0.30
CAD	1.92 ± 0.39	2.08 ± 0.14
CAMKV	0.29 ± 0.03	0.13
CASK	0.37	0.30 ± 0.17
CCNH	1.67 ± 0.63	2.03 ± 0.21
CDK5	0.51 ± 0.16	0.67 ± 0.11
CHEK1	2.00 ± 0.14	2.42 ± 0.62
CSK	2.33 ± 0.84	1.96 ± 0.41
CSNK2A1	1.95 ± 1.03	1.72 ± 0.08
CSNK2A2	2.01 ± 0.20	1.61 ± 0.10
CSNK2B	1.58 ± 0.14	1.53 ± 0.26
DDR1	4.76	6.01 ± 3.53
EIF2AK2	1.60 ± 0.17	2.34 ± 0.52
ERBB3	0.03	0.52
EXOSC10	1.72 ± 0.35	3.41 ± 1.38
GAK	2.27 ± 0.55	1.84 ± 0.28
GNE	1.68 ± 0.6	4.00
GRK6	1.98	2.80 ± 1.82
HK2	2.44 ± 0.65	1.91 ± 0.34
HSPB8	15.71 ± 11.95	8.75 ± 8.66
ITPK1	1.64 ± 0.98	4.54 ± 1.80
MAGI3	3.46	4.23
MASTL	1.70 ± 0.38	2.06 ± 0.27
PFKM	2.72 ± 0.69	9.45 ± 5.44
PFKP	0.43 ± 0.01	0.54 ± 0.05
PIK3R4	3.90 ± 3.51	3.45
POLR2A	3.53 ± 1.58	1.92 ± 0.18
POLR2B	1.94 ± 0.68	2.46 ± 0.44
POLR2C	1.96 ± 0.10	2.30 ± 1.25
POLR2E	1.96 ± 0.39	1.90 ± 0.40
POLR2G	1.65 ± 0.07	2.58 ± 1.16
POLR2H	2.02 ± 0.90	3.24 ± 0.75
POLR2L	2.00 ± 0.46	7.07 ± 4.33
PRKACB	0.10 ± 0.00	0.14
PTK2	2.11 ± 0.90	6.21
PTK7	15.93 ± 13.90	1.71 ± 0.10
RIOK1	2.48 ± 0.69	5.33 ± 2.00
STK25	0.56	0.42
STK26	0.59 ± 0.16	0.44 ± 0.20
TAF9	4.35	2.85
TK1	1.71 ± 0.64	1.84 ± 0.20
TRIM27	3.85	2.26 ± 1.31
UCK2	1.56 ± 0.43	3.38 ± 1.14
WNK1	1.51 ± 0.33	2.85 ± 1.66

Figure 5.1. A PRM-based targeted proteomic approach for quantifying the differential expression of kinase proteins in two pairs of radioresistant and parental breast cancer cells.

(a) Experimental strategy for the PRM-based targeted proteomic approach. (b) A scatter plot showing the correlation between the ratios of kinase protein expression in C6/MCF-7 cells obtained from forward and reverse SILAC labeling experiments. (c) A scatter plot displaying the correlation between the ratios of kinase protein expression in C5/MDA-MB-231 cells obtained from forward and reverse SILAC labeling experiments.

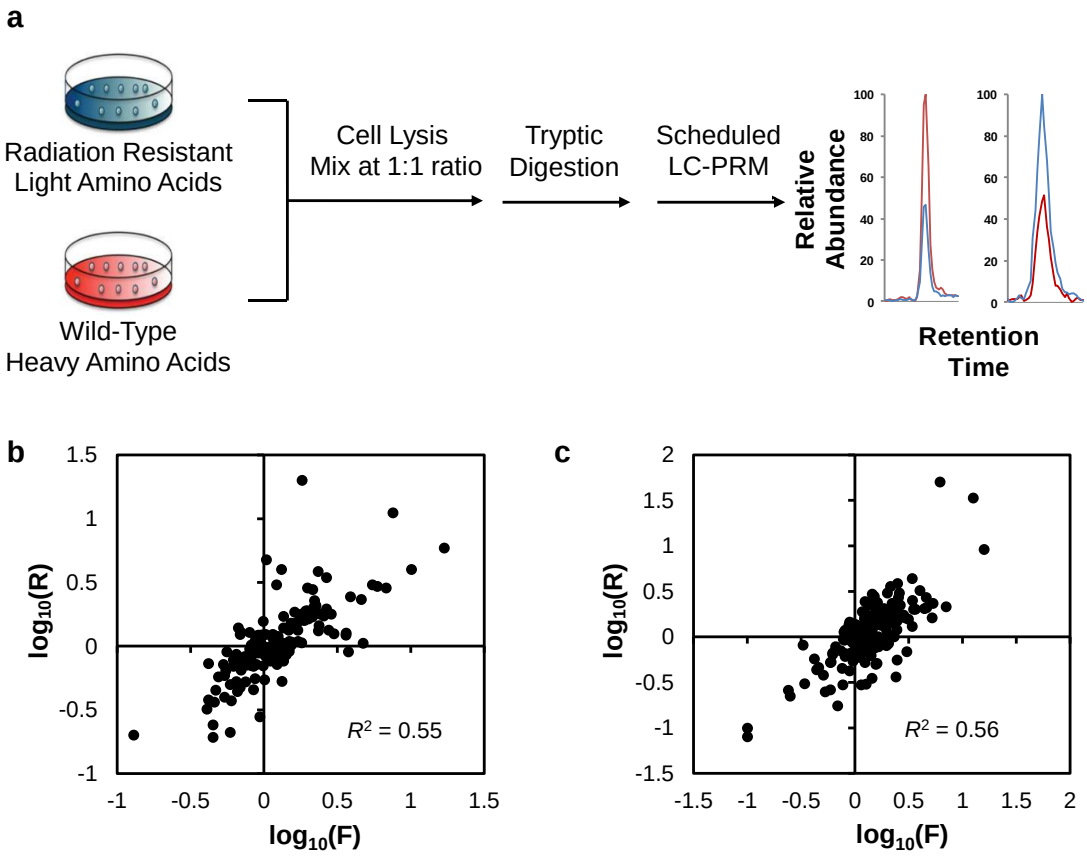


Figure 5.2. Differential expression of kinase proteins in C6/MCF-7 (a) and C5/MDA-MB-231 (b) pairs of breast cancer cells.

The kinase protein expression data represent the means of results obtained from two forward and two reverse SILAC labeling experiments. The ratios obtained from individual measurements are listed in Table S1.

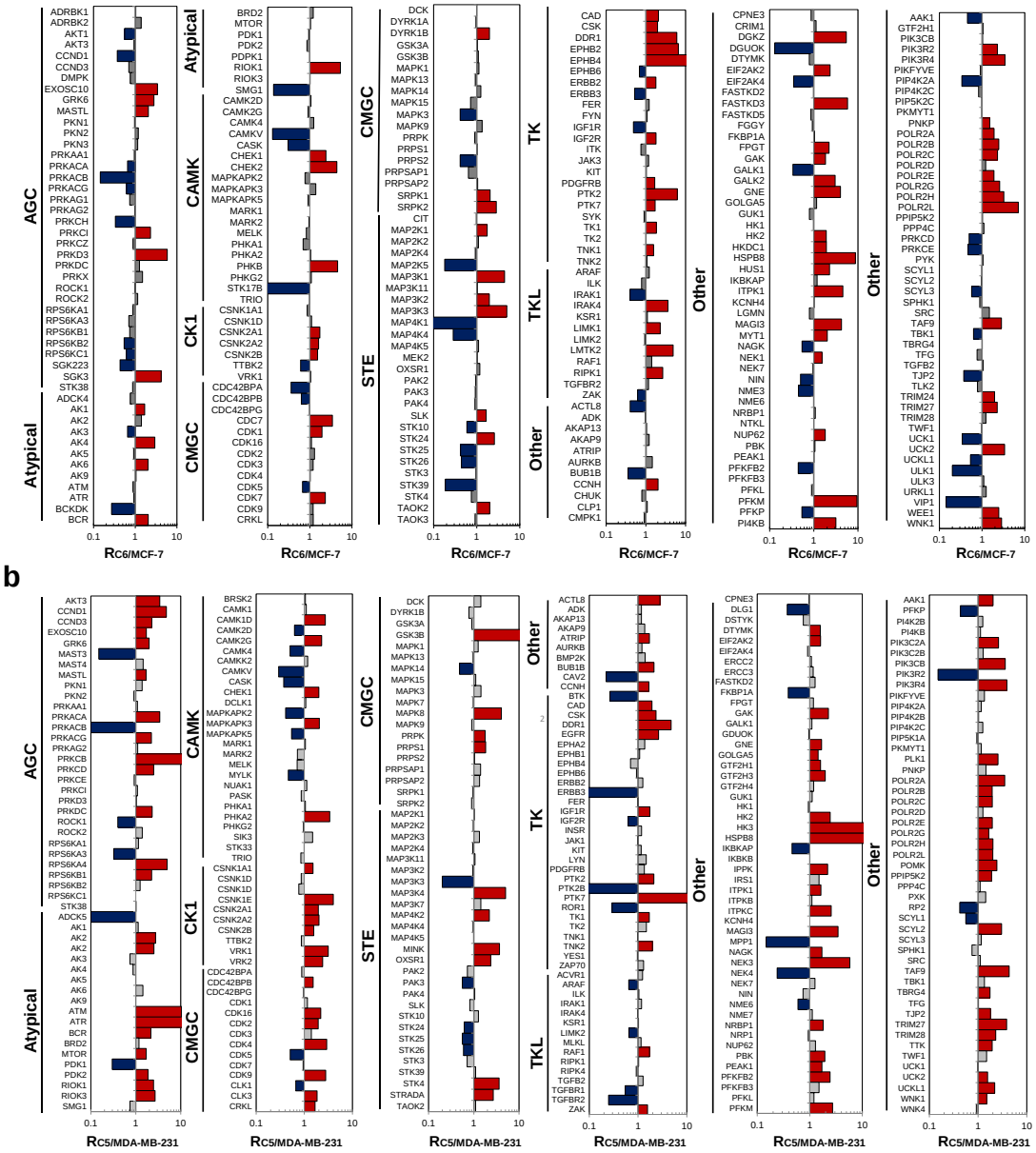


Figure 5.3. Comparison of MRM and PRM result.

(a) A Venn diagram displaying the overlap between quantified kinases from MRM and PRM analysis in C6/MCF-7 pair of breast cancer cells. (b) A scatter plot showing the correlation between the expression ratios of kinases obtained from MRM and PRM analysis in C6/MCF-7 pair of breast cancer cells.

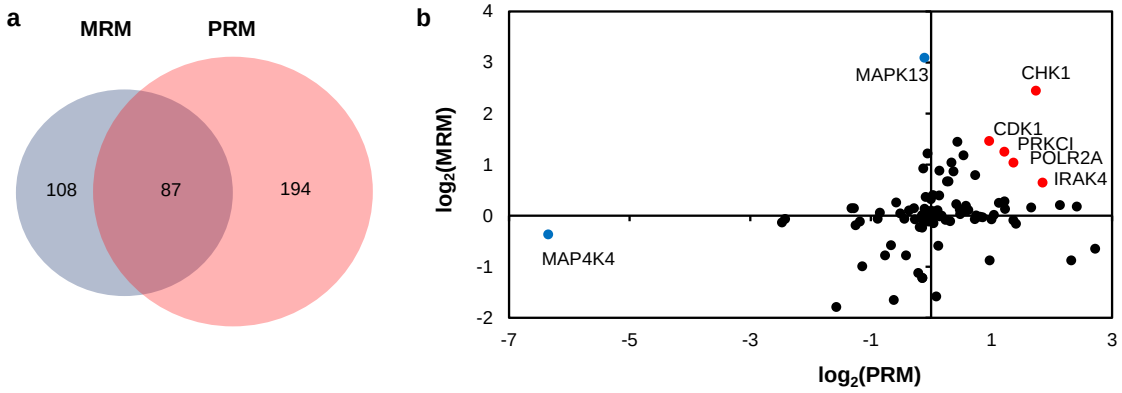


Figure 5.4. A comparison of the quantified kinases in the two pairs of breast cancer cells.

(a) A Venn diagram displaying the overlap between quantified kinases from C5/MDA-MB-231 and C6/MCF-7 pairs of breast cancer cells. (b) A scatter plot showing the correlation between the expression ratios of kinases obtained for the two pairs of radioresistant/parental (WT) cells. (c) A heatmap showing the differences in expression of the commonly altered kinases in the two pairs of radioresistant/WT breast cancer cell lines. Genes were clustered according to Euclidean distance. The data represent the means of results obtained from two forward and two reverse SILAC labeling results, and Table 1 lists the ratios. (d) GO and KEGG pathway analysis of commonly perturbed kinases in the two pairs of WT/radioresistant breast cancer cell lines.

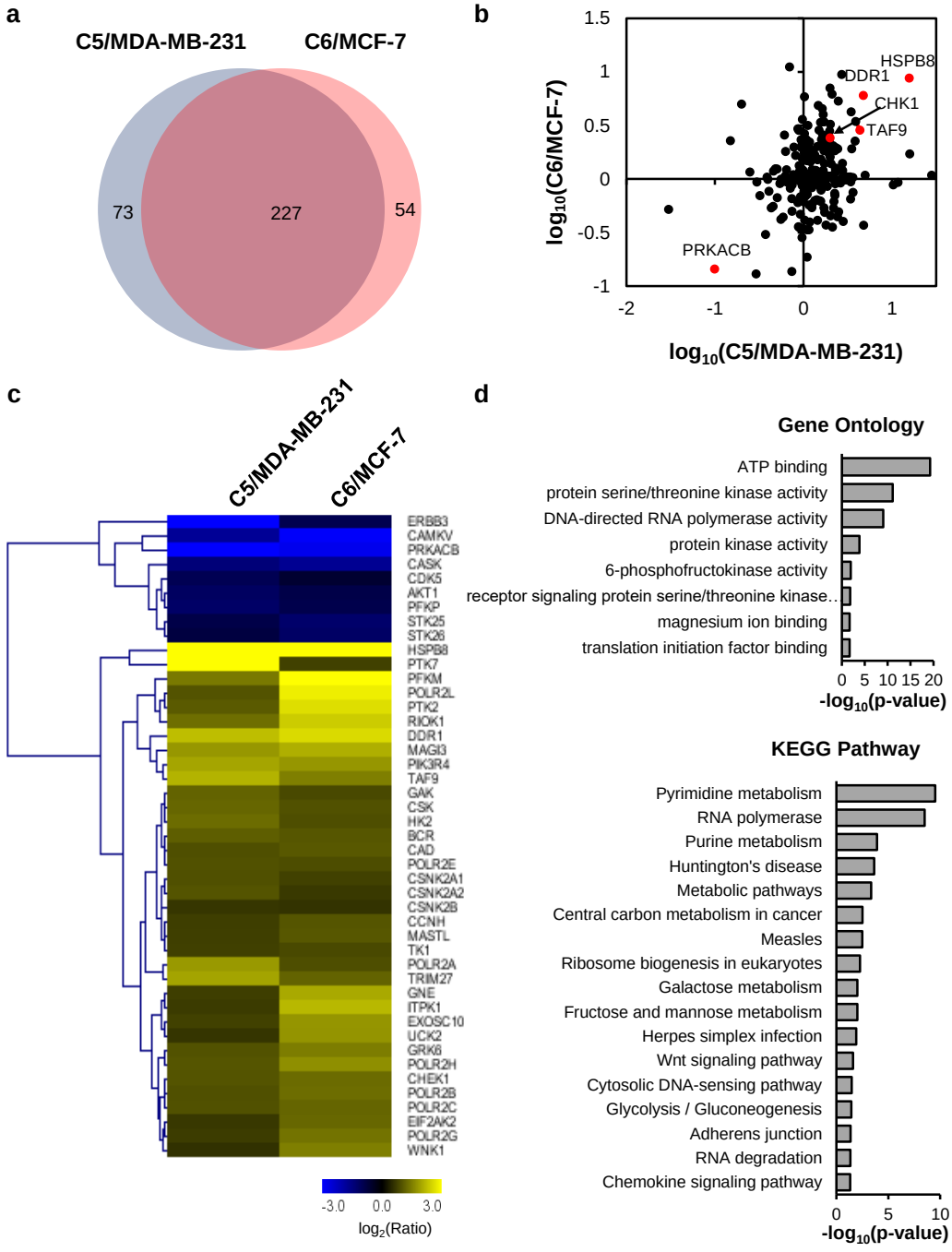

a
C5/MDA-MB-231
C6/MCF-7
73
227
54
b
HSPB8
DDR1
CHK1
TAF9
PRKACB
log10(C6/MCF-7)
log10(C5/MDA-MB-231)
c
C5/MDA-MB-231
C6/MCF-7
ERBB3
CAMKV
PRKACB
CASK
CDK5
AKT1
PFKP
STK25
STK26
HSPB8
PTK7
PFKM
POLR2L
PTK2
RIOK1
DDR1
MAGI3
PIK3R4
TAF9
GAK
CSK
HK2
BCR
CAD
POLR2E
CSNK2A1
CSNK2A2
CSNK2B
CCNH
MASTL
TK1
POLR2A
TRIM27
GNE
ITPK1
EXOSC10
UCK2
GRK6
POLR2H
CHEK1
POLR2B
POLR2C
EIF2AK2
POLR2G
WNK1
log2(Ratio)
d
Gene Ontology
ATP binding
protein serine/threonine kinase activity
DNA-directed RNA polymerase activity
protein kinase activity
6-phosphofructokinase activity
receptor signaling protein serine/threonine kinase..
magnesium ion binding
translation initiation factor binding
-log10(p-value)
KEGG Pathway
Pyrimidine metabolism
RNA polymerase
Purine metabolism
Huntington's disease
Metabolic pathways
Central carbon metabolism in cancer
Measles
Ribosome biogenesis in eukaryotes
Galactose metabolism
Fructose and mannose metabolism
Herpes simplex infection
Wnt signaling pathway
Cytosolic DNA-sensing pathway
Glycolysis / Gluconeogenesis
Adherens junction
RNA degradation
Chemokine signaling pathway
-log10(p-value)

Figure 5.5. Differential expression of CHK1 and TAF9 in acquired radioresistance.

Patient survival correlates with *POLR2H* (a) and *CHEK1* (b) expression in the METABRIC cohort with radiotherapy treatment. p value was calculated by using the log-rank test. (c) Western blot for the validation of the relative expression levels of CHK1 and TAF9 proteins in the paired breast cancer cells. (d) PRM traces for the quantifications of CHK1 and TAF9 proteins in the paired breast cancer cells. (e) Quantitative comparison of the ratios of CHK1 and TAF9 in the paired breast cancer cells obtained from PRM and Western blot analysis. The data represent the mean ± S. D. of the quantification results (n = 3). (f) Box-and-whisker plot showing the comparison of mRNA levels of *CHK1* and *TAF9* gene in cell lines derived from normal breast tissues and primary breast cancer tissues. Data were retrieved from cell lines derived from 114 normal breast tissues and 1097 breast cancer tissues from UALCAN. Box plots represent interquartile range (IQR) including minimum, 25th percentile, median, 75th percentile and maximum values. Outliers are excluded from the plot. The *t*-test was performed using a PERL script with Comprehensive Perl Archive Network (CPAN) module "Statistics::TTest".

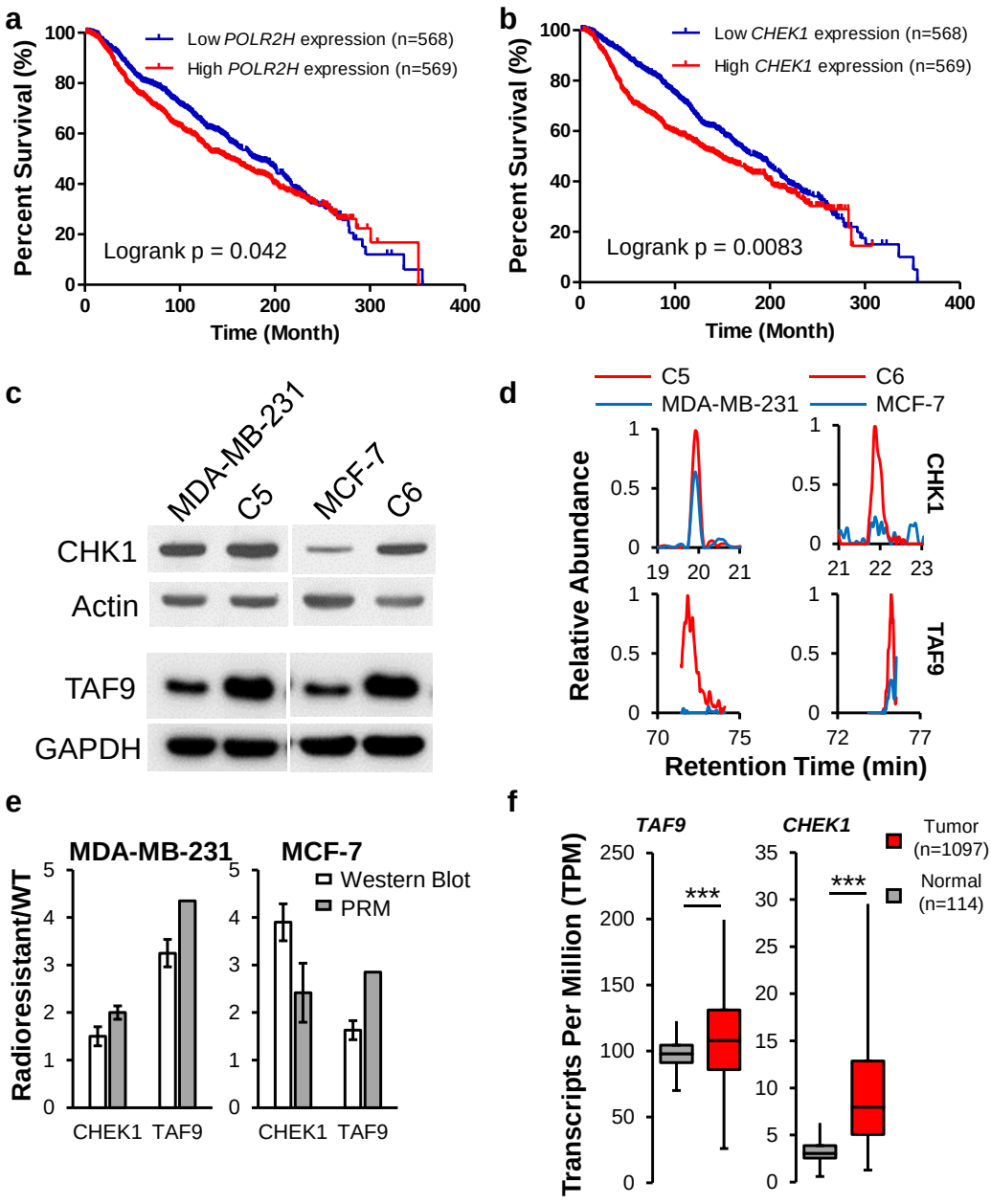
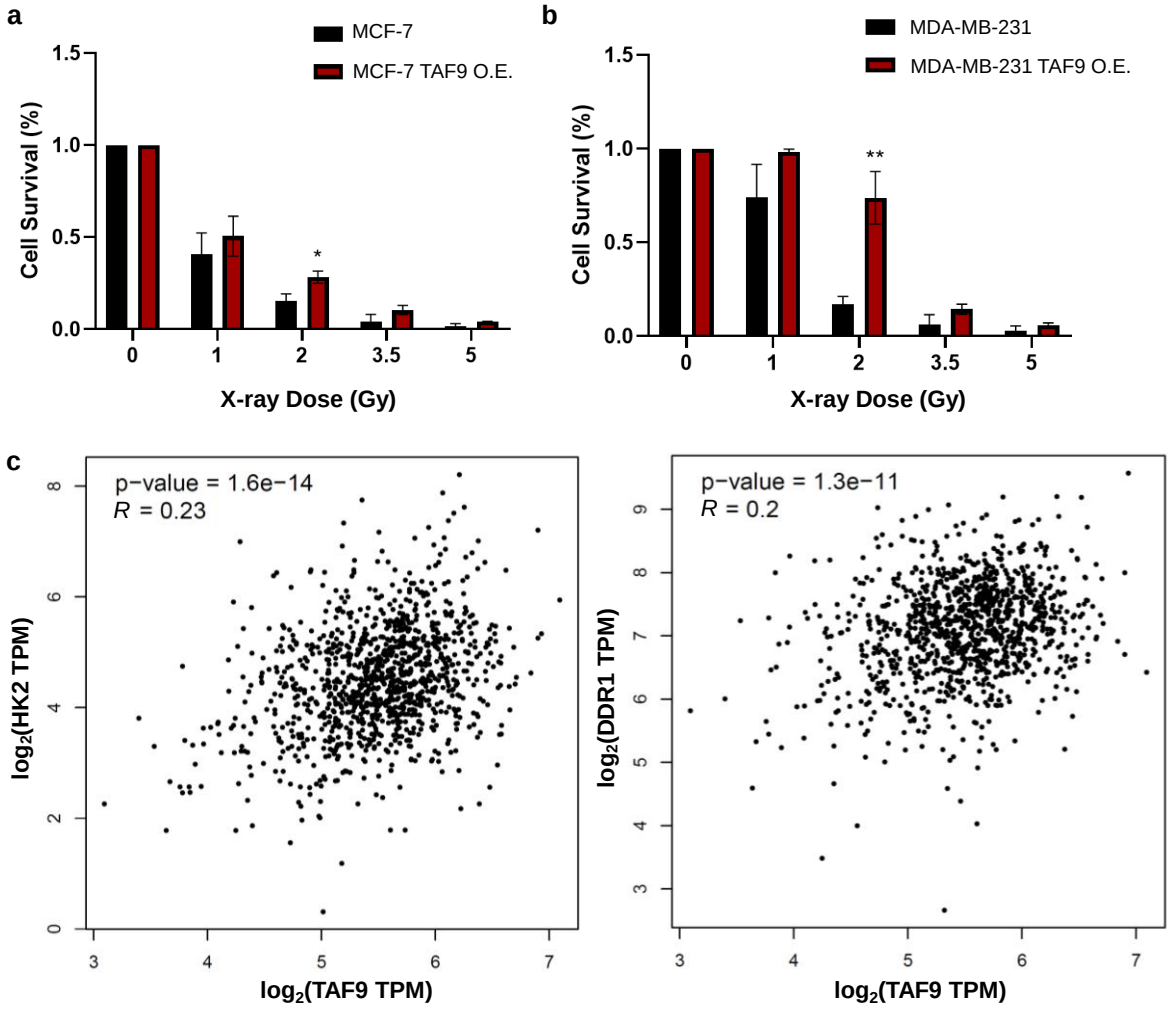

a
Low POLR2H expression (n=568)
High POLR2H expression (n=569)
Percent Survival (%)
Logrank p = 0.042
Time (Month)
b
Low CHEK1 expression (n=568)
High CHEK1 expression (n=569)
Percent Survival (%)
Logrank p = 0.0083
Time (Month)
c
MDA-MB-231
C5
MCF-7
C6
CHK1
Actin
TAF9
GAPDH
d
C5
MDA-MB-231
C6
MCF-7
Relative Abundance
CHK1
TAF9
Retention Time (min)
e
MDA-MB-231
MCF-7
Western Blot
PRM
Radioresistant/WT
CHEK1
TAF9
CHEK1
TAF9
f
TAF9
CHEK1
Tumor (n=1097)
Normal (n=114)
Transcripts Per Million (TPM)
***
***

Figure 5.6. TAF9 modulates the acquisition of radiation resistance in breast cancer cells.

(a) Relative cell survival of MCF-7 (a) and MDA-MB-231 (b) breast cancer cells with exposure to different doses of X rays after ectopic overexpression TAF9 or transfection with an empty vector (control). The *p* values were calculated based on unpaired, two-tailed Student's *t*-test: #, $p \geq 0.05$; *, $0.01 \leq p < 0.05$; **, $0.001 \leq p < 0.01$; ***, $p < 0.001$. (c) Quantitative comparison of mRNA expression between *TAF9* and *HK2*, *TAF9* and *DDR1* obtained from 1085 breast cancer patients.



Chapter 6: Targeted Quantitative Profiling of GTP-binding Proteins Associated with Metastasis of Melanoma Cells

1. Introduction

Melanoma is a principal cause of mortality in the United States, accounting for an estimated 87,110 new cases and 9,730 deaths in 2017 alone (1). Melanoma is typically curable in the early stages, where the five-year survival rate is 99% for localized melanoma; the prognosis, however, worsens considerably to approximately 27% for distant metastatic cases (2). Fortunately, reported mortality for melanoma in the United States decreased dramatically since the FDA approved new therapies, such as checkpoint inhibitor ipilimumab and BRAF inhibitor vemurafenib, for metastatic cancer (3,4). Further investigation of other proteins responsible for regulating melanoma metastasis is of paramount importance to advance cancer therapy and improve overall survival of advanced melanoma patients.

GTPases usually serve as molecular switches in cell signaling pathways (5). Aberrant expression of GTPases is known to be associated with melanoma metastasis. In 2018, Huang *et al.* (6) reported a targeted proteomic method combining SILAC labeling, SDS-PAGE fractionation with MRM analysis to quantitate small GTPases in melanoma cells. Based on LC-MS results and bioinformatic analysis, they revealed the potential regulatory role of RAB38 in melanoma metastasis through the upregulation of Matrix Metallopeptidase 2 (MMP2) and MMP9. Ras Homolog Family Member C (RhoC) has also been identified as a driver of melanoma metastasis, where induction of RHOC expression promotes invasion and metastasis, and the phenotype is reversed upon inhibition (7,8).

Despite these findings, there is a continued need to discover novel drivers and suppressors for melanoma metastasis.

We previously developed isotope-coded desthiobiotin-GTP probes for the enrichment and quantitative profiling of GTP-binding proteins. Combining with an MRM-based proteomic method, we were able to quantify variations in GTP-binding proteins in a pair of primary and metastatic colon cancer cell lines (9). Several up- or down-regulated proteins identified with the method were previously reported to play essential roles in regulating colon cancer metastasis. Considering the effectiveness and efficiency of this method and the importance of discovering novel regulators of melanoma metastasis, we applied this quantitative analysis in two matched pairs of primary and metastatic melanoma cancer cell lines. A few GTP-binding proteins identified from the quantification result were already reported to regulate melanoma cancer metastasis, including Ras-Related Nuclear Protein (RAN)and Ras-Related Protein Rab-27A (RAB27A) (10,11). We selected AK4, one of the most up-regulated proteins, to validate our method's effectiveness and explore its potential role in the migration and invasion of melanoma cells. Based on migration and invasion assays and wound healing scratch assays, the increased AK4 expression level was associated with elevated migration and invasion of melanoma cells.

2. Materials and Methods

2.1 Cell Lysates Labeling with the GTP Probes

WM-115 and WM-266-4 cells were obtained from ATCC, IGR39 and IGR37 cells were generous gifts from Prof. Peter H. Duesberg (University of California, Berkeley, CA).

Cells were cultured in Dulbecco's Modified Eagle Medium (DMEM, Invitrogen-Gibco) supplemented with 10% fetal bovine serum (FBS, Invitrogen-Gibco) and penicillin/streptomycin (100 IU/mL) and maintained at 37°C in a humidified atmosphere containing 5% CO_2. All chemicals were purchased from Sigma-Aldrich unless specifically noted. The preparation of isotope-coded desthiobiotin-GTP affinity probes, cell lysates preparation and labeling were performed as previously reported.**Error! Bookmark not defined.** Briefly, approximately 2×10^7 cells were lysed with 1 mL lysis buffer (0.1% Triton X-100, 50 mM HEPES (pH 7.4), 100 mM NaCl) containing protease inhibitor cocktail. After removal of endogenous nucleotides using NAP-5 columns (Amersham Biosciences), approximately 1 mg lysates in 1 mL lysis buffer were treated with 5 mM EDTA for 5 min, followed by 20 μM GTP probe and 20 mM $MgCl_2$ at room temperature for 2 h. In the forward labeling experiments, the primary and metastatic melanoma cells were labeled with light or heavy GTP probes, respectively. The labeling conditions were swapped in the reverse experiments. After the labeling reaction, the unreacted probes were quenched with 100 mM glycine for 30 min at room temperature. The labeled lysates for a pair of primary and metastatic melanoma cells were combined in 1:1 ratio for further analysis.

2.2 Scheduled MRM analysis

The LC-MS/MS samples were prepared following previously reported procedures. **Error! Bookmark not defined.** Briefly, after probe labeling, the lysates were digested with trypsin using filter-aided sample preparation (FASP) method. The probe-labeled peptides were enriched using high-capacity streptavidin beads (Sigma-Aldrich) and analyzed on a TSQ Vantage triple-quadrupole mass spectrometer (Thermo Fisher) coupled with a

nanoelectrospray ionization source and an EASY n-LC-II HPLC system. Samples were automatically loaded onto a pre-column packed with ~4 cm of 5 µm C18 120 Å reversed-phase material (ReproSil-Pur 120 C18-AQ, Dr. Maisch) at a 3 µL/min flow rate. The pre-column was connected to a 20-cm fused-silica analytical column (75 µm i.d.) packed with 3 µm C18 120 Å reversed-phase material. Peptides were separated with a 140-min HPLC gradient from 10-35% buffer B (acetonitrile with 0.1 % formic acid) in buffer A (water with 0.1 % formic acid) at a flow rate of 230 nL/min. The spray voltage was 1.9 kV. Q1 and Q3 resolutions were 0.7 Da and the cycle time was 5 s.

We employed our previously developed scheduled LC-MRM method for targeted peptide detection in a pre-selected retention time window that was calculated based on iRT calibration using tryptic peptides of BSA as the standards. The quantification results were analyzed with Skyline and manually checked to ensure that the intensity distribution of selected transitions match with theoretical distributions in the spectral library (dotp higher than 0.8) (12). The sum of peak areas for all selected transitions were used for quantification of the heavy over light ratios. The transition lists for peptides in the library and the raw files for LC-MRM analyses for paired melanoma cancer cells were deposited into PeptideAtlas with the identifier number of PASS01639 (http://www.peptideatlas.org/PASS/PASS01639).

2.3 Stable Cell Line Generation

The stable cell lines for knockdown or overexpression of AK4 in melanoma cancer cells were generated with lentivirus. The shRNA sequences were selected from Sigma and cloned into pLKO.1-Puro (Addgene). The loop regions for shRNAs were CTCGAG, the

sense strands were 5′-GCCAGTCATTGAATTATACAA-3′ for shAK4-2e, 5′-GCCAGGCTAAGACAGTACAAA-3′ for shAK4-1, and 5′-TCCTAAGGTTAAGTCGCCCTCG-3′ for scrambled control (shCtrl). The cDNA sequence of AK4 was PCR-amplified from a cDNA library prepared from WM-266-4 cells and cloned into a pLJM1-EGFP vector (Addgene). The lentiviruses were packaged in HEK 293T cells and collected after a 48-h incubation. Primary melanoma cells were infected with lentivirus for AK4 overexpression or empty vector, and metastatic melanoma cells were infected with lentivirus for AK4 knockdown or shCtrl for 48 h. Successfully infected cells were selected with 1 μg/ mL puromycin until the corresponding untransduced cells dead completely. The AK4 expression levels were tested using western blot prior to further experiments.

2.4 Cell Migration and Invasion Assays

For the migration assay, WM or IGR pairs of melanoma cells were suspended in serum-free media, and 2.5×10^4 (for the WM pair) or 4×10^4 (for the IGR pair) cells were added into the chamber of the transwell insert (Corning). Complete DMEM media supplemented with 10% FBS was added into the lower chamber under the insert. After incubation at 37°C for 24 hours (WM cells) or 48 hours (IGR cells), the inserts were washed with PBS and the unmigrated cells inside the insert were gently removed with a cotton swab. The migrated cells were fixed with 70% ethanol followed by staining with 0.5% crystal violet and imaged under an inverted microscope. The WM cells were imaged under bright field and the IGR cells were imaged using the TxRed lens. Cells in four randomly selected fields were counted and averaged. The invasion assay was performed side by side with the migration

assay under the same conditions, except that the inner surface of the transwell insert was coated with Matrigel basement membrane matrix (Corning), at a concentration of 200-400 μg/ml, at 37°C for an hour prior to seeding of cells.

2.5 Wound Healing Assays

Cells were seeded in 6-well plates at a density of 5×10^5 cells per well and incubated until more than 90% confluency. A scratch was made in the cell monolayer with a 100 μl pipette tip and the disrupted cells were removed by gently wash with 1×PBS twice. The remaining cells were cultured at 37°C in DMEM media without FBS to prevent cell proliferation. Images were then taken under an inverted microscope at 0, 24 and 48 hours, to assess the level of cell migration. Images were then analyzed with Image-Pro v10 (Media Cybernetics).

3. Results and Discussion

3.1 Quantitative Profiling of GTP-binding Proteins in Melanoma Cells

We have previously reported the application of GTP acyl phosphate probes for the covalent labeling of GTP binding proteins (Figure 6.1A). These probes were designed to label the nucleophilic lysine residues in the GTP-binding pockets, such as lysines in the highly conserved Walker A motif GxxxxGKT/S in many GTPases, with a desthiobiotin affinity tag. By introducing an isotope-coded γ-amino-butyryl (GABA) linker between GTP and the affinity tag, targeted proteins could be labeled with either a heavy or a light probe to facilitate quantitative analysis. Based on shotgun proteomic data, we constructed a peptide library encompassing 217 probe-labeled GTP-binding proteins and developed a

quantification method using multiple reaction monitoring (MRM) technique (Figure 6.1B). Primary and metastatic melanoma cancer cells were labeled with light or heavy probes, respectively, and mixed at a 1:1 ratio, followed by tryptic digestion. The peptide mixtures were then treated with streptavidin beads to enrich labeled peptides for LC-MRM analysis. At least three top-ranking y-ions were selected for each peptide in the library. Targeted peptides were detected in pre-selected retention time windows based on iRT calibration with 10 selected tryptic peptides from bovine serum albumin (BSA) as external standards. The data were then analyzed with skyline.

3.2 Scheduled MRM Analysis of GTP-binding Proteins in Matched Metastatic and Primary Melanoma Cancer Cells

Metastasis is one of the primary reasons leading to the mortality of melanoma. To evaluate the impact of GTP-binding protein expression on melanoma metastasis, we applied our scheduled MRM method on two pairs of matched metastatic and primary melanoma cell cancer lines. WM-115 and WM-266-4 were derived from primary tumor and skin metastasis from the same 55 years old female patient (13). IGR39 and IGR37 were derived from primary tumor and lymph node metastasis from the same 26 years old male patient (14).

A total of 64 GTP-binding proteins were quantified in both pairs of cell lines (Figure 6.2). Among these proteins, several were previously shown to regulate melanoma progression, including Ras-Related Nuclear Protein (RAN) (11). Another example of this is RAB27A, which has been found to be upregulated in metastatic cancer and promote melanoma metastasis through regulating the MET network (10,15). For most quantified

peptides, the peptide intensity ratio was reasonably consistent between forward and reverse experiments (Figure 6.3A, B). We also observed a significant correlation for differential regulation of proteins between the WM and IGR cell lines pairs (Figure 6.3C).

We next validated the altered expression of AK4 and Transglutaminase 2 (TGM2) between primary and metastatic cell lines. By employing Western blot analysis, we determined that the results about the differential expressions of AK4 and TGM2 in the paired primary/metastatic melanoma cells are consistent with our MRM data, further validating quantification accuracy of our method (Figure 6.3D, 6.7, 6.8). Hence, we chose to further investigate AK4 for its effects on promoting melanoma metastasis.

3.3 AK4 Promoted Migration and Invasion in Melanoma Cells

We next sought to determine whether AK4 can promote the migration and invasion capabilities of metastatic melanoma cell lines. For this purpose, we employed the transwell migration and invasion assay. To employ this method, we first modified the WM-115 and IGR39 cell lines to stably overexpress AK4. We compared these cell lines against WM-115 and IGR39 cell lines containing the empty vector sequence. Using the transwell assay detailed above, we determined that the invasion of WM-115 cells was significantly increased upon overexpression of AK4 (Figure 6.4B, C). This finding was further corroborated with the IGR39 cells, where overexpression of AK4 led to a significant increase in cell migration and invasion (Figure 6.8B, C).

We subsequently sought to determine whether heightened AK4 expression in metastatic melanoma cells can promote melanoma cell migration and invasion. We prepared WM-266-4 and IGR37 cell lines with stable knockdown of AK4. This

knockdown's effectiveness was assessed against cell lines containing an shCtrl sequence with Western blot analysis (Figure 6.4A, 6.8A). Our results demonstrated that the knockdown of AK4 in WM-266-4 cells resulted in a significant decrease in these cells' ability to migrate and invade (Figure 6.4D, E). We observed a similar change with IGR37 cells with stable knockdown of AK4, wherein migration and invasion were significantly decreased (Figure 6.8D, E). Together these results suggest the vital importance of AK4 in promoting cell invasion and migration.

To further investigate this effect, we employed a wound-healing scratch assay to determine if AK4 expression is necessary for melanoma cells to migrate. Using the WM-115 AK4 stable overexpression line, we observed a significant increase in cell migration compared to the control at 48 hours after the scratch (Figure 6.5A, B). Consistent with this trend, we observed an increase in cell migration with the IGR39 cells following cell perturbation (Figure 6. 9A, B). Reciprocal experiment with WM-266-4 and IGR37 cells with stable knockdown of AK4 showed a significant diminution in cell migration (Figure 6.5C, D, 6. 9C, D).

There is considerable research showing that increased AK4 is associated with a poor prognosis in cancer. In lung cancer, elevated AK4 expression promotes metastasis through regulating activating transcription factor 3 (ATF3). In one mechanism, this occurs through the AK4-induced down-regulation of ATF3, which leads to the induction of MMP2 (16). In the same vein, increased AK4 expression has also been documented to promote the stabilization of HIF-1α, thus driving the epithelial-to-mesenchymal transition (EMT) in lung cancer cells (17). Altogether, these results provide significant evidence of AK4's role

in promoting melanoma cell metastasis through supporting cell invasion and migration capabilities.

4. Conclusions

In this study, we developed a high-throughput chemoproteomic method for the quantitative analysis of GTP-binding proteins, where the method involves the use of stable isotope-coded desthiobiotin-GTP acyl phosphate probes and a peptide library encompassing 217 probe-labeled peptides derived from GTP-binding proteins. We subsequently used this library to design an MRM method for the targeted quantification of these proteins. By employing this method, we were able to quantify 64 GTP-binding proteins across two primary and metastatic melanoma cells pairs. From these data, we found that AK4 and TGM2 proteins were significantly up- and downregulated, respectively, in the metastatic cell lines compared to their primary counterparts. We further demonstrated through migration/invasion assays and the wound healing scratch assay that the upregulation of AK4 is necessary for the increased migration and invasion observed in the metastatic cell lines.

References

1. Cancer Facts & Figures 2017 [Internet]. [cited 2021 Feb 18]. Available from: https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2017.html

2. Melanoma Survival Rates | Melanoma Survival Statistics [Internet]. [cited 2021 Feb 18]. Available from: https://www.cancer.org/cancer/melanoma-skin-cancer/detection-diagnosis-staging/survival-rates-for-melanoma-skin-cancer-by-stage.html

3. Tarhini AA, Toor K, Chan K, McDermott DF, Mohr P, Larkin J, et al. A matching-adjusted indirect comparison of combination nivolumab plus ipilimumab with BRAF plus MEK inhibitors for the treatment of BRAF-mutant advanced melanoma☆. ESMO Open. 2021 Feb 5;6(2):100050.

4. Shirkavand A, Mohajerani E, Farivar S, Ataie-Fashtami L, Ghazimoradi MH. Monitoring the Response of Skin Melanoma Cell Line (A375) to Treatment with Vemurafenib: A Pilot In Vitro Optical Spectroscopic Study. Photobiomodulation, Photomedicine, and Laser Surgery. 2021 Feb 15;

5. Song S, Cong W, Zhou S, Shi Y, Dai W, Zhang H, et al. Small GTPases: Structure, biological function and its interaction with nanoparticles. Asian Journal of Pharmaceutical Sciences. 2019 Jan;14(1):30–39.

6. Huang M, Qi TF, Li L, Zhang G, Wang Y. A Targeted Quantitative Proteomic Approach Assesses the Reprogramming of Small GTPases during Melanoma Metastasis. Cancer Res. 2018 Sep 15;78(18):5431–5445.

7. Boone B, Van Gele M, Lambert J, Haspeslagh M, Brochez L. The role of RhoC in growth and metastatic capacity of melanoma. J Cutan Pathol. 2009 Jun;36(6):629–636.

8. Clark EA, Golub TR, Lander ES, Hynes RO. Genomic analysis of metastasis reveals an essential role for RhoC. Nature. 2000 Aug 3;406(6795):532–535.

9. Cai R, Huang M, Wang Y. Targeted Quantitative Profiling of GTP-Binding Proteins in Cancer Cells Using Isotope-Coded GTP Probes. Anal Chem. 2018 Dec 18;90(24):14339–14346.

10. Guo D, Lui GYL, Lai SL, Wilmott JS, Tikoo S, Jackett LA, et al. RAB27A promotes melanoma cell invasion and metastasis via regulation of pro-invasive exosomes. Int J Cancer. 2019 Jun 15;144(12):3070–3085.

11. Caputo E, Wang E, Valentino A, Crispi S, De Giorgi V, Fico A, et al. Ran signaling in melanoma: implications for the development of alternative therapeutic strategies. Cancer Lett. 2015 Feb 1;357(1):286–296.

12. MacLean B, Tomazela DM, Shulman N, Chambers M, Finney GL, Frewen B, et al. Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. Bioinformatics. 2010 Apr 1;26(7):966–968.

13. Balaban G, Herlyn M, Guerry D, Bartolo R, Koprowski H, Clark WH, et al. Cytogenetics of human malignant melanoma and premalignant lesions. Cancer Genet Cytogenet. 1984 Apr;11(4):429–439.

14. Aubert C, Rougé F, Galindo JR. Tumorigenicity of human malignant melanocytes in nude mice in relation to their differentiation in vitro. J Natl Cancer Inst. 1980 May;64(5):1029–1040.

15. Peinado H, Alečković M, Lavotshkin S, Matei I, Costa-Silva B, Moreno-Bueno G, et al. Melanoma exosomes educate bone marrow progenitor cells toward a pro-metastatic phenotype through MET. Nat Med. 2012 Jun;18(6):883–891.

16. Jan Y-H, Tsai H-Y, Yang C-J, Huang M-S, Yang Y-F, Lai T-C, et al. Adenylate kinase-4 is a marker of poor clinical outcomes that promotes metastasis of lung cancer by downregulating the transcription factor ATF3. Cancer Res. 2012 Oct 1;72(19):5119–5129.

17. Jan Y-H, Lai T-C, Yang C-J, Lin Y-F, Huang M-S, Hsiao M. Adenylate kinase 4 modulates oxidative stress and stabilizes HIF-1α to drive lung adenocarcinoma metastasis. J Hematol Oncol. 2019 Jan 29;12(1):12.

Figure 6.1. The design of stable isotope-encoded desthiobiotin-GTP probes and the quantification workflow.

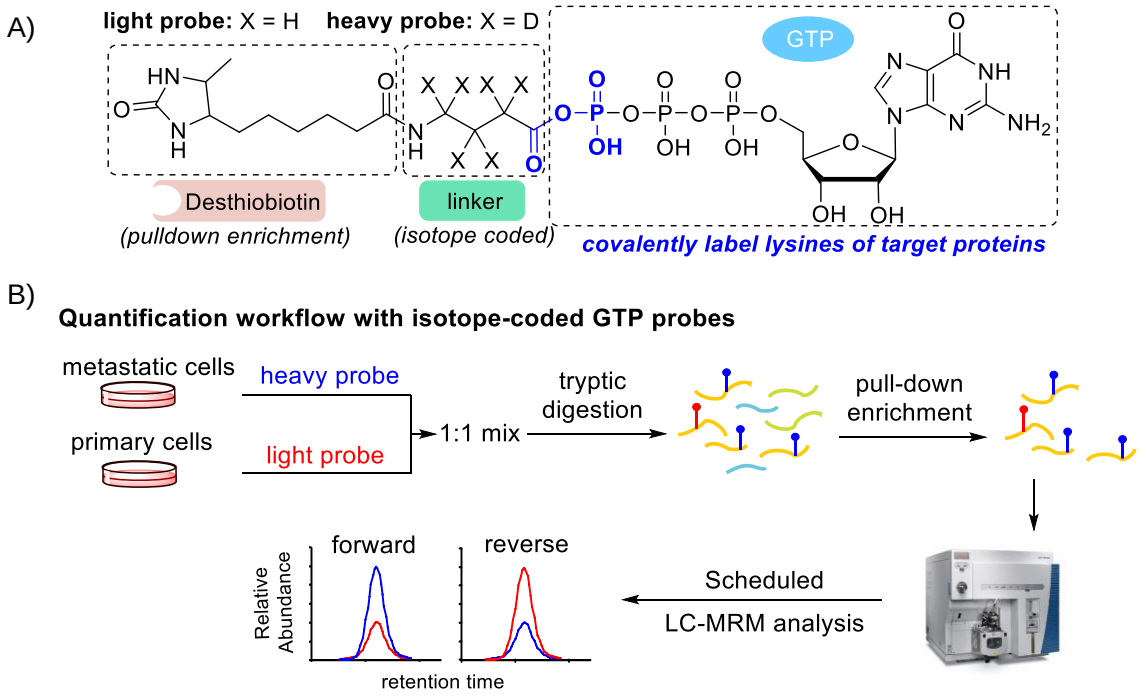


Figure 6.2. A heatmap showing the quantification results for GTP-binding proteins in all four experiments.

The values for WM-F and WM-R are results obtained from the forward and reverse probe labeling experiments in the WM pair of cell lines, respectively. The values for IGR-F and IGR-R represent the corresponding quantification results for the IGR pair of cell lines.

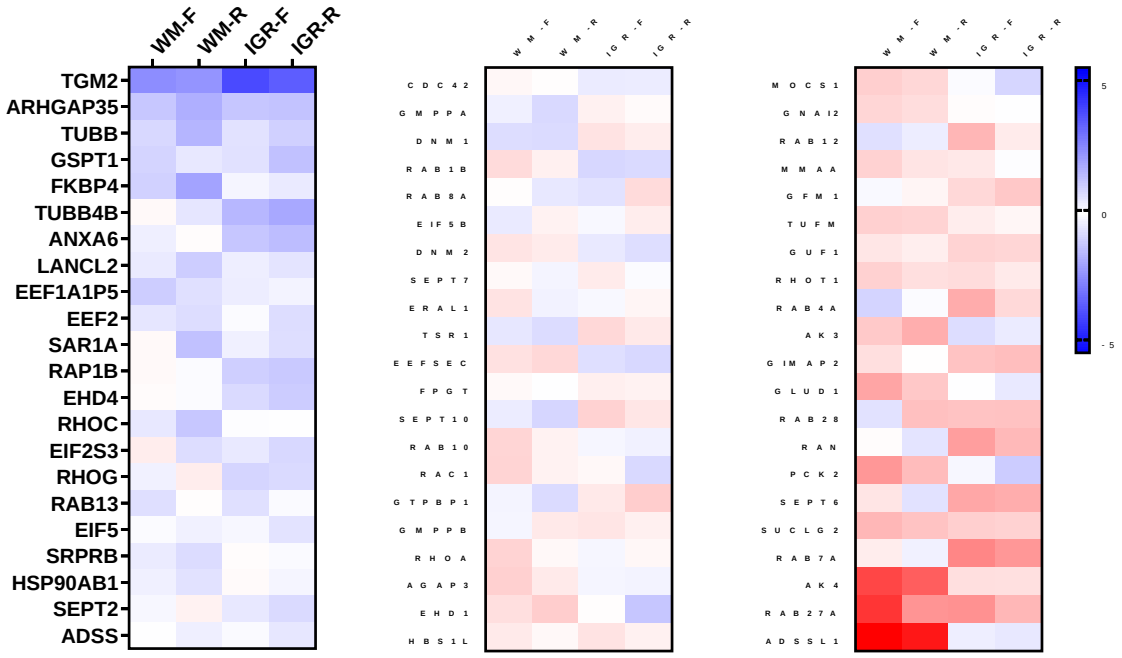


Figure 6.3. Analysis and validation of the MRM quantification results of labeled GTP-binding proteins.

A) comparison between the forward and reverse labeling experiments of the WM-pair cell lines; B) comparison between the forward and reverse labeling experiments of the IGR-pair cell lines; C) comparison between the average ratios for WM- and IGR-pair cell lines; D) validation of expression levels for selected GTP-binding proteins with Western blot.

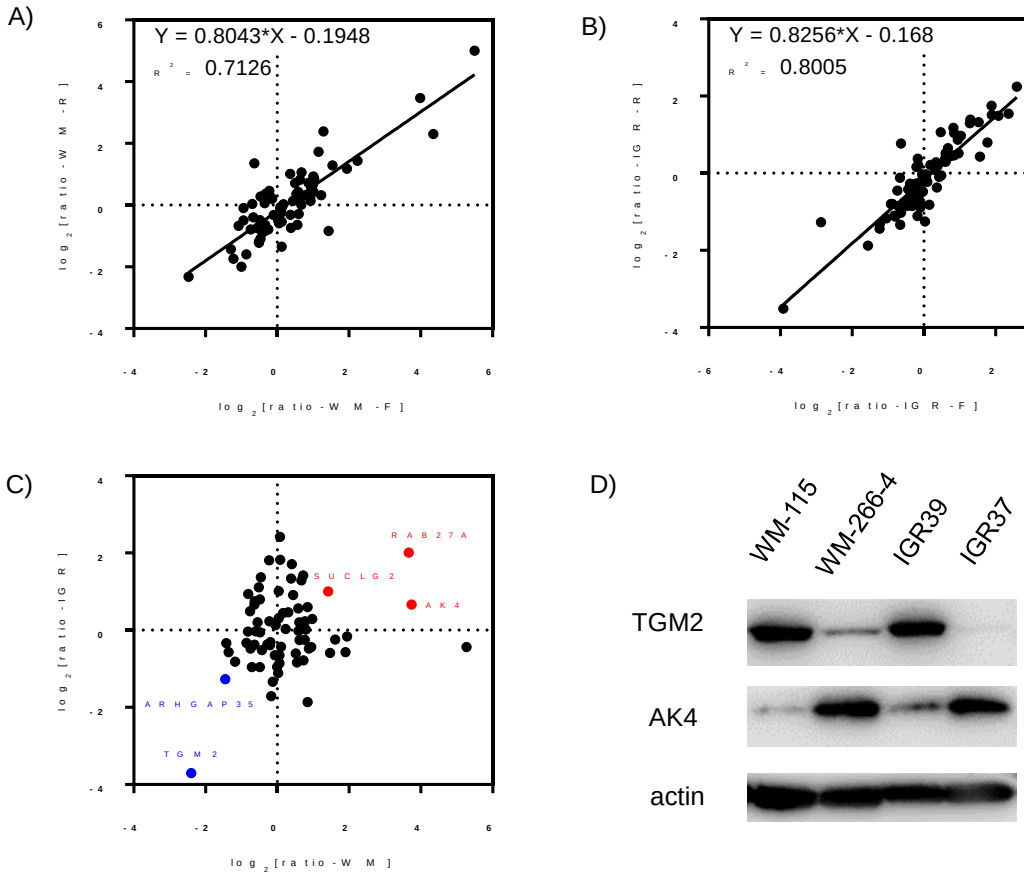


Figure 6.4. AK4 promoted migration and invasion in the WM-pair cell lines.

A) Western blot showing the expression level of AK4 after its stable knockdown in WM-266-4 cells; B) representative images and C) quantification results of the migration and invasion assay showing the influence of AK4 overexpression in WM-115 cells; D) representative images and E) quantification results of the migration and invasion assay showing the influence of stable AK4 knockdown (with two different shRNAs) compared with shCtrl in WM-266-4 cells.

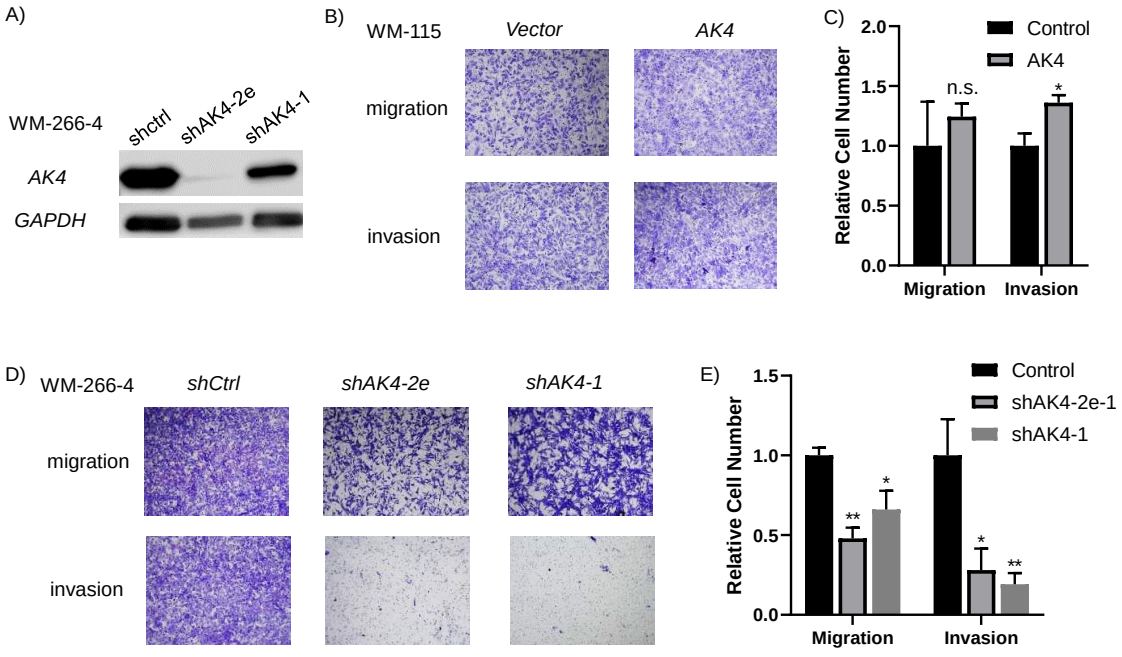


Figure 6.5. Wound healing assay for the WM-pair cell lines.

A) representative images and B) quantification results showing the influence of AK4 overexpression in WM-115 cells; C) representative images and D) quantification results showing the influence of AK4 knockdown in WM-266-4 cells.

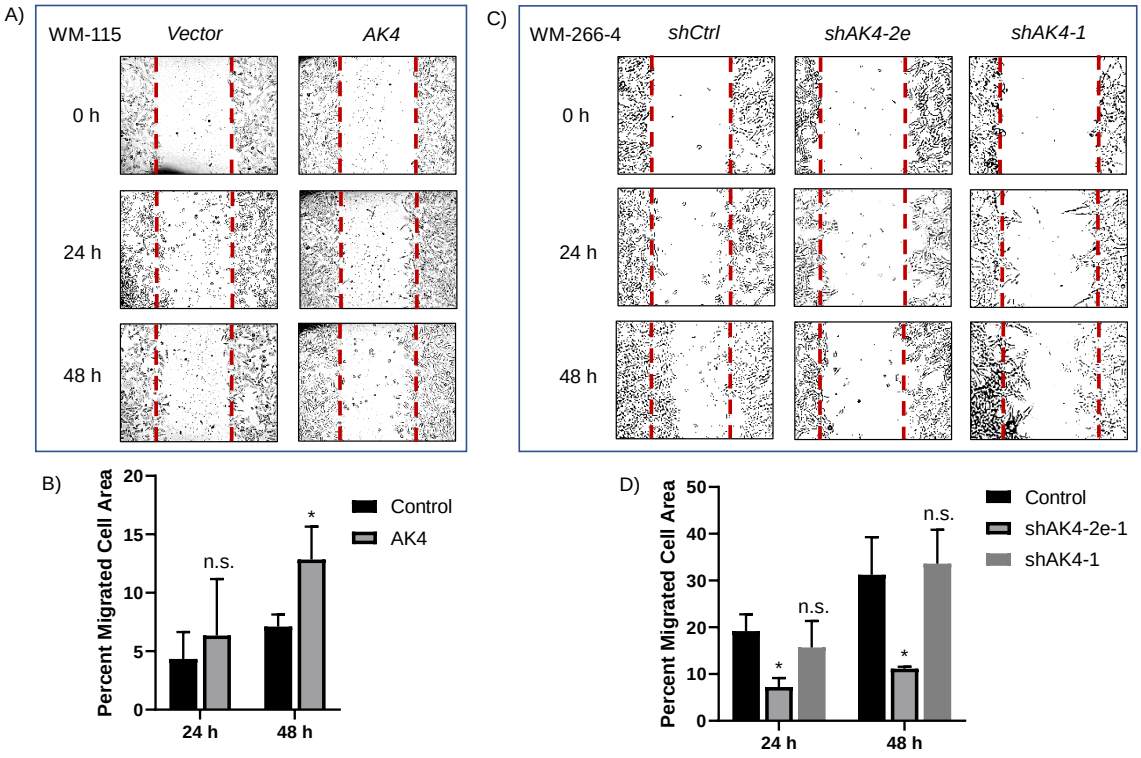


Figure 6.6. Selected-ion chromatograms for the light and heavy forms of the targeted TGM2 peptide.

Data are obtained from A) forward experiment of the WM-pair of cell lines; B) reverse experiment of the WM-pair of cell lines; C) forward experiment of the IGR-pair of cell lines; D) reverse experiment of the IGR-pair of cell lines.

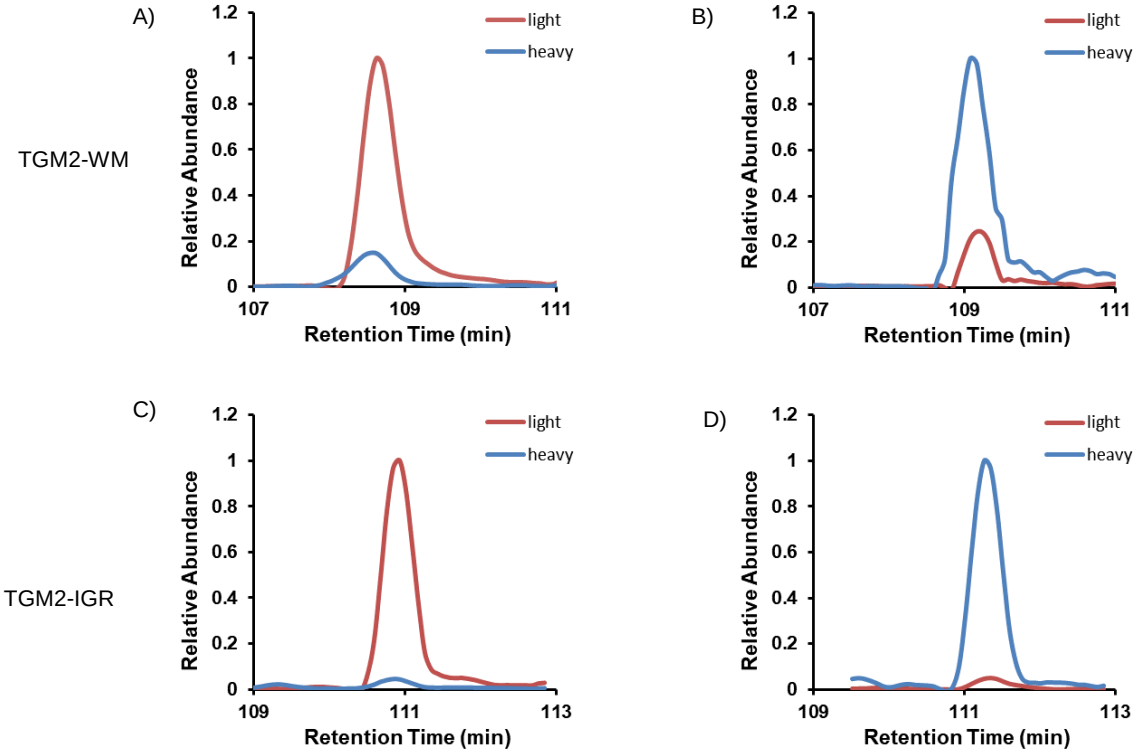


Figure 6.7. Selected-ion chromatograms for the light and heavy forms of the targeted AK4 peptide.

Data are obtained from A) forward experiment of the WM-pair of cell lines; B) reverse experiment of the WM-pair of cell lines; C) forward experiment of the IGR-pair of cell lines; D) reverse experiment of the IGR-pair of cell lines.

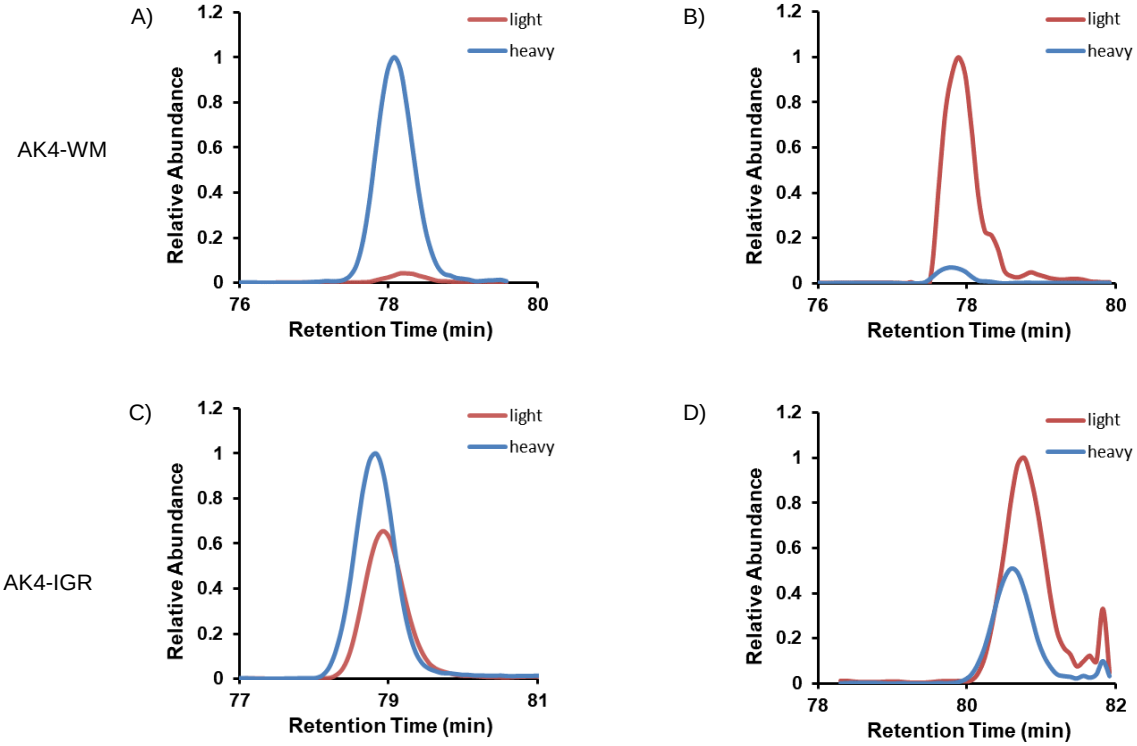


Figure 6.8. AK4 promoted migration and invasion in the IGR-pair cell lines.

A) western blot showing the expression level of AK4 after overexpression in IGR39 cells and stable knockdown in IGR37 cells; B) representative images and C) quantification results of the migration and invasion assay showing the influence of AK4 overexpression in IGR39 cells; D) representative images and E) quantification results of the migration and invasion assay showing the influence of stable AK4 knockdown (with two different shRNAs) compared with shCtrl in IGR37 cells.

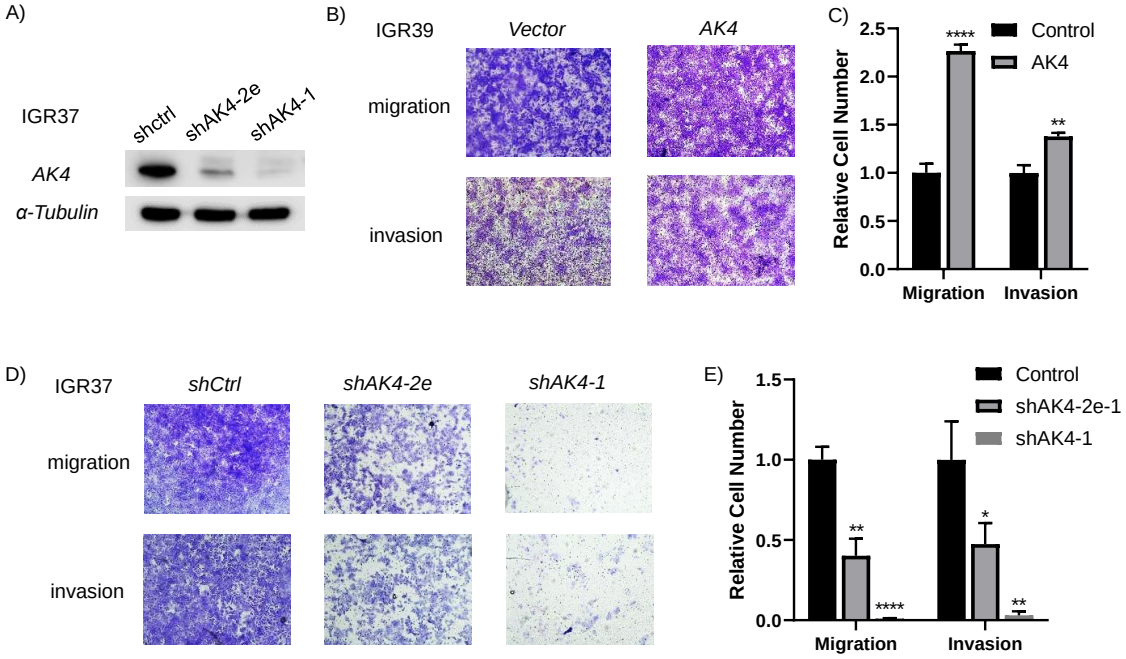
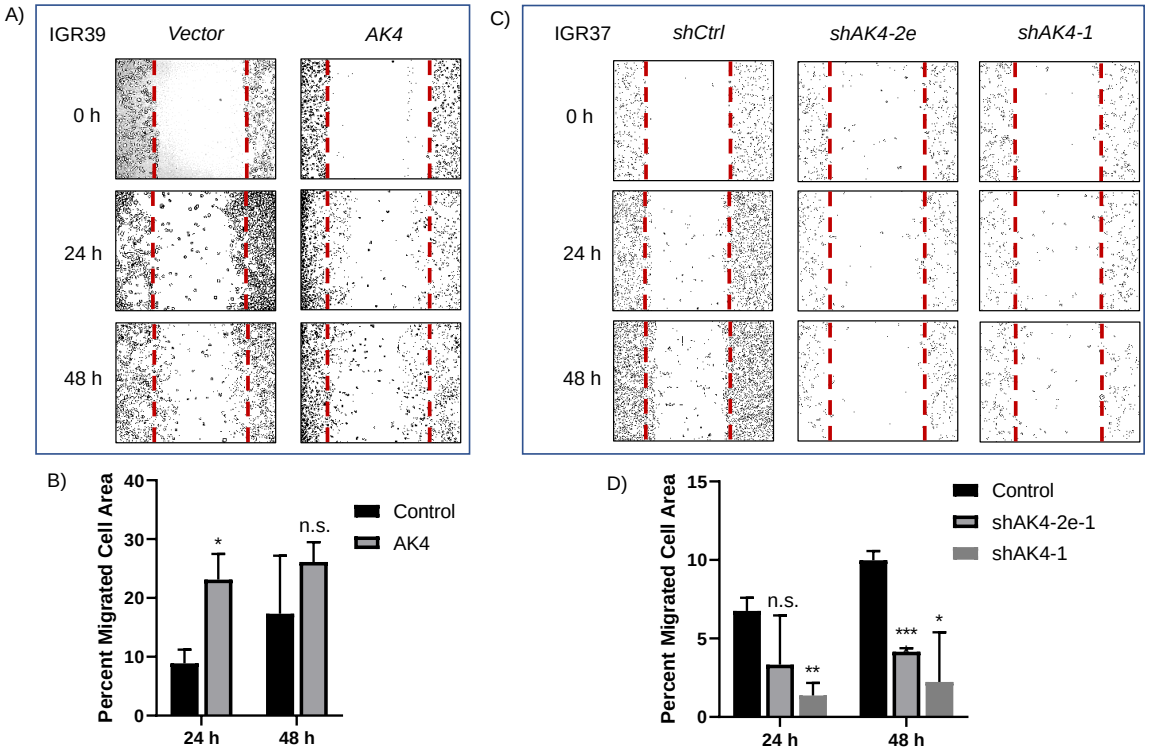


Figure 6.9. Wound healing assay for the IGR-pair cell lines.

A) representative images and B) quantification results showing the influence of AK4 overexpression in WM-115 cells; C) representative images and D) quantification results showing the influence of AK4 knockdown in WM-266-4 cells.



Chapter 7: Conclusions and Future Remarks

The research in this dissertation focused on two main goals. In the first avenue, we sought to identify novel protein substrates for two types of protein methylation, histidine methylation and α-N-methylation, as well as to elucidate how the installation of the latter type of methylation is regulated. Our research initially focused on discovering how NTMT1 expression is regulated by the m^6A-mediated epitranscriptomic pathway, and how this affects the α-N-methylation of a novel substrate, MRG15. We subsequently sought to identify novel substrates of both α-N-methylation and histidine methylation, both PTMs that are as-of-yet poorly understood. Our second primary avenue of research was to discover drivers and suppressors of metastasis in melanoma cells and acquired radioresistance in breast cancer cells.

In Chapter Two, we sought to determine if NTMT1, and by extension the entire process of α-N-methylation is regulated epitranscriptomically. To this end, we used isogenic cells with genetic knockout of writers/readers/erasers of m^6A, one of the most well studied mRNA modifications. In cells with a loss of METTL3 expression, we observed a dramatic increase in NTMT1 expression. This finding was mirrored when testing isogenic cells without FTO expression. Furthermore, using cells lacking expression of YTHDF1/2, we observed an increase in NTMT1 expression, a trend that was opposite in cells without YTHDF3. Together, these data showed that NTMT1 expression and the α-N-methylation that it is responsible for are regulated in an epitranscriptomic manner. Subsequently, we sought to identify a new protein substrate of NTMT1. MRG15 contains the N-terminal XPK motif that is common among NTMT1

substrates. By using affinity enrichment and mass spectrometry analysis, we discovered that MRG15 exhibits α-N-methylation. Moreover, to determine if MRG15 is indeed a NTMT1 substrate, we conducted the same experiment using siRNA to deplete NTMT1 expression. Upon depletion of NTMT1, we observed a loss of MRG15 α-N-methylation. This was further confirmed by expressing an MRG15 plasmid harboring a K4Q mutation, known to block NTMT1 methyltransferase activity. Conducting the same experiment, we observed a loss of MRG15 N-terminal methylation, confirming NTMT1 is the primary protein responsible for this methylation. We further tested whether MRG15 α-N-methylation is regulated epitranscriptomically. Using the isogenic cells with genetic ablations of the m^6A writers/readers/erasers, we observed promotion or suppression of MRG15 α-N-methylation consistent with previous data. Altogether, we demonstrated that NTMT1 expression is regulated epitranscriptomically, including its ability to target a novel protein substrate, MRG15.

In Chapters Three and Four, we discuss the mining of publicly available mass spectrometric datasets for the identification of novel protein substrates exhibiting α-N-methylation or histidine methylation. This method consisted of comparing spectra from publicly available datasets against the theoretical peptide fragments produced from all proteins in the human proteome with MaxQuant to determine novel PTM substrates. Our results uncovered 219 instances of N-terminal methylation across 196 proteins. Furthermore, we used the affinity purification and mass spectrometric analysis to confirm the discovery of VAMP4 as a substrate for α-N-methylation. Given the knowledge that VAMP4 contains a conserved XPK motif known to be targeted by NTMT1, we

subsequently sought to determine if NTMT1 is responsible for this methylation. Through employing isogenic cells lacking expression of NTMT1, or a K4Q mutant of VAMP4, we observed a dramatic loss of VAMP4 α-N-methylation. We subsequently applied the same method of screening publicly available datasets for the identification of novel protein substrates of histidine methylation. Using this method, we uncovered 33 instances of histidine methylation among 26 proteins. Moreover, using the affinity purification method detailed above, we confirmed that RBM22 is a novel protein with histidine methylated. Together, our results have significantly expanded on the pool of potential α-N-methylated and histidine methylated proteins, including VAMP4 and RBM22. Future experiments should focus on confirming other potential histidine methylated proteins and determining how the α-N-methylation of VAMP4 and histidine methylation of RBM22 modulate the functions of these proteins.

In Chapter Five, we used SILAC labeling and a PRM-based LC-MS/MS method to determine the differential expression of kinase proteins between breast cancer cell lines MCF7 and MDA-MB-231 and their radioresistance clones (C6 and C5 respectively). By employing this method, we were able to quantify the relative expression levels of 300 and 281 protein kinases in C5/MDA-MB-231 and C6/MCF-7 pairs of breast cancer cells, respectively. Among the proteins shown to have increased expression in the radioresistance clones, we identified TAF9 as a potential driver of acquired radioresistance. Through employing clonogenic survival assays, we further confirmed TAF9 as a driver of radioresistance in breast cancer cell lines. Further studies would

focus on confirming TAF9s role in radioresistance, as well as to determine whether TAF9 effects radioresistance in animal models.

In Chapter Six, we employed an MRM-based LC-MS/MS method for the quantitative profiling of GTP-binding proteins. For this purpose, we used stable isotope-coded desthiobiotin-GTP probes for the targeted enrichment of GTP-binding proteins. A total of 64 GTP-binding proteins were quantified in both pairs of cell lines, including several proteins such as RAB27A which are known drivers of melanoma metastasis. From these proteins, we identified AK4 as a potential driver of metastasis as it was one of the most upregulated proteins in the metastatic cell lines. We further discovered that AK4 expression is necessary for melanoma cell invasion and migration.