UNIVERSITY OF CALIFORNIA
RIVERSIDE

Quantitative Proteomic Profiling of TBC Domain-Containing Proteins and Epitranscriptomic Reader, Writer and Eraser Proteins

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

Doctor of Philosophy

in

Environmental Toxicology

by

Tianyu Qi

March 2022

Dissertation Committee:
Dr. Yinsheng Wang, Chairperson
Dr. Wenwan Zhong
Dr. Joseph Genereux

The Dissertation of Tianyu Qi is approved:

Committee Chairperson

University of California, Riverside

ACKNOWLEDGMENTS

Completing a PhD is a journey full of hardship, challenge, and reward. I could not have accomplished it without people who have helped and supported me.

First, I would like to express my deepest gratitude to my doctoral advisor and mentor Dr. Yinsheng Wang. He is the one who opened the door of Mass Spectrometry-based proteomics to me, allowing me to see the beauty of combining analytical chemistry and chemical biology. Whenever I experience difficulties in research, he is always there to help and guide me through it. Whenever I have manuscript drafts of journal articles requiring edits, he always provides thorough revisions in a short amount of time (within 24 hours). He is quite patient during this process, willing to provide edits over more than six rounds until the finished product is polished. His lifetime dedication to research, curiosity and ingenuity in learning new techniques, meticulous scholarship, and extensive collaboration with other researchers inspires me to become a better researcher every day.

Next, I would like to pass my thankfulness to my oral qualifying committee members Dr. Wenwan Zhong, Dr. Jeff Perry, Dr. Ernest Martinez, and Dr. Zhenbiao Yang, and my guidance committee Dr. Cheng Quan and Dr. Ernest Martinez for their invaluable advice and support on my projects. I would like to especially thank Dr. Joseph Genereux for kindly agreeing to serve on my defense committee following Dr. Perry’s departure from UCR. I would like to acknowledge Dr. Genereux and Dr. Min Xue for sharing their lab equipment with me, Dr. Clay Clark and Dr. David Carter for their extensive training and service at Institute for Integrative Genome Biology at UCR. I would also like to acknowledge the

Environmental Toxicology Graduate Program at UCR and National Institutes of Health (NIH) for their financial support, without which this moment would not have been possible.

I owe a debt of gratitude to Dr. Ming Huang and Dr. Weili Miao, who guided me in performing MS-based proteomics. I discovered not only how to conduct specific experiments from them, but I also observed the critical characteristics of being great researchers from them. I saw the importance of perseverance when facing research difficulties from Ming, who was extremely hardworking and published four papers in a row in her last year of PhD. I gathered the necessity of setting clear short- and long-term research goals and being highly motivated and efficient from Weili, the most productive student graduating from Wang lab to date. I am grateful to have them as my PhD peer mentors.

I would like to pass my thankfulness to Wang lab members, Dr. Yuxiang Cui, Dr. Gwendolyn Gonzales, Dr. Xuejiao Dong, Dr. Quanqing Zhang, Jiekai Yin, and Yen-Yu Yang, who are my "Mass Spec" buddies in Wang lab. Dr. Yuxiang Cui taught me a lot about mass spectrometer maintenance. Jiekai and Yen-Yu are incredibly patient and generous in sharing their experience to help me troubleshoot error messages from mass spectrometers. I am so grateful for their help in mass spectrometry and their comradery and support in daily life as great friends.

Special thanks must be given to previous lab member Dr. Nathan Price who provided me with tremendous help in scientific writing and daily life, such as grocery shopping when I did not have a car. My sincere appreciation also goes to my "deskmate" Jun Yuan, who helped me in experimental troubleshooting and bioinformatics. In addition, I would like to thank Dr. Xiaochuan Liu, Dr. Yuxiang Sun, Dr. Feng Tang, Dr. Lin Li, Dr. Xiaoxia Dai, Dr. Rong Cai, and Dr. Xiaomei He for sharing their expertise in different biological experiments. I am also grateful for the companionship with Zi Gao, Shiyuan Guo, Xin Wang, Dr. Andrew Kullem, Dr. Yun Xiong, Xingyuan Chen, Dr. Songbo Wei, Garit Clabaugh, Zhongwen Cao, Ting Zhao, and Chen Wang.

I would not have made it through this program without the friendship of so many people that helped to keep me grounded. I would like to offer my profound thankfulness to my dearest friends Yinan Dong, and Yangsong Li for their constant company, help and support. I am so lucky to have them as “besties”. I also greatly cherish my friendship with Younan Liang, Sheri Tao, Craig Douglass, Dr. Siwen Wang, Yuwei Cui, Liuqing Yang, Bi Chen, Daniel Shields, Jianan Sun, and Mengyuan Xi.

I would like to express my most profound thankfulness to my parents, who love me unconditionally. They are unselfish and supportive in sending their only daughter to study abroad. I could not have come to pursue my PhD studies at UCR without their unconditional love. At last, I am the luckiest to find my previous labmate, best friend, mentor in politics, cooking partner (chef), listener of my opinions of the stock market, and my husband Dr. David L. Bade. Life has ups and downs, as does research. David is the one who always pays attention and understands my complaints, and constantly passes his positive attitude to me. His jokes and positive attitude make difficulties in research easier to handle. Without his constant company and support, I could not finish my PhD. Thanks to David, I got to know the Bade family, the Costales family, and the Farley family. They provide tremendous help in our wedding preparation and are always welcome me to visit

and stay over at their places, making me feel the warmth of family in the U.S while being so far from my family.

I am extremely grateful to all the people mentioned above and wish them the best of luck in their future endeavors!

COPYRIGHT ACKNOWLEDGEMENTS

The text of this dissertation, in part or in full, is a reprint of the materials as they appear in the following publications:

Chapter 2: 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; 20: e1900347.

Chapter 3: Qi TF, Miao W, Wang Y. Targeted Profiling of Epitranscriptomic Reader, Writer, and Eraser Proteins Accompanied with Radioresistance in Breast Cancer Cells. Anal Chem. 2022; 94: 1525-30.

DEDICATION

To my parents, Yanxiang Li and Guangqi Qi,

who support me with unconditional love;

To my husband, David L. Bade,

who is the best company and supporter.

ABSTRACT OF THE DISSERTATION

Quantitative Proteomic Profiling of TBC Domain-Containing Proteins and Epitranscriptomic Reader, Writer and Eraser Proteins

by

Tianyu Qi

Doctor of Philosophy, Graduate Program in Environmental Toxicology
University of California, Riverside, March 2022
Dr. Yinsheng Wang, Chairperson

Recent advances in mass spectrometry instrumentation and sample preparation methods have enabled robust identification and quantification of proteins at the entire proteome level. The focus of this dissertation is placed on two groups of proteins, the TBC domain-containing proteins as well as epitranscriptomic reader, writer and eraser (RWE) proteins.

In Chapter 2, I utilized a shotgun quantitative proteomic method to assess, at the global proteome scale, differential protein expression in a matched pair of primary/metastatic melanoma cell lines (i.e. WM-115/WM-266-4). I found TBC1D7 may play a role in melanoma cell invasion.

In Chapter 3, I established a liquid chromatography–parallel-reaction monitoring (LC-PRM) method for high-throughput profiling of approximately 150 epitranscriptomic RWE proteins. I employed this LC-PRM method coupled with stable isotope labeling by amino

acids in cell culture (SILAC) to examine the differences in expression levels of the proteins in two matched pairs of radioresistant/wild type (MDA-MB-231/C5 and MCF-7/C6) breast cancer cells. This method allows for the quantifications of 70% and 65% of the epitranscriptomic RWE proteome. Among them, TRMT1 (an $m^{2,2}G$ writer) may assume a crucial role in enhancing breast cancer radioresistance.

In Chapter 4, I further applied this LC-PRM method to assess the expression of epitranscriptomic RWE proteins in modulating colorectal cancer (CRC) metastasis. I was able to quantify 74% of the epitranscriptomic RWE proteome; among them, 48 and 5 were up- and down-regulated by over 1.5-fold in metastatic SW620 relative to primary SW480 CRC cells, respectively.

In Chapter 5, I modified the LC-PRM method by employing a mixture of 48 stable isotope-labeled (SIL) peptides representing RWE proteins as internal or surrogate standards. I utilized this method to explore potential crosstalk between N^6-methyladenosine (m^6A) and other modified ribonucleosides by assessing the epitranscriptomic RWE proteome in *ALKBH5*$^{-/-}$, *FTO*$^{-/-}$, *METTL3*$^{-/-}$ cells, and their isogenic parental HEK293T. *NOP2*, *PUS3*, *TGS1* and *RBMX* were altered by more than 1.5-fold in the opposite directions in *ALKBH5*$^{-/-}$ and *METTL3*$^{-/-}$ cells relative to isogenic HEK293T cells.

Together, the research described in this dissertation documented the power of quantitative proteomics in revealing new functions of cellular proteins in modulating cancer metastasis, radioresistance, and the epitranscriptome.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. Chapter 1. Introduction

1.1 General Overview

The advances in Mass spectrometry (MS)-based proteomics in the past decade provides powerful tools for studying protein identification and quantification, posttranslational modification (PTM) characterization, protein-protein interaction, protein structure and function, and biomarker discovery (1-4). Bottom-up proteomics analyzes proteins that are proteolytically cleaved prior to mass spectrometric analysis. In this chapter, I will discuss common peptide detection methods in bottom-up proteomics, i.e., unbiased and untargeted methods, including data-dependent acquisition (DDA) and data-independent acquisition (DIA), and targeted methods, including selected-reaction monitoring (SRM) and parallel-reaction monitoring (PRM). I will discuss the basic principle, method development, workflow, and comparison among these different methods.

Next, I will describe quantitative proteomics, which encompasses relative and absolute quantifications and involves label-free and labeling-based methods. The overview of MS-based proteomics is summarized in Figure 1.1. Labeling-based quantification methods include metabolic labeling, spike-in of isotope-labeled standards, and isotope-tagging by chemical reactions, where I will elaborate on stable isotope labeling by amino acids in cell culture (SILAC), stable isotope-labeled (SIL) peptide, and tandem mass tags (TMT), respectively. In the end, I will discuss absolute quantification using absolute quantification of proteins (AQUA) and protein standard absolute quantification (PSAQ). I will explain the basic principle, advantages, and disadvantages of each method.

Next, I will move on to introduce the most abundant internal modifications in mRNA, N^6-methyladenosine (m^6A), and its epitranscriptomic reader, writer, and eraser (RWE) proteins. I will also discuss other modified nucleosides in RNA. In the end, I will introduce the scope of the dissertation.

1.2 Bottom-up proteomics

1.2.1 Data-dependent acquisition (DDA)

In shotgun proteomics, also referred to as discovery-mode proteomics, mass spectrometers acquire data in the data-dependent acquisition (DDA) mode (Figure 1.2a). It is an unbiased and untargeted analysis allowing for identifying thousands of proteins in a complex sample (5). In the DDA mode, MS survey scan is initially performed, then based on the intensity acquired from the survey scan, the instrument is programmed to fragment precursor ions with the highest abundances (i.e., top 20) to obtain their MS/MS individually; subsequently, another MS survey scan is conducted, followed by acquiring top N individual MS/MS, where each MS^2 is a snapshot of a specific peptide precursor. The precursor space is sampled by MS^2 discontinuously in neither the RT dimension nor *m/z* dimension. (Figure 1.2b). After data acquisition, MS and MS/MS can be searched using search engines, e.g., MaxQuant and MASCOT (6, 7), against database with protein sequences where peptides are identified based on peptide spectrum matches.

Compared with targeted proteomics, DDA analysis interrogates the full scope of the proteins in the sample, and no information needs to be provided from the researcher to the instrument before data acquisition. However, it lacks sensitivity because DDA only fragments high-abundance precursor ions and leaves low-abundance ones in the MS survey scan unanalyzed. In addition, DDA lacks reproducibility because low-abundance precursor ions sampled in one biological replicate may not be selected for fragmentation in the second biological replicate, resulting in large variations among replicates. It is worth noting that DDA-based quantifications are based on extracted-ion chromatograms (XICs) from MS scan, which may be more susceptible to interference of co-eluting ions than quantification based on MS/MS.

In bottom-up proteomics, traditional proteolytic digestion is performed either “in-gel” or “in-solution” using trypsin (8). Both digestion methods have disadvantages, where in-gel digestion may suffer from significant sample loss and low peptide recovery, and in-solution digestion may be incomplete owing to interferences from sample matrices (9). To overcome these disadvantages, Wis´niewski *et al.* (9) introduced filter-aided sample preparation (FASP). FASP was employed to examine the cleanup efficiency of sample solubilized in 4% sodium dodecyl sulfate (SDS). The successful removal of residual SDS during buffer exchange in a filter-based device is crucial because the presence of SDS, even at low amounts, affects trypsin digestion efficiency and interferes with peptide quantification due to the high ionization efficiency of SDS.

1.2.2 Selected-reaction monitoring (SRM)

Western blot analysis for target protein quantification has been used for decades. It is highly sensitive, but suffers from low throughput, and potential antibody specificity and availability issues. In comparison, MS-based targeted proteomics has the advantages of high throughput, where up to hundreds of peptides can be quantified in one LC-MS/MS run, and high specificity, where unique tryptic peptides representing targeted proteins can be quantified.

Targeted proteomics can be a better strategy if researchers have a pre-defined list of proteins of interest since it offers better sensitivity, selectivity, and reproducibility than DDA. Among targeted proteomic approaches, selected-reaction monitoring (SRM), also referred to as multiple-reaction monitoring (MRM), is known for its highly sensitive and reproducible peptide measurement (10). SRM is conducted on a triple-quadrupole mass spectrometer (Figure 1.2a). The first (Q1) and third quadrupole (Q3) serve as mass filters for precursor ion and fragment ion selection, respectively, whereas the second quadrupole (Q2) operates in radio frequency (RF)-only mode and serves as a collision cell for fragmentation of precursor ions. Scheduled SRM is programmed to target a pre-defined list of peptides of interest, including the information of precursor *m/z*, fragment *m/z*, collision energy, and potential elution time. Each MS^2 is a sampling of one fragment ion in Q3. Precursor space is sampled by MS^2 signals continuously in the RT dimension, but not in *m/z* dimension (Figure 1.2b). Because SRM is conducted in a quadrupole, the resolution is lower than data acquired in a high-resolution mass analyzer, e.g., Orbitrap. The complexity of the sample usually needs to be decreased through sample fractionation (11).

In order to establish an SRM library, three types of information are required: selection of target proteins, selection of peptides to represent target proteins, and selection of transitions (Figure 1.3) (12). Unlike discovery proteomics, the selection of target proteins in SRM is driven by research hypothesis (13). The selection of peptides can be deduced from discovery proteomics in-house or from data repositories. Two major criteria need to be considered: (a) Peptide should be unique from a specific protein for high selectivity. Online tools, such as neXtProt - Peptide uniqueness checker (https://www.nextprot.org/viewers/peptide-uniqueness-checker/app/index.html) (14), can be used to ensure the uniqueness of the peptides. (b) Three to five peptide precursor ions of a single protein with the maximum intensities should be selected for high sensitivity. The selection of precursors can be determined by mining previously acquired discovery proteomic data in-house or from data repositories, including PeptideAtlas (www.peptideatals.org/), ProteomeXchange (http://www.proteomexchange.org/), and PRIDE (https://www.ebi.ac.uk/pride/). To use PeptideAtlas (15) as an example, upon searching a protein name, a table titled "Distinct Observed Peptide" is exhibited where peptide sequences are displayed in descending order of empirical suitability score. Top sequences on the list representing the targeted protein, which have been frequently detected in different experiments and instruments from separate studies, may be selected to monitor in the SRM experiment. It is worth noting that the top sequences on the list does not guarantee the highest ion intensities on the triple quadrupole instrument to conduct the SRM experiment. Therefore, it is more reliable to obtain proteomic data in-house using the same instrument as conducting SRM experiment than from public data repositories. In addition to the above-described empirical data, peptides can also be selected from prediction algorithms, for instance, ESP predictor (https://www.genepattern.org/esppredictor) (16). Prediction algorithms are less accurate and it can be used as the last resort when empirical data are not available.

In addition to two primary criteria mentioned above, below are the other criteria when selecting tryptic digested peptides to establish an SRM library (17): (a) Peptide length should maintain 7-25 amino acids; (b) Avoid missed cleavage sites of trypsin; (c) Avoid two enzymatic cleavage sites adjacent to each other, such as KK, KR, RR, or RK; (d) Avoid D/E adjacent to the cleavage site; (e) Avoid P after K/R, such as KP or RP; and (f) Avoid frequently modified amino acids, for instance, M and W, which are frequently oxidized.

To select three-to-five transitions to represent each precursor, the transition needs to be unique and exhibits the highest intensities. For a peptide encompassing ten amino acids, fragment ions y_9 and y_8 are more likely to be unique and reliable than y_1 and y_2. Therefore, if they display the same intensity, the selection of y_9 and y_8 to represent the precursor may be a better choice than y_1 and y_2. To obtain the highest intensities of fragment ions from each peptide precursor, researchers can use the MS/MS obtained from discovery proteomics generated in-house or from public data deposited to online repositories. In addition to the aforementioned online repositories, other databases containing SRM data such as SRMAtlas (www.srmatlas.org) and PASSEL (www.peptideatlas.org/passel/) can also be used.

In summary, SRM is a highly sensitive and reproducible targeted proteomic method that can be used to profile tens-to-hundreds of targeted proteins. The disadvantage of SRM is that it can be labor-sensitive since prior knowledge of proteins is needed. Additionally, samples may need to be fractionated to enhance signal-to-noise ratio due to the low-resolution data obtained from quadrupole. However, once the SRM method is established, it may be efficiently applied to other research projects without extensive further optimization.

1.2.3 Parallel-reaction monitoring (PRM)

Parallel-reaction monitoring (PRM) is another MS-based proteomic method, also known as targeted MS/MS (tMSMS), MRM-HR, pseudo selected reaction monitoring (pSRM) or targeted full-scan MS/MS. Distinct from MRM, it is usually performed on a high-resolution and accurate-mass (HRAM) mass analyzer such as Orbitrap or time of flight (TOF), coupled with quadrupole or ion trap for precursor ion selection (Figure 1.2a) (18). The major difference between PRM and SRM is that HRAM full scan MS/MS (<10 ppm) are obtained in PRM; on the contrary, low resolution (0.7 *m/z*) discrete transitions are measured in SRM. PRM has a significantly reduced time in method development because a full-scan MS/MS is acquired for all fragment ions of a peptide precursor ion, and it does not require *a priori* information of fragment ions as SRM does.

Without the need to indicate fragment ions *m/z*, only predefined *m/z* values of precursor ions need to be included in the PRM method. This information can be retrieved from previous shotgun proteomics obtained in-house or from public data repositories, similar to the processes described in the SRM section. While operating a quadrupole-Orbitrap mass spectrometer in the PRM mode, precursor ions are selected in quadrupole, transferred through the C-trap, and injected into the higher-energy collisional dissociation (HCD) cells for fragmentation. The ensuing fragment ions are sent back to the C-trap for focusing and are injected into the Orbitrap for high-resolution mass measurement. Each MS^2 is sampled from one targeted precursor ion at one point in time, and MS^2 signal of each precursor is acquired over time across the chromatogram. Similar to SRM, precursor space is sampled by MS^2 signals continuously in the RT dimension, but not in *m/z* dimension (Figure 1.2b).

The goal of the targeted proteomics is to achieve the best possible sensitivity, accuracy, and as many number of precursors as possible. However, there is a trade-off among these three parameters. Dwell time or accumulation time is the time given the instrument to accumulate signal for a specific transition (for SRM) or for obtaining full-scan MS/MS of a specific precursor (for PRM). The longer the dwell time, the higher the signal-to-noise ratio. Cycle time is the sum of the dwell time to loop through the transition list (for MRM) or the precursor list (for PRM) (Figure 1.2b). It is also an indication of the sampling rate across the chromatographic peak. For instance, for a 30-second chromatogram, if 15 points or more are preferable to obtain accurate quantification, the cycle time should not exceed 2 seconds. Therefore, the shorter the cycle time, the more accurate the quantification is. The number of scanned transitions (for MRM) or precursors (for PRM) equals cycle time divided by dwell time. In order to maximize the number of transitions or precursors to scan among tens-to-hundreds of peptides in a target list, without reducing dwell time (i.e., sacrificing sensitivity) or increasing cycle time (i.e., sacrificing quantification accuracy), RT scheduling is necessary.

The determination of normalized retention time (iRT) (19) of each peptide precursor is fundamental for RT scheduling. After knowing the elution time of the peptide precursor in discovery proteomics in-house, iRT of each peptide precursor can be obtained from the correlation between ten predefined iRT of bovine serum albumin (BSA) tryptic peptides and their corresponding RT. Upon determining iRT of each peptide, peptide precursors are monitored in a scheduled retention time window. Therefore, the number of concurrent precursor ions is reduced at any point in time, allowing many more targets to be analyzed in a single LC run at a given dwell and cycle time.

Skyline (www.skyline.ms) is a freely available, open-source software suitable for targeted proteomic analysis (20). SRM or PRM library of targeted peptides, containing spectrum library and iRT information of each peptide, is first developed using Skyline. Subsequently, the isolation list of PRM or transition list of SRM can be automatically exported from Skyline as an inclusion list for the LC-PRM or LC-SRM method. After data acquisition, the obtained raw data are imported to Skyline for peak integration.

For confident SRM or PRM peak identification, several criteria can be employed (Figure 1.4): (a) Transitions of a specific precursor should co-elute and exhibit similar peak shape; (b) If a heavy isotope-labeled standard is employed for an analyte of interest, the heavy form and its light counterpart should co-elute; (c) dot plot (dotp) value is employed to gauge the similarity of the fragment ion distribution between acquired MS/MS and reference MS/MS in the spectral library (21). The closer dotp value to 1, the more these two MS/MS resemble. Dotp value of 0.7 is often used as a cutoff, but 0.9 is preferable. It should be noted that each transition of SRM or PRM needs to be examined by the researcher to rule out the use of potential interfering ions for quantification. The most frequently identified interfering ions are y_1 ion since it is less selective. Moreover, PRM is acquired on a high-resolution mass analyzer whose mass accuracy is less than 10 ppm. Transitions with mass accuracy larger than 20 ppm are usually not reliable.

In conclusion, both SRM and PRM are highly accurate, reproducible, and have a wide dynamic range of four to five orders of magnitude (22). SRM excels in its high sensitivity, whereas PRM offers high selectivity. SRM usually requires prefractionation to reduce sample complexity, such as using SDS-PAGE to separate and excise desired bands of interest or using isoelectric focusing, because it is conducted in a low-resolution quadrupole mass analyzer. However, because of its high selectivity, PRM can deal with samples with high complexity, such as whole-cell lysate or plasma (23, 24). In addition, PRM is less labor-extensive compared with MRM because fragment ion information does not need to be determined *a priori*.

1.2.4 Data-independent acquisition (DIA)

Data-independent acquisition (DIA) or sequential window acquisition of all theoretical mass spectra (SWATH) (25) has gained increasing attention in the scientific field in the past decade. Unlike DDA, which is instrument-driven and biased toward detecting highly abundant species, DIA measures all fragment ions, even the least abundant species, for all precursors in a defined relatively large isolation window (for instance, 20 *m/z*). DIA provides untargeted, unbiased, and consistent acquisition of MS/MS for all analytes across chromatographic time scale. In theory, DIA offers complete peptide identification in a given sample. However, it also results in a highly complicated MS/MS that contains all fragment ions from different precursors in the same isolation window (26).

There are two major ways to identify peptides from DIA data. The most traditional one is peptide-centric analysis (25). A reference spectrum library can be established *a priori* after learning the precursor and fragment ion information of specific peptides from shotgun proteomic analysis. Similar to PRM data analysis, the acquired DIA MS/MS are subsequently compared with the established spectral library for peptide identification. This peptide-centric analysis can be carried out in open-source software such as OpenSWATH or Skyline (20, 27). The disadvantage of this method is that peptides not included in the reference library cannot be identified. The other way is spectrum-centric analysis. By using this spectral library-free approach, DIA MS/MS can be de-convoluted to pseudo-MS^2 spectra using scoring algorithms such as DIA-Umpire or Group-DIA (28, 29). The de-convoluted spectra can be searched against traditional protein sequence database similar to DDA data analysis.

DIA is usually conducted on a hybrid instrument where a mass-selective quadrupole is followed with a high-resolution mass analyzer such as TOF or Orbitrap (Figure 1.2a). DIA continuously samples in both RT dimension and *m/z* dimension, which allows for a full coverage of the precursor ion space (Figure 1.2b) (30). The sensitivity of DIA is good, slightly worse than SRM but better than DDA (25). The precursor selectivity of DIA is lower than DDA or SRM/PRM since it has a much larger isolation window and MS/MS reflects co-fragmented precursor ions. Like SRM/PRM, DIA offers accurate and reproducible peptide quantification. In addition, the most significant feature of DIA is its flexibility of data re-analysis or re-mining since information of all the peptides in a given sample are documented in DIA data. In comparison, SRM/PRM is hypothesis-driven; data need to be re-acquired when the hypothesis is changed.

In summary, DIA offers unbiased, untargeted, comprehensive, sensitive, accurate and reproducible quantification of peptides, which combines the merits of DDA and

SRM/PRM analysis. However, one major drawback of DIA is that it produces a complex mixture of MS/MS, which requires a relatively involved peptide-centric analysis or spectrum-centric analysis. DDA, SRM, PRM, or DIA should be chosen based on research goals.

1.3 Quantitative Proteomics

1.3.1 Relative quantification

1.3.1.1 Label-free quantification

MS-based quantitative proteomics usually provides two types of quantification information, i.e., relative and absolute quantification, with the former being more widely used. MS-based quantitative proteomics has two underlying methodologies: label-free and label-based.

The label-free approach is cost- and work-effective since samples are analyzed separately without the need for sample multiplexing or the addition of stable isotope-labeled peptides (Figure 1.5a). It does not increase sample complexity compared with label-based method. However, label-free approach may suffer from quantitative inaccuracy and inconsistency because many variations can be introduced among samples in different sample preparation processes, such as proteolytic digestion, cleanup, or variations in different LC-MS/MS runs. Therefore, careful sample preparation and stable chromatography and mass spectrometer performance are critical in obtaining satisfying label-free quantification results. In addition, the quantification results from the label-free method may be normalized using the ion intensities of peptides from a specific protein encoding the housekeeping gene or the total ion intensities in a given sample (31).

1.3.1.2 Label-based quantification

Stable isotope-labeling approach introduces a heavy isotope-labeled counterpart, namely, ^{13}C, ^{15}N, or ^{18}O, to its endogenous target in a given sample. The heavy form is chemically identical to its light form. Therefore, both forms exhibit the same behavior in sample preparation, chromatography, ionization, and fragmentation. Isotope labeling-based quantifications, including metabolic labeling, spike-in of isotope-labeled standards, and isotope-tagging by chemical reactions, improve quantitative accuracy and producibility compared with label-free quantification.

Stable isotope labeling by amino acids in cell culture (SILAC) is the most frequently used metabolic labeling method (Figure 1.5b) (32). In SILAC, heavy isotope-labeled amino acids, such as ^{13}C- and ^{15}N-labeled lysine and arginine, are added to the cell culture medium for several cell doubling times until they are nearly fully incorporated into proteins through translation. Lysates of cells grown in heavy- and light-medium, from two biological states, are mixed at 1:1 ratio by mass. Upon tryptic digestion, nearly all the peptide precursor ions in heavy-medium are labeled because they bear ^{13}C- and ^{15}N-labeled lysine and arginine, except for some of the C-terminus peptides. Since the chemical properties of the heavy- and light-labeled forms are the same, both forms should have the same tryptic digestion efficiency and LC-MS/MS performance, namely with the same retention time and the same ionization efficiency, but can be distinguished by mass differences in MS and MS/MS.

The ratio of peak intensities between heavy and light isotope-labeled peptides in the spectrum, within the linear dynamic range of the peptide, reflects the difference in the abundance of a particular protein between the two sample sets. High precision and accuracy are distinct benefits of SILAC because heavy and light isotope-labeled proteins are mixed at the very early stage of the sample preparation process (Figure 1.5b), which reduces experimental and analytical inconsistencies among samples. However, SILAC has limitations for samples that are not amenable to metabolic labeling, such as biofluid or tissues. In addition, it significantly increases sample complexity by combining light- and heavy-labeled cells, which reduces sensitivity.

In targeted proteomics, stable isotope-labeled (SIL) standards, such as tryptic peptide bearing ^{13}C- and ^{15}N- labeled lysine and arginine of targeted proteins, can be synthesized and spiked into samples post tryptic digestion (Figure 1.5c) (33). This method increases the confidence of peptide identification because the SIL standard and the peptide of interest share the same elution time (Figure 1.4c). In addition, this method does not increase the complexity of the sample as much as SILAC does. However, since SIL standards are usually introduced post proteolytic digestion, it can only adjust analytical variations. In addition, even though the purity of the SIL peptide does not need to be high (>70% in peptide purity and >99% in isotopic purity), it still can be costly, especially when a large number of peptides need to be synthesized.

Tandem mass tags (TMT) labeling is one of the most prevalent chemical labeling techniques for tryptic peptides (Figure 1.5d) (34). TMT reagents have three groups: mass reporter, mass normalizer, and NH_2-reactive group (Figure 1.6a). Different combinations of stable heavy isotopes (^{13}C and/or ^{15}N) are introduced to different positions of mass reporters and mass normalizers (Figure 1.6b). NH_2-reactive group reacts with lysine residues and peptide N-termini. TMT reagents offer isobaric tags to the peptide but yield mass reporters carrying different masses upon fragmentation in mass spectrometers. In this context, TMT-labeled peptides are quantified based on the signal intensities of mass reporter ions in the MS/MS. It should be noted that TMT reporter ions can only be generated from HCD or electron transfer dissociation (ETD), but not from collision-induced dissociation (CID) (35).

TMT was first introduced as TMT duplex in 2003 (34). Since then, the multiplexing capacity of TMT reagents has been expanded to 6plex, 10plex, 11plex, even pro16plex (36, 37). The workflow of TMT-labeling is illustrated in Figure 1.6c. The higher level of multiplexing requires a high-resolution mass spectrometer, for instance, a resolving power of 50,000 at *m/z* 130, to distinguish ^{13}C from ^{15}N labeling (38). Different from quantification of PRM data using extracted-ion chromatograms (XICs) to integrate MS^2 signal intensity across the chromatographic elution time of a peptide precursor, quantification of TMT-labeled peptides employs reporter ion signal intensities at one point in time in the MS/MS spectrum.

Due to the high level of multiplexing, TMT labeling enables quantifications across multiple samples in a single LC-MS/MS run. TMT labeling is usually coupled with pre-fractionation methods, such as high-pH reversed-phase fractionation (39), prior to LC-MS/MS, to reduce sample complexity and improve proteome coverage. One major drawback of TMT labeling is ratio compression through impure MS precursor isolation, which may reduce quantification accuracy (40). This ratio compression issue can be improved by further fragmentation to MS^3 or employing a narrower precursor isolation window (41, 42).

1.3.2 Absolute quantification

Absolute quantification determines the absolute amount of proteins in a mixture, which can be employed to address research questions, such as protein complex stoichiometries, sports doping testing, and biomarker development (43-45). Label-based absolute quantification is more accurate and frequently used than label-free absolute quantification. Employing peptides or proteins labeled with heavy stable isotopes is commonly seen in label-based absolute quantification.

In targeted proteomics, heavy isotope-labeled absolute quantification (AQUA peptide) mimicking a tryptic peptide of targeted protein can be synthesized and added an absolute amount to the sample as an internal standard (46). The heavy isotope can be incorporated into AQUA peptides not only through lysine or arginine residue on the C-termini, but at another amino acid for specific PTM studies (47). The same chemical properties between AQUA peptide and its native form ensure that they have the same retention time, ionization efficiency, and fragmentation in an LC-SRM experiment. After a calibration curve of peak intensity versus peptide concentration is generated using a dilution series of isotope-labeled reference peptides (48), the absolute amount of the native peptide can be derived.

Since AQUA peptide is introduced after proteolytic digestion, it can correct for variations among samples in LC-MS/MS performance, but not in sample preparation procedures. In addition, the accuracy of AQUA peptide quantification relies on the assumption that proteins are completely extracted from the sample and entirely digested with trypsin. In reality, it is inevitable that global protein losses from incomplete cell lysis or local protein losses from low solubility of membrane proteins. In order to ensure that the measured peptide level faithfully represents the protein level, an optimal quantotypic peptide should be selected as an AQUA peptide (49). The selection of quantotypic peptide from a specific protein has a stricter rule, for instance, ensuring complete digestion, than that of SIL peptide used in relative quantification. Moreover, AQUA peptides may result in losses due to adsorption to surfaces and degradation through freezing-thawing cycles. Therefore, handling AQUA peptides with caution is crucial, such as using low-binding tubes and pipet tips, and aliquot stock solution to avoid freezing-thawing cycles.

In addition to the spike-in of heavy isotope-labeled peptides, heavy isotope-labeled proteins can also be used for absolute quantification (50). Protein standard absolute quantification (PSAQ) method involves adding the *in vitro*-synthesized full-length isotope-labeled recombinant protein to the sample as an internal standard before trypsin digestion. PSAQ could provide more accurate quantification results than AQUA peptide since it accounts for digestion efficiency and protein loss during cleanup. However, it is worth noting that neither PSAQ nor AQUA can fully represent the native protein or peptide because the endogenous one may have PTMs (51).

Label-free absolute quantification is based on spectral counting (52), and precursor or fragment ion intensity (53). Several methods have been developed, for instance, absolute protein expression (APEX) and “Top 3” (54, 55). Label-free absolute quantification suffers from low accuracy and low reproducibility; therefore, it is much less reliable compared with label-based absolute quantification.

In summary, absolute quantification using either heavy isotope-labeled peptide or protein is more challenging to achieve accurate and reliable quantification than relative quantification. Major challenges, including incomplete protein extraction, incomplete protein digestion, and protein loss during sample preparation, needs to be addressed carefully.

1.4 Epitranscriptomic reader, writer, eraser (RWE) proteins

1.4.1 m^{6}A and its RWE proteins

Epigenetic modifications of DNA and histones have been well studied; however, less is known about the roles of RNA modifications in cellular processes. More than 170 types of chemical modifications have been identified in multiple cellular RNA species (56). The most abundant internal modification in eukaryotic mRNA is m^{6}A, first discovered in 1974 (57-60). METTL3 was uncovered in bacteria, and its function was demonstrated as an m^{6}A methyltransferase in 1997 (61). However, it was not until 2012, when advances in high-throughput sequencing made transcriptome-wide m^{6}A profiling achievable (62, 63), this modification started to attract much more attention in the scientific community. In recent years, the field of epitranscriptomics has been quickly evolving, and publication is seen on a monthly basis.

Transcriptome-wide m^{6}A site mapping studies revealed that m^{6}A is primarily localized near stop codon in 3′ untranslated region (3′UTR) (63). m^{6}A is selectively enriched in some mRNAs (64). Most m^{6}A-containing mRNAs have only one m^{6}A site, where the consensus sequence is DR(m^{6}A)CH (D = A, G, or U; R = G or A; H = A, C or U) (63).

m^6A in mRNA is primarily installed by a methyltransferase complex or "writer complex", which consists of a heterodimer of a catalytic subunit METTL3 and its allosteric activator METTL14 (65), together with WTAP, ZC3H13, RBM15/15B, and VIRMA (66-69). METTL16 is another m^6A writer in pre-mRNAs and non-coding RNAs (70). METTL5-TRMT112 complex and ZCCHC4 are m^6A writers on rRNAs (71, 72). m^6A modification is reversible, which can be removed by a demethylase or "eraser", FTO or ALKBH5 (73, 74). A later study demonstrated that FTO selectively demethylates N^6,2′-O-dimethyladenosine (m^6A_m) over m^6A at a rate of 100 times higher *in vitro* (75), rendering ALKBH5 the only well-accepted m^6A eraser in mRNA in the scientific community. ALKBH5 is known to impact RNA metabolism and mouse fertility (76).

m^6A-binding proteins or "readers" exerts regulatory roles in mRNA splicing, export, stability, and translation (Figure 1.7). YTH domain-containing proteins (i.e., YTHDF1, YTHDF2, YTHDF3, YTHDC1, and YTHDC2) are direct binders to m^6A, where YTHDF1, YTHDF2, and YTHDF3 bind to m^6A in the cytoplasm, whereas YTHDC1 and YTHDC2 bind to m^6A in the nucleus. The roles of YTHDF1, YTHDF2, and YTHDF3 had been demonstrated to regulate mRNA translation efficiency, mRNA degradation, and both mRNA translation efficiency and degradation, respectively (77-79), until another study reported conflicting findings that YTHDF1, YTHDF2, and YTHDF3 work redundantly on mRNA degradation, and their functions of regulating mRNA stability are obvious when they are depleted simultaneously (80). YTHDC1 regulates mRNA splicing and mRNA export in the nucleus (81, 82). YTHDC2 is primarily expressed in testes and assumes important roles in spermatogenesis (83).

Other m^6A readers translate m^6A mark into different functions. heterogeneous nuclear ribonucleoproteins, hnRNPA2B1 and hnRNPC, regulate mRNA splicing and primary microRNA processing in the nucleus (84, 85). Insulin-like growth factor 2 mRNA binding proteins, IGF2BP1/2/3, bind to the m^6A-modified transcripts enhancing their stabilities in the nucleus (86). Eukaryotic translation initiation factor 3 (eIF3) binds to the m^6A sites on a 5′ untranslated region (5′UTR) to promote cap-independent translation (87).

Three models have been established to explain how m^6A promotes translation efficiency (Figure 1.7). The first model is through m^6A reader protein YTHDF1 (77). YTHDF1 interacts with translation initiation factor complex 3 (eIF3), a component of newly-assembled 43S pre-initiation complex. eIF3 is associated with eIF4G, which binds both cap-binding protein eIF4E at the 5′-end and poly(A) binding protein at the 3′-end. eIF4G, as a bridge of 5′- and 3′-end, forms a loop structure of mRNA. Therefore, YTHDF1 was proposed to regulate the translation initiation complex through the binding to eIF3 and the loop structure formed through eIF4G. The detailed mechanism of the loop structure still needs further validation (77). Another model was also a mRNA-looping model, but through a direct interaction between METTL3, bound to m^6A at the 3′-end, and eIF3h, bound to cap-binding proteins at the 5′-end (88). This loop structure is further substantiated by the proximity of METTL3 to polyribosome and 5′ cap-binding proteins in the cytoplasm using electron-microscopy (88). METTL3, a well-known catalytic subunit of m^6A writer complex located in the nucleus, is shown here to serve as an m^6A reader in the cytoplasm. The third model is called cap-independent translation, which involves direct binding of eIF3 to m^6A at the 5′UTR upstream of the start codon (87). This m^6A-mediated translation does not require the recruitment of eIF3 to eIF4E at the 5′cap, which is crucial for cap-dependent translation. This mechanism was proposed to occur upon stress induction in the m^6A site located at 5′UTR of mRNAs, such as mRNAs encoding heat shock proteins.

In addition to regulating mRNA metabolism, many studies have investigated aberrant regulation of m^6A in cancer development. For instance, m^6A writer complex component METTL3, overexpressed in gastric cancer, promotes the epithelial-to-mesenchymal transition and metastatic transformation of gastric cancer through regulating the stability of *ZMYM1* mRNA (89). In addition, overexpression of m^6A eraser ALKBH5 promotes the development of acute myeloid leukemia (AML) and self-renewal of leukemia stem/initiating cells maintenance by modulating the stability of *TACC3* mRNA (90). Moreover, m^6A readers IGF2BP1/2/3 play oncogenic roles in cancer cells by enhancing mRNA stability of *MYC* through recruiting mRNA stabilizers ELAVL1 and MATRIN3 (86). In summary, aberrant expression of m^6A RWE proteins may result in cancer tumorigenesis or progression through affecting mRNA stabilities or expression of cancer-related genes (91).

1.4.2 Other nucleoside modifications in RNA

Besides the most well-studied m^6A, more than 170 nucleoside modifications are reported to date (92). Transfer RNA (tRNA) has the most dynamic modifications among all RNA species, with an average of 13 modifications per molecule in eukaryotic cells (93). In this vein, other major modified nucleosides in mRNA include 7-methylguanosine (m^7G), 2′-*O*-methylation N_m, 5-methylcytosine (m^5C), N^1-methyladenosine (m^1A), N^4-acetylcytidine (ac^4C), inosine (I), pseudouridine (Ψ), and 5-hydroxymethylcytosine (hm^5C) (Figure 1.8).

Among them, m^1A regulates RNA folding and stability, ribosome biosynthesis, and translation (94-96). m^5C modulates the export (97), stability (98), and translation of mRNA (99). Ψ, primarily located in tRNA and rRNA (100), is the most abundant internal modification in cellular RNA, and it affects RNA structure and translation (101, 102). Moreover, ribosome can read through Ψ in the stop codon through unusual base-pairing with tRNA, thereby modulating mRNA coding (103). RNA modifications studies focused not only on understanding the aforementioned biological regulations, but also on its association with cancer. Many RWE proteins have been known to promote or inhibit the hallmarks of cancer (104).

1.5 Scope of the dissertation

Melanoma is the deadliest type of skin cancer because of its strong tendency to metastasize. We set out to identify new protein players that may drive or suppress melanoma metastasis. In Chapter 2, we utilized an unbiased mass spectrometry-based quantitative proteomic method to assess differential protein expression in a matched pair of primary/metastatic melanoma cell lines (i.e., WM-115/WM-266-4) derived from the same patient. We found that TBC1D7 is overexpressed in metastatic over primary melanoma cells, and elevated expression of TBC1D7 promotes the invasion of these melanoma cells *in vitro*, partly through modulating the activities of secreted matrix metalloproteinases 2 and 9. Additionally, interrogation of publicly available data showed that higher mRNA expression of TBC1D7 predicts poorer survival in melanoma patients. Together, our results suggest TBC1D7 as a driver for melanoma cell invasion, which is an essential element in melanoma metastasis.

Epitranscriptomic RWE proteins recognize, install, and remove modified nucleosides in RNA, which play crucial roles in RNA export, splicing, translation, and stability. In Chapter 3, we established an LC-PRM method, for the first time, for high-throughput profiling of a total of 152 epitranscriptomic RWE proteins. We also applied the LC-PRM method, in conjunction with SILAC, to quantify these proteins in two pairs of matched parental/radioresistant breast cancer cells (i.e., MDA-MB-231 and MCF-7 cells, and their corresponding radioresistant C5 and C6 clones), with the goal of assessing the roles of these proteins in radioresistance. We found that eight epitranscriptomic RWE proteins were commonly altered by over 1.5-fold in the two pairs of breast cancer cells. Among them, TRMT1 (an $m^{2,2}G$ writer) may play a role in promoting breast cancer radioresistance due to its clinical relevance and its correlation with DNA repair gene sets.

Aberrant expressions of some epitranscriptomic RWE proteins are associated with cancer initiation and progression. In Chapter 4, we employed the PRM-based targeted proteomic method, in conjunction with SILAC, to comprehensively examine the differential expression of epitranscriptomic RWE proteins in a matched pair of primary/metastatic CRC cells (i.e., SW480/SW620). We were able to quantify 113 non-redundant epitranscriptomic RWE proteins; among them, 48 and 5 were up- and down-regulated by at least 1.5-fold in SW620 over SW480 cells, respectively. Particularly, NAT10, hnRNPC, and DKC1 were markedly up-regulated in metastatic CRC cells, and the potential roles of these proteins in driving CRC metastasis were documented in recent studies. Interrogation of the Clinical Proteomic Tumor Analysis Consortium (CPTAC) data revealed that the elevated expressions of these and several other RWE proteins are also accompanied with CRC initiation, suggesting the dual roles of these proteins in the initiation and metastatic transformation of CRC.

To our knowledge, there have been no systematic investigations about the crosstalk between m^6A and other modified nucleosides in RNA. In Chapter 5, we modified the LC-PRM method by employing a mixture of 48 SIL peptides representing 45 RWE proteins as internal or surrogate standards for profiling epitranscriptomic RWE proteins. We were able to reproducibly detect a total of 114 RWE proteins in HEK293T cells with the genes encoding m^6A eraser proteins (i.e., *ALKBH5*, *FTO*) and the catalytic subunit of the m^6A writer complex (i.e., *METTL3*) being individually ablated. Notably, eight proteins were altered by more than 1.5-fold in the opposite directions in *ALKBH5*$^{-/-}$ and *METTL3*$^{-/-}$ cells relative to isogenic HEK293T cells. Bioinformatic analysis of published m^6A mapping results revealed the presence of m^6A in the mRNAs of four of these genes, namely, *NOP2*, *PUS3*, *TGS1*, and *RBMX*. We are in the process of interrogating if the differential expression of these proteins emanates from the altered stabilities of their transcripts.

We envision that the LC-PRM method reported in this dissertation is applicable for studying the roles of the epitranscriptomic RWE proteins in the metastatic transformation and therapeutic resistance of other cancer types. In conjunction with the use of SIL peptides, the LC-PRM is applicable to study epitranscriptomic RWE proteins in biofluid or tissues.

References

1. Zhang Y, Fonslow BR, Shan B, Baek MC, Yates JR, 3rd. Protein analysis by shotgun/bottom-up proteomics. Chemical reviews. 2013;113:2343-94.

2. Aebersold R, Mann M. Mass-spectrometric exploration of proteome structure and function. Nature. 2016;537:347-55.

3. Zhou W, Petricoin EF, 3rd, Longo C. Mass Spectrometry-Based Biomarker Discovery. Methods Mol Biol. 2017;1606:297-311.

4. Low TY, Syafruddin SE, Mohtar MA, Vellaichamy A, NS AR, Pung YF, et al. Recent progress in mass spectrometry-based strategies for elucidating protein-protein interactions. Cellular and molecular life sciences : CMLS. 2021;78:5325-39.

5. Cravatt BF, Simon GM, Yates JR, 3rd. The biological impact of mass-spectrometry-based proteomics. Nature. 2007;450:991-1000.

6. 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;26:1367-72.

7. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis. 1999;20:3551-67.

8. Shevchenko A, Wilm M, Vorm O, Mann M. Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. Anal Chem. 1996;68:850-8.

9. Wisniewski JR, Zougman A, Nagaraj N, Mann M. Universal sample preparation method for proteome analysis. Nat Meth. 2009;6:359-62.

10. Peterson AC, Russell JD, Bailey DJ, Westphall MS, Coon JJ. Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics. Molecular & cellular proteomics : MCP. 2012;11:1475-88.

11. Shi T, Su D, Liu T, Tang K, Camp DG, 2nd, Qian W-J, et al. Advancing the sensitivity of selected reaction monitoring-based targeted quantitative proteomics. Proteomics. 2012;12:1074-92.

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

13. Picotti P, Bodenmiller B, Aebersold R. Proteomics meets the scientific method. Nat Meth. 2013;10:24-7.

14. Schaeffer M, Gateau A, Teixeira D, Michel PA, Zahn-Zabal M, Lane L. The neXtProt peptide uniqueness checker: a tool for the proteomics community. Bioinformatics (Oxford, England). 2017;33:3471-2.

15. Desiere F, Deutsch EW, King NL, Nesvizhskii AI, Mallick P, Eng J, et al. The PeptideAtlas project. Nucleic Acids Res. 2006;34:D655-D8.

16. Fusaro VA, Mani DR, Mesirov JP, Carr SA. Prediction of high-responding peptides for targeted protein assays by mass spectrometry. Nat Biotechnol. 2009;27:190-8.

17. Hoofnagle AN, Whiteaker JR, Carr SA, Kuhn E, Liu T, Massoni SA, et al. Recommendations for the Generation, Quantification, Storage, and Handling of Peptides Used for Mass Spectrometry-Based Assays. Clin Chem. 2016;62:48-69.

18. 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;11:1475-88.

19. 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.

20. 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 (Oxford, England). 2010;26:966-8.

21. de Graaf EL, Altelaar AFM, van Breukelen B, Mohammed S, Heck AJR. 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.

22. Dong M, Lih T-SM, Höti N, Chen S-Y, Ponce S, Partin A, et al. Development of Parallel Reaction Monitoring Assays for the Detection of Aggressive Prostate Cancer Using Urinary Glycoproteins. J Proteome Res. 2021;20:3590-9.

23. Brzhozovskiy A, Kononikhin A, Bugrova AE, Kovalev GI, Schmit PO, Kruppa G, et al. The Parallel Reaction Monitoring-Parallel Accumulation-Serial Fragmentation (prm-PASEF) Approach for Multiplexed Absolute Quantitation of Proteins in Human Plasma. Anal Chem. 2022;94:2016-22.

24. 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.

25. Gillet LC, Navarro P, Tate S, Röst H, Selevsek N, Reiter L, et al. Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. Mol Cell Proteomics. 2012;11:O111.016717.

26. Bern M, Finney G, Hoopmann MR, Merrihew G, Toth MJ, MacCoss MJ. Deconvolution of mixture spectra from ion-trap data-independent-acquisition tandem mass spectrometry. Anal Chem. 2010;82:833-41.

27. Röst HL, Rosenberger G, Navarro P, Gillet L, Miladinović SM, Schubert OT, et al. OpenSWATH enables automated, targeted analysis of data-independent acquisition MS data. Nat Biotechnol. 2014;32:219-23.

28. Tsou C-C, Avtonomov D, Larsen B, Tucholska M, Choi H, Gingras A-C, et al. DIA-Umpire: comprehensive computational framework for data-independent acquisition proteomics. Nat Meth. 2015;12:258-64.

29. Li Y, Zhong CQ, Xu X, Cai S, Wu X, Zhang Y, et al. Group-DIA: analyzing multiple data-independent acquisition mass spectrometry data files. Nat Methods. 2015;12:1105-6.

30. Egertson JD, MacLean B, Johnson R, Xuan Y, MacCoss MJ. Multiplexed peptide analysis using data-independent acquisition and Skyline. Nature Protocols. 2015;10:887-903.

31. Wang W, Zhou H, Lin H, Roy S, Shaler TA, Hill LR, et al. Quantification of proteins and metabolites by mass spectrometry without isotopic labeling or spiked standards. Anal Chem. 2003;75:4818-26.

32. Ong SE, 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;1:376-86.

33. Percy AJ, Byrns S, Pennington SR, Holmes DT, Anderson NL, Agreste TM, et al. Clinical translation of MS-based, quantitative plasma proteomics: status, challenges, requirements, and potential. Expert review of proteomics. 2016;13:673-84.

34. Thompson A, Schäfer J, Kuhn K, Kienle S, Schwarz J, Schmidt G, et al. Tandem Mass Tags: A Novel Quantification Strategy for Comparative Analysis of Complex Protein Mixtures by MS/MS. Anal Chem. 2003;75:1895-904.

35. Hung C-W, Tholey A. Tandem Mass Tag Protein Labeling for Top-Down Identification and Quantification. Anal Chem. 2012;84:161-70.

36. McAlister GC, Huttlin EL, Haas W, Ting L, Jedrychowski MP, Rogers JC, et al. Increasing the multiplexing capacity of TMTs using reporter ion isotopologues with isobaric masses. Anal Chem. 2012;84:7469-78.

37. Thompson A, Wölmer N, Koncarevic S, Selzer S, Böhm G, Legner H, et al. TMTpro: Design, Synthesis, and Initial Evaluation of a Proline-Based Isobaric 16-Plex Tandem Mass Tag Reagent Set. Anal Chem. 2019;91:15941-50.

38. McAlister GC, Huttlin EL, Haas W, Ting L, Jedrychowski MP, Rogers JC, et al. Increasing the multiplexing capacity of TMTs using reporter ion isotopologues with isobaric masses. Anal Chem. 2012;84:7469-78.

39. Yang F, Shen Y, Camp DG, 2nd, Smith RD. High-pH reversed-phase chromatography with fraction concatenation for 2D proteomic analysis. Expert review of proteomics. 2012;9:129-34.

40. Karp NA, Huber W, Sadowski PG, Charles PD, Hester SV, Lilley KS. Addressing accuracy and precision issues in iTRAQ quantitation. Mol Cell Proteomics. 2010;9:1885-97.

41. Bai B, Tan H, Pagala VR, High AA, Ichhaporia VP, Hendershot L, et al. Deep Profiling of Proteome and Phosphoproteome by Isobaric Labeling, Extensive Liquid Chromatography, and Mass Spectrometry. Methods Enzymol. 2017;585:377-95.

42. Ting L, Rad R, Gygi SP, Haas W. MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics. Nat Methods. 2011;8:937-40.

43. Schmidt C, Lenz C, Grote M, Lührmann R, Urlaub H. Determination of protein stoichiometry within protein complexes using absolute quantification and multiple reaction monitoring. Anal Chem. 2010;82:2784-96.

44. Kirsch S, Widart J, Louette J, Focant JF, De Pauw E. Development of an absolute quantification method targeting growth hormone biomarkers using liquid chromatography coupled to isotope dilution mass spectrometry. J Chromatogr A. 2007;1153:300-6.

45. Politi L, Groppi A, Polettini A. Applications of liquid chromatography-mass spectrometry in doping control. Journal of analytical toxicology. 2005;29:1-14.

46. Gerber SA, Rush J, Stemman O, Kirschner MW, Gygi SP. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. Proc Natl Acad Sci U S A. 2003;100:6940-5.

47. Kirkpatrick DS, Gerber SA, Gygi SP. The absolute quantification strategy: a general procedure for the quantification of proteins and post-translational modifications. Methods. 2005;35:265-73.

48. Bowers GN, Jr., Fassett JD, White Et. Isotope dilution mass spectrometry and the National Reference System. Anal Chem. 1993;65:475r-9r.

49. Worboys JD, Sinclair J, Yuan Y, Jørgensen C. Systematic evaluation of quantotypic peptides for targeted analysis of the human kinome. Nat Meth. 2014;11:1041-4.

50. Brun V, Dupuis A, Adrait A, Marcellin M, Thomas D, Court M, et al. Isotope-labeled protein standards: toward absolute quantitative proteomics. Mol Cell Proteomics. 2007;6:2139-49.

51. Li Y, Shu Y, Peng C, Zhu L, Guo G, Li N. Absolute quantitation of isoforms of post-translationally modified proteins in transgenic organism. Molecular & cellular proteomics : MCP. 2012;11:272-85.

52. Liu H, Sadygov RG, Yates JR, 3rd. A model for random sampling and estimation of relative protein abundance in shotgun proteomics. Anal Chem. 2004;76:4193-201.

53. Higgs RE, Knierman MD, Gelfanova V, Butler JP, Hale JE. Comprehensive label-free method for the relative quantification of proteins from biological samples. J Proteome Res. 2005;4:1442-50.

54. Lu P, Vogel C, Wang R, Yao X, Marcotte EM. Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. Nat Biotechnol. 2007;25:117-24.

55. Silva JC, Gorenstein MV, Li GZ, Vissers JP, Geromanos SJ. Absolute quantification of proteins by LCMSE: a virtue of parallel MS acquisition. Mol Cell Proteomics. 2006;5:144-56.

56. Kadumuri RV, Janga SC. Epitranscriptomic Code and Its Alterations in Human Disease. Trends Mol Med. 2018;24:886-903.

57. Roundtree IA, Evans ME, Pan T, He C. Dynamic RNA modifications in gene expression regulation. Cell. 2017;169:1187-200.

58. Desrosiers R, Friderici K, Rottman F. Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. Proc Natl Acad Sci U S A. 1974;71:3971-5.

59. Desrosiers RC, Friderici KH, Rottman FM. Characterization of Novikoff hepatoma mRNA methylation and heterogeneity in the methylated 5' terminus. Biochemistry. 1975;14:4367-74.

60. Furuichi Y, Morgan M, Shatkin AJ, Jelinek W, Salditt-Georgieff M, Darnell JE. Methylated, blocked 5 termini in HeLa cell mRNA. Proc Natl Acad Sci U S A. 1975;72:1904-8.

61. 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 (New York, NY). 1997;3:1233-47.

62. 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;485:201-6.

63. Meyer KD, Saletore Y, Zumbo P, Elemento O, Mason CE, Jaffrey SR. Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. Cell. 2012;149:1635-46.

64. Linder B, Grozhik AV, Olarerin-George AO, Meydan C, Mason CE, Jaffrey SR. Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. Nat Meth. 2015;12:767-72.

65. Liu J, Yue Y, Han D, Wang X, Fu Y, Zhang L, et al. A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. Nat Chem Biol. 2014;10:93-5.

66. Ping XL, Sun BF, Wang L, Xiao W, Yang X, Wang WJ, et al. Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. Cell Res. 2014;24:177-89.

67. Wen J, Lv R, Ma H, Shen H, He C, Wang J, et al. Zc3h13 Regulates Nuclear RNA m(6)A Methylation and Mouse Embryonic Stem Cell Self-Renewal. Mol Cell. 2018;69:1028-38.e6.

68. Patil DP, Chen CK, Pickering BF, Chow A, Jackson C, Guttman M, et al. m(6)A RNA methylation promotes XIST-mediated transcriptional repression. Nature. 2016;537:369-73.

69. Yue Y, Liu J, Cui X, Cao J, Luo G, Zhang Z, et al. VIRMA mediates preferential m(6)A mRNA methylation in 3'UTR and near stop codon and associates with alternative polyadenylation. Cell discovery. 2018;4:10.

70. Warda AS, Kretschmer J, Hackert P, Lenz C, Urlaub H, Hobartner C, et al. Human METTL16 is a N(6)-methyladenosine (m(6)A) methyltransferase that targets pre-mRNAs and various non-coding RNAs. EMBO Rep. 2017;18:2004-14.

71. van Tran N, Ernst FG M, Hawley BR, Zorbas C, Ulryck N, Hackert P, et al. The human 18S rRNA m6A methyltransferase METTL5 is stabilized by TRMT112. Nucleic Acids Res. 2019;47:7719-33.

72. Ma H, Wang X, Cai J, Dai Q, Natchiar SK, Lv R, et al. N(6-)Methyladenosine methyltransferase ZCCHC4 mediates ribosomal RNA methylation. Nat Chem Biol. 2019;15:88-94.

73. 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;7:885-7.

74. Zheng GQ, Dahl JA, Niu YM, Fedorcsak P, Huang CM, Li CJ, et al. ALKBH5 Is a Mammalian RNA Demethylase that Impacts RNA Metabolism and Mouse Fertility. Mol Cell. 2013;49:18-29.

75. Mauer J, Luo X, Blanjoie A, Jiao X, Grozhik AV, Patil DP, et al. Reversible methylation of m^6A_m in the 5' cap controls mRNA stability. Nature. 2017;541:371-5.

76. Zheng G, Dahl JA, Niu Y, Fedorcsak P, Huang CM, Li CJ, et al. ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. Mol Cell. 2013;49:18-29.

77. Wang X, Zhao BS, Roundtree IA, Lu Z, Han D, Ma H, et al. N(6)-methyladenosine Modulates Messenger RNA Translation Efficiency. Cell. 2015;161:1388-99.

78. Wang X, Lu Z, Gomez A, Hon GC, Yue Y, Han D, et al. N6-methyladenosine-dependent regulation of messenger RNA stability. Nature. 2014;505:117-20.

79. Shi H, Wang X, Lu Z, Zhao BS, Ma H, Hsu PJ, et al. YTHDF3 facilitates translation and decay of N(6)-methyladenosine-modified RNA. Cell Res. 2017;27:315-28.

80. Zaccara S, Jaffrey SR. A Unified Model for the Function of YTHDF Proteins in Regulating m(6)A-Modified mRNA. Cell. 2020;181:1582-95.e18.

81. Xiao W, Adhikari S, Dahal U, Chen YS, Hao YJ, Sun BF, et al. Nuclear m(6)A Reader YTHDC1 Regulates mRNA Splicing. Mol Cell. 2016;61:507-19.

82. Roundtree IA, Luo GZ, Zhang Z, Wang X, Zhou T, Cui Y, et al. YTHDC1 mediates nuclear export of N(6)-methyladenosine methylated mRNAs. Elife. 2017;6.

83. Hsu PJ, Zhu Y, Ma H, Guo Y, Shi X, Liu Y, et al. Ythdc2 is an N(6)-methyladenosine binding protein that regulates mammalian spermatogenesis. Cell Res. 2017;27:1115-27.

84. Alarcon CR, Lee H, Goodarzi H, Halberg N, Tavazoie SF. N6-methyladenosine marks primary microRNAs for processing. Nature. 2015;519:482-5.

85. Liu N, Dai Q, Zheng G, He C, Parisien M, Pan T. N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. Nature. 2015;518:560-4.

86. Huang H, Weng H, Sun W, Qin X, Shi H, Wu H, et al. Recognition of RNA N(6)-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. Nat Cell Biol. 2018;20:285-95.

87. Meyer KD, Patil DP, Zhou J, Zinoviev A, Skabkin MA, Elemento O, et al. 5' UTR m(6)A Promotes Cap-Independent Translation. Cell. 2015;163:999-1010.

88. Choe J, Lin S, Zhang W, Liu Q, Wang L, Ramirez-Moya J, et al. mRNA circularization by METTL3–eIF3h enhances translation and promotes oncogenesis. Nature. 2018;561:556-60.

89. Yue B, Song C, Yang L, Cui R, Cheng X, Zhang Z, et al. METTL3-mediated N6-methyladenosine modification is critical for epithelial-mesenchymal transition and metastasis of gastric cancer. Molecular cancer. 2019;18:142.

90. Shen C, Sheng Y, Zhu AC, Robinson S, Jiang X, Dong L, et al. RNA Demethylase ALKBH5 Selectively Promotes Tumorigenesis and Cancer Stem Cell Self-Renewal in Acute Myeloid Leukemia. Cell stem cell. 2020;27:64-80.e9.

91. He L, Li H, Wu A, Peng Y, Shu G, Yin G. Functions of N6-methyladenosine and its role in cancer. Molecular cancer. 2019;18:176.

92. Boccaletto P, Machnicka MA, Purta E, Piatkowski P, Baginski B, Wirecki TK, et al. MODOMICS: a database of RNA modification pathways. 2017 update. Nucleic Acids Res. 2018;46:D303-d7.

93. Zhang W, Foo M, Eren AM, Pan T. tRNA modification dynamics from individual organisms to metaepitranscriptomics of microbiomes. Mol Cell. 2022.

94. Helm M, Giegé R, Florentz C. A Watson−Crick Base-Pair-Disrupting Methyl Group (m1A9) Is Sufficient for Cloverleaf Folding of Human Mitochondrial tRNALys Biochemistry. 1999;38:13338-46.

95. Sharma S, Watzinger P, Kötter P, Entian KD. Identification of a novel methyltransferase, Bmt2, responsible for the N-1-methyl-adenosine base modification of 25S rRNA in Saccharomyces cerevisiae. Nucleic Acids Res. 2013;41:5428-43.

96. Dominissini D, Nachtergaele S, Moshitch-Moshkovitz S, Peer E, Kol N, Ben-Haim MS, et al. The dynamic N1-methyladenosine methylome in eukaryotic messenger RNA. Nature. 2016;530:441-6.

97. Yang X, Yang Y, Sun BF, Chen YS, Xu JW, Lai WY, et al. 5-methylcytosine promotes mRNA export - NSUN2 as the methyltransferase and ALYREF as an m(5)C reader. Cell Res. 2017;27:606-25.

98. Yang Y, Wang L, Han X, Yang WL, Zhang M, Ma HL, et al. RNA 5-Methylcytosine Facilitates the Maternal-to-Zygotic Transition by Preventing Maternal mRNA Decay. Mol Cell. 2019;75:1188-202.e11.

99. Tuorto F, Liebers R, Musch T, Schaefer M, Hofmann S, Kellner S, et al. RNA cytosine methylation by Dnmt2 and NSun2 promotes tRNA stability and protein synthesis. Nat Struct Mol Biol. 2012;19:900-5.

100. Zhao BS, Roundtree IA, He C. Post-transcriptional gene regulation by mRNA modifications. Nat Rev Mol Cell Biol. 2017;18:31-42.

101. Han L, Kon Y, Phizicky EM. Functional importance of Ψ38 and Ψ39 in distinct tRNAs, amplified for tRNAGln(UUG) by unexpected temperature sensitivity of the s2U modification in yeast. RNA (New York, NY). 2015;21:188-201.

102. Jack K, Bellodi C, Landry DM, Niederer RO, Meskauskas A, Musalgaonkar S, et al. rRNA pseudouridylation defects affect ribosomal ligand binding and translational fidelity from yeast to human cells. Mol Cell. 2011;44:660-6.

103. Fernández IS, Ng CL, Kelley AC, Wu G, Yu Y-T, Ramakrishnan V. Unusual base pairing during the decoding of a stop codon by the ribosome. Nature. 2013;500:107-10.

104. Barbieri I, Kouzarides T. Role of RNA modifications in cancer. Nat Rev Cancer. 2020;20:303-22.

Figure 1.1 Overview of MS-based proteomics in the quantification level, study aim, provided information, and underlying methodology perspective.

The figure is adapted from Rozanova, S. *et al.*, Quantitative Methods in Proteomics. Springer US; 2021. p. 85-116.

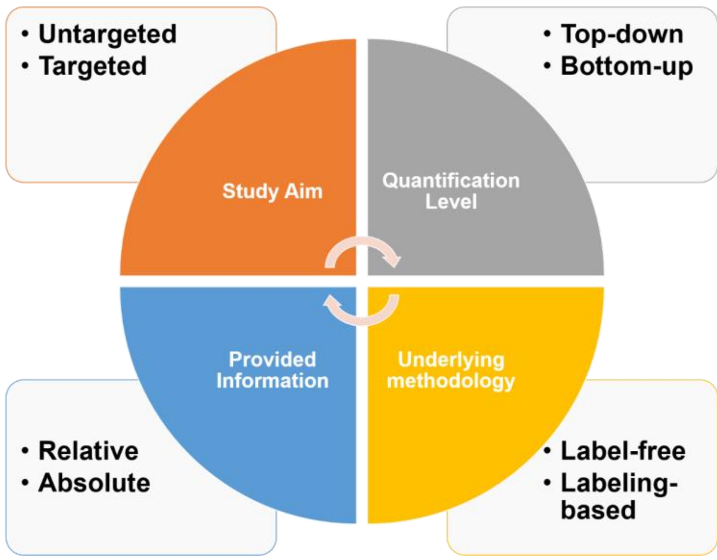


Figure 1.2 (a) Mass spectrometry instrumentation in MRM, PRM, DDA, and DIA. (b) Summary of MS2 sampling of the precursor space in the retention time dimension and *m/z* dimension.

Panel (b) is partially adapted from Egertson, J.D. *et al.*, Multiplexed peptide analysis using data-independent acquisition and Skyline. Nature Protocols. 2015; 10: 887-903.

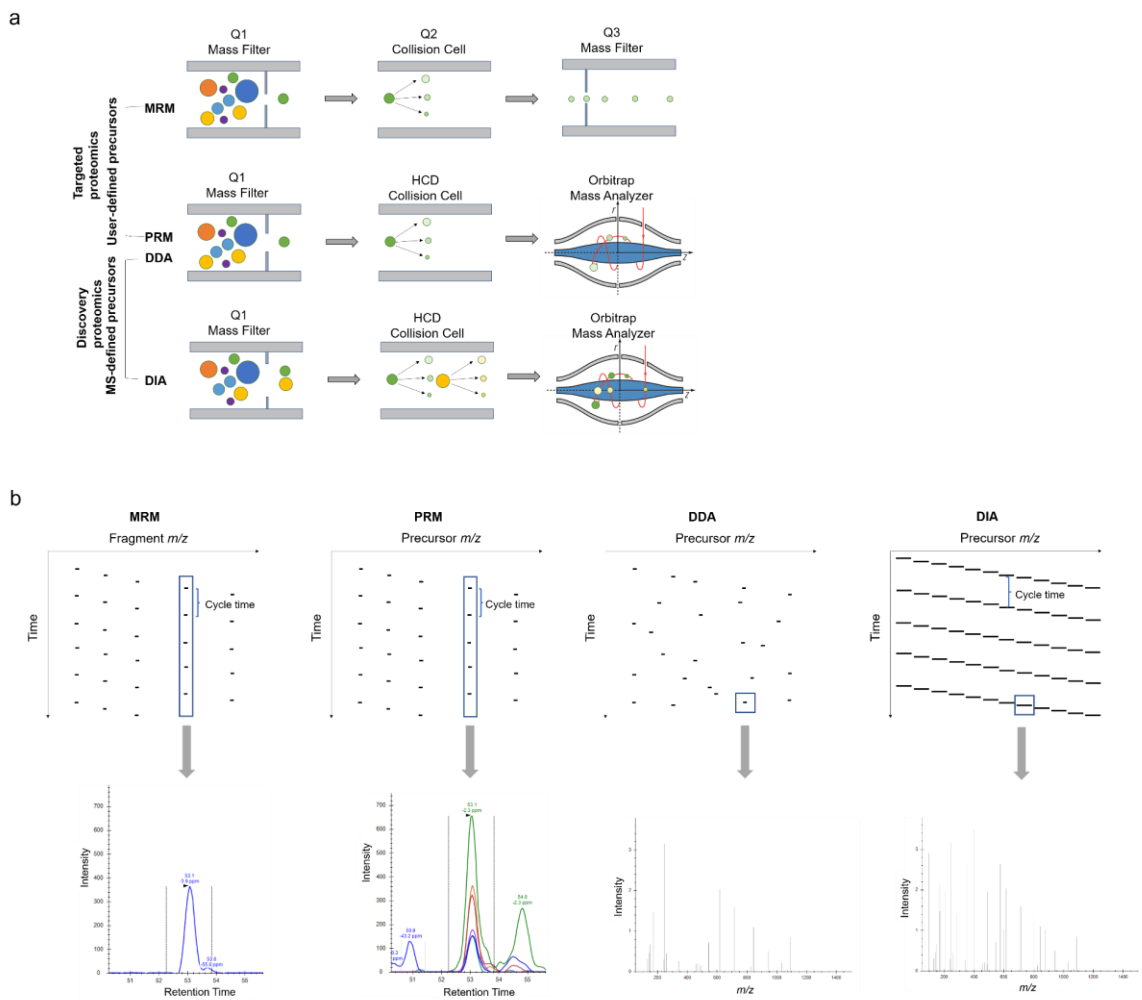


Figure 1.3 Workflow of SRM library development.

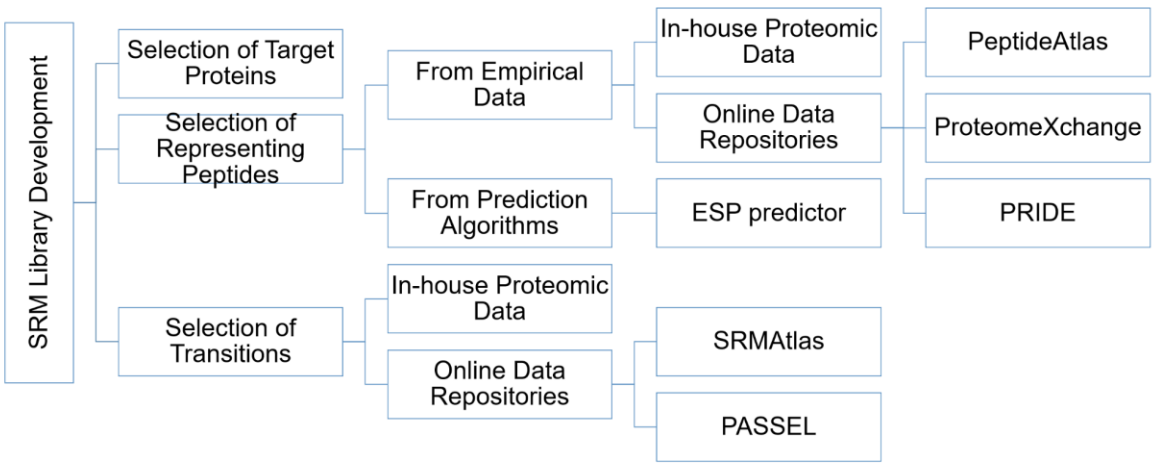


Figure 1.4 Criteria for SRM or PRM peak identification.

(a) Co-elution of all fragment ions from one precursor ion. y_{10} ion is a potential interfering ion that needs to be excluded for quantification (b) Co-elution of light- and heavy- isotope labeled precursor ions. (c) dotp value > 0.7. It gauges the similarities in fragment patterns in the acquired MS/MS (left) and reference MS/MS (right) in the library.

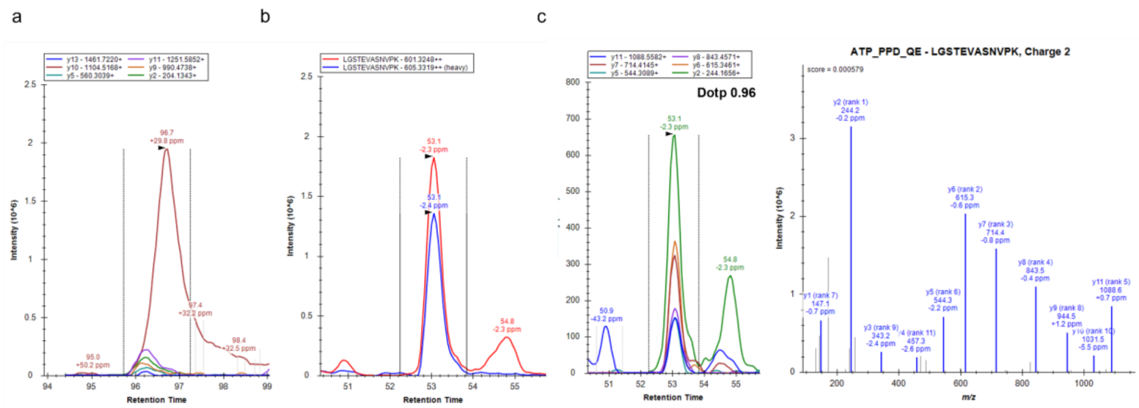


Figure 1.5 Workflow of quantitative proteomics in label-free quantification (a), SILAC (b), SIL standards (c), and TMT labeling (d).

Orange and red triangles indicate non-labeling, whereas light and dark blue triangles indicate heavy stable isotope-labeling. TMTduplex is shown for simplicity. The figure is partially adapted from Bantscheff, M. *et al.*, Quantitative mass spectrometry in proteomics: a critical review. Anal. Bioanal. Chem. 2007; 389: 1017-31.

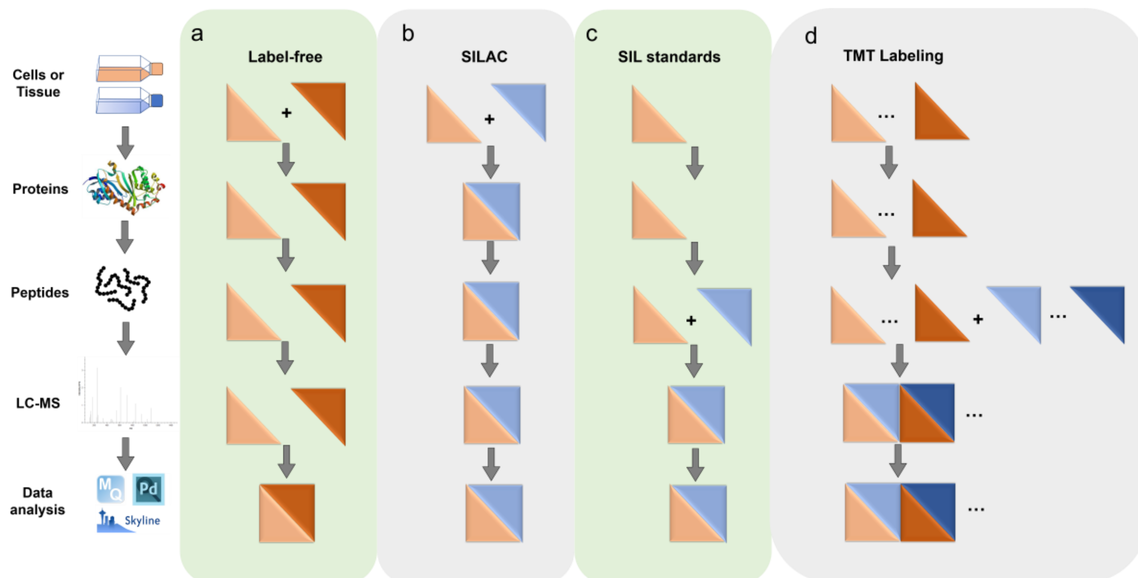


Figure 1.6 TMT-labeling in quantitative proteomics.

(a) Structure of the TMT reagent (b) Regents of TMT6plex. Mass reporter group, mass normalizer, and NH_2-reaction group are highlighted individually in yellow, green, and blue boxes. *m/z* value of the mass reporter group of each regent is labelled in the bottom-left corner. ^{13}C and ^{15}N are labeled in red and blue asterisks, respectively. (c) Workflow of TMT-labeling.

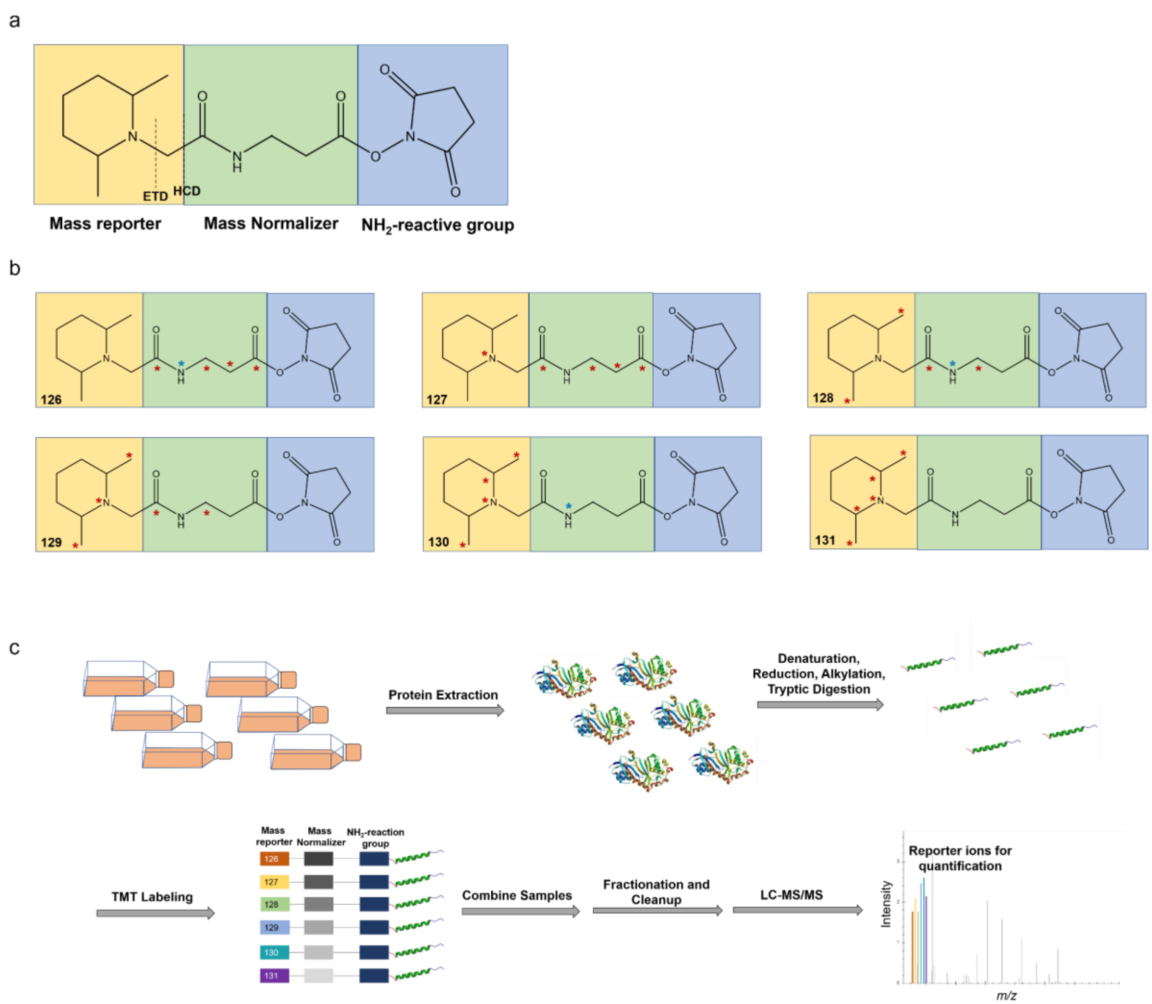


Figure 1.7 Overview of an m^{6}A writer complex, m^{6}A erasers, and major m^{6}A reader proteins.

m^{6}A reader proteins assume essential roles in mRNA splicing, export, degradation, and translation. Three current models on how m^{6}A affects translation are shown.

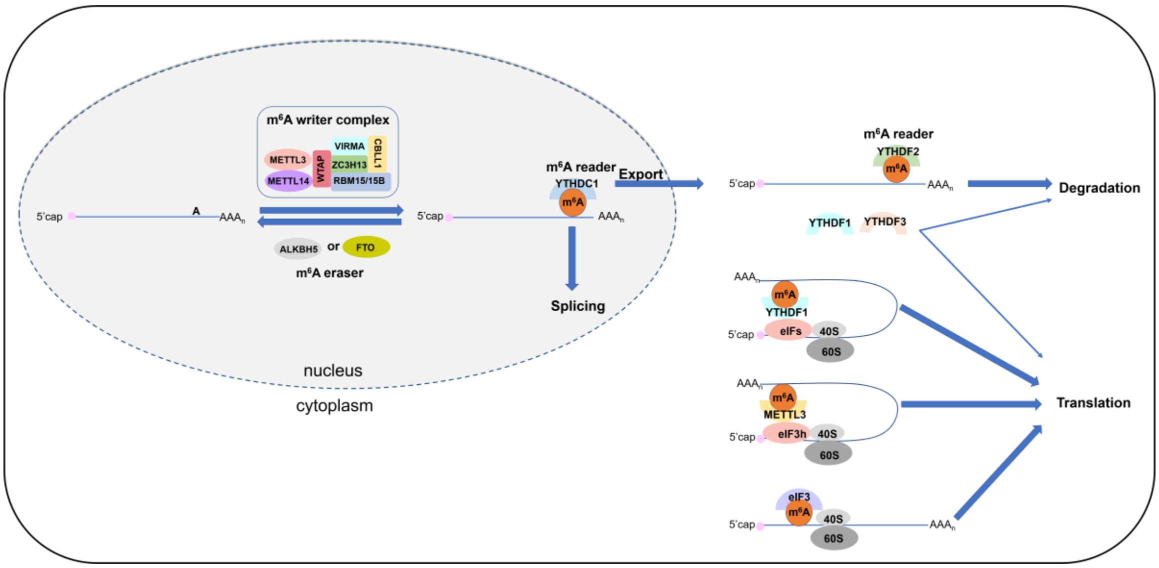
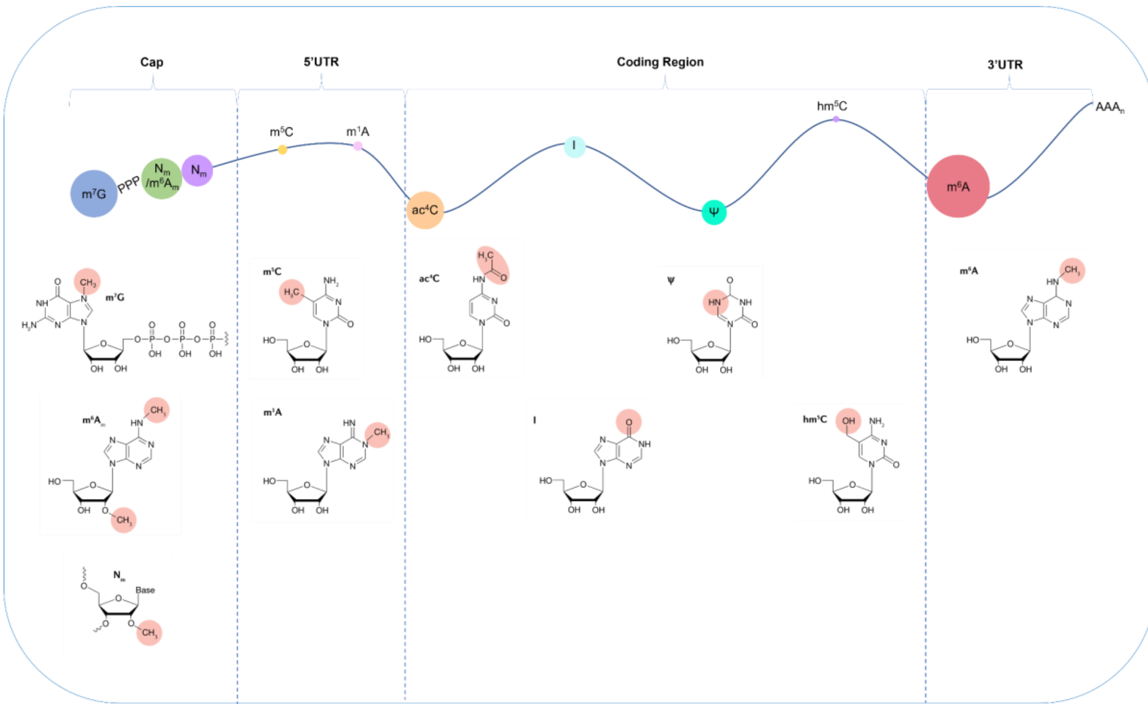


Figure 1.8 Overview of nucleoside modifications distribution in mRNA.

The dimension of the circles in mRNA schematically represents the abundance of the modification. For modifications located in different regions of mRNA, only the major location is shown for simplicity. The figure is adapted from Zaccara, S. *et al.*, Reading, writing and erasing mRNA methylation. Nat. Rev. Mol. Cell Biol. 2019; 20: 608-24.



2. Chapter 2. Discovery of TBC1D7 as a potential driver for melanoma cell invasion

2.1 Introduction

Melanoma is the least common type of skin cancer, but it is the deadliest because it is more likely to metastasize, where the five-year survival rate for distant-stage melanoma is as low as 15-20% (1). In the local invasion stage of the metastasis cascade, cells must escape from the primary site by acquiring a more motile phenotype. These cells then secrete matrix metalloproteinases (MMPs) to degrade the basement membrane of the surrounding extracellular matrix, thereby allowing cancer cells to migrate and invade adjacent tissues. The cells then enter the lymphatic system or circulation to reach distant organs to produce secondary tumors (2).

Rab subfamily of small GTPases are mainly responsible for membrane trafficking, e.g. vesicle budding, transport, and fusion (3). Like other small GTPases, Rab GTPases exist in GDP-bound inactive state or GTP-bound active state. Due to tight binding of Rab proteins with GDP, guanine nucleotide-exchange factors (GEFs) activate Rab proteins by catalyzing the exchange of the GTPase-bound GDP with GTP, which in turn facilitates the binding of downstream effector proteins (4). On the other hand, owing to the slow inherent rate of GTP hydrolysis mediated by GTPases, GTPase-activating proteins (GAPs) inactivate Rab proteins by catalyzing GTP hydrolysis (5).

Most identified eukaryotic Rab GAPs contain a Tre2–Bub2–Cdc16 (TBC) domain (6, 7). The TBC domain harbors approximately 200 amino acids and it was first discovered as a conserved domain that is shared among the Tre-2 oncogene product and the yeast cell cycle regulators Bub2 and Cdc16 (6). The TBC domain-containing proteins (TBC proteins) have 44 predicted members based on the sequence homology of the TBC domain (8). It was demonstrated that TBC proteins are essential in regulating intracellular trafficking, particularly in integrating signal between RABs or between RABs and other small GTPases (8), thereby modulating many cellular processes, e.g. autophagy (9), primary cilium formation (10), and exosome secretion (11).

A few TBC proteins have been reported with roles in tumorigenesis or tumor progression. TBC1D3 was identified as an oncoprotein in prostate cancer (12), TBC1D8 was reported to play a role in the tumorigenesis of ovarian cancer (13), and TBC1D16 is known to enhance melanoma progression by targeting epidermal growth factor receptor (EGFR) (14). However, there is no proteome-wide studies to identify the roles of TBC proteins in melanoma metastasis. The goal of the present study is to employ a quantitative proteomic method to identify the TBC proteins that may drive or suppress melanoma metastasis.

2.2 Material and method

2.2.1 Cell culture

WM-115 and WM-266-4 cells (WM pair) were obtained from the American Type Culture Collection (ATCC) (15), and this pair of cell lines are no longer available from ATCC. WM-115 cells were established from the vertical growth phase from the primary melanoma site (right anterior leg) of a 55-year-old female, whereas WM-266-4 cells were derived from the skin cutaneous metastasis to the right thigh of the same patient (15).

For the stable isotope labeling by amino acids in cell culture (SILAC) experiments (16), light or heavy lysine ([$^{13}C_6$, $^{15}N_2$]-L-lysine) and arginine ([$^{13}C_6$]-L-arginine), respectively, along with dialyzed FBS (Invitrogen), were added to the lysine, arginine-depleted SILAC medium to yield the light and heavy DMEM media. WM-115 and WM-266-4 cells were cultured in the heavy DMEM media for at least 6 cell doublings to ensure complete heavy isotope labelling. In the forward SILAC experiment, the lysate of light-labeled WM-266-4 cells and that of the heavy-labeled WM-115 cells were combined at 1:1 ratio (w/w), whereas the heavy-labeled WM-266-4 cell lysate was mixed equally with the light-labeled WM-115 cell lysate in the reverse SILAC experiment. Following the filter-aided sample preparation (FASP) protocol (17), proteins in the cell lysates were denatured, reduced, alkylated and digested with trypsin, as described previously (18).

2.2.2 LC-MS/MS

The tryptic digestion mixture was subjected to an off-line strong cation exchange (SCX) separation by using a PolySulfoethyl A SCX column (9.4 × 200 mm, 5 μm, 200 Å, PolyLC, Columbia, MD). Twenty-one fractions were collected over an elution with a 90-min linear gradient of 0 – 500 mM ammonium acetate in 0.1% formic acid. Each fraction was desalted using OMIX C18 pipette tips (Agilent). The peptide fractions were analyzed individually using reversed-phase liquid chromatography coupled with an LTQ-Orbitrap Velos mass spectrometer, and the detailed experimental conditions for LC-MS/MS were described elsewhere (18). LC-MS/MS data were searched against the International Protein Index (IPI) human database version 3.68, which contained 87,083 entries, using MaxQuant for protein identification and quantification (19).

2.2.3 Western Blot

Melanoma cells were cultured in DMEM (Invitrogen-Gibco) supplemented with 10% FBS (Invitrogen-Gibco) and 1% penicillin/streptomycin. The cells were trypsinized, rinsed twice with PBS, and the resultant cell pellet stored at - 80°C until analysis. The cells were lysed in CelLytic M (Sigma-Aldrich) supplemented with a protease inhibitor cocktail. After centrifugation, total proteins in the supernatant were quantified using the Quick Start Bradford Protein Assay (Bio-Rad). Approximately 10-20 µg total protein was loaded and separated on a 10% SDS-PAGE gel, transferred to a nitrocellulose membrane, and blocked with 5% milk in PBS-T. The membranes were subsequently incubated with primary antibodies targeting human TBC1D7 (D8K1Y, Cell Signaling Technology, 1:10,000), β-actin (Cell Signaling Technology, 1:10,000), or MITF (Abcam, Ab12039, 1:10,000). Membranes were then incubated with donkey anti-rabbit secondary antibody (Thermo Fisher Scientific, 1:10,000), or anti-mouse secondary antibody (Santa Cruz, m-IgGκ BP-HRP, 1:10,000). Amersham ECL™ Western Blot Detecting Reagent (GE Healthcare) was used to visualize the protein bands, following the vendor’s instructions.

2.2.4 Migration and invasion assay

FLAG-TBC1D7 plasmid was generously provided by Dr. Alexandre Reymond (20). WM-115 cells were transfected with a FLAG-TBC1D7 plasmid or empty vector using TransIT 2020 transfection reagent (Mirus) and incubated for 24 h. WM-266-4 cells were transfected with siTBC1D7 or non-targeting siRNA control (NTsiCtrl) using Lipofectamine RNAiMAX reagent (Thermo Fisher Scientific) and incubated for 72 h. The sequence of siTBC1D7 was 5′-GAACAAGUGCAGAGAAGAUA-3′ (21). At 24 h or 72 h post-transfection, WM-115 (2.5×10^4) and WM-266-4 (5×10^4) cells suspended in serum-free DMEM medium were seeded into the upper chamber of a transwell insert (Corning). DMEM medium supplemented with 10% FBS and 1% penicillin/streptomycin was added to the lower chamber as chemoattractant. For the migration assay, the cells, following incubation at 37°C for 24 h (for WM-115 and WM-266-4 cells), that migrated through the insert were fixed with 70% ethanol followed by staining with 0.5% crystal violet. The cells were counted using an inverted microscope by randomly selecting four fields of each insert and the numbers of cells from these four fields were subsequently averaged. Cell numbers were compared between treatment groups (i.e. with ectopic expression or siRNA knockdown of TBC1D7) and controls (i.e. empty vector or non-targeting control siRNA (NTsiCtrl)).

The invasion assay was conducted under the same conditions as the migration assay except that the transwell insert was coated with a matrigel basement membrane matrix (Corning). For this purpose, matrigel (200 – 400 μg/mL) in serum-free medium was coated on the top of the membrane of the transwell inserts at 37°C for 1 – 1.5 h. The matrigel was removed from the top surface of the membrane before seeding cells onto the upper chamber of the transwell inserts.

2.2.5 Gelatin zymography assay

After transfection of WM-115 cells with FLAG-TBC1D7 plasmid for 24 h, or transfection of WM-266-4 cells with siTBC1D7 for 72 h, the culture medium was replaced with FBS-free medium and the cells were cultured for another 24 h. Conditioned medium was collected and concentrated by around 40-fold using Microcon centrifugal filter units with a molecular weight cutoff of 30 kDa (EMD Millipore). Total protein concentration in the conditioned media was quantified using the Quick Start Bradford Protein Assay (Bio-Rad). One to two μg of total protein was separated using a 7.5% polyacrylamide gel containing 1 mg/mL gelatin under non-reducing conditions. After electrophoresis, the gel was washed with the zymography washing buffer (2.5% Triton X-100, 50 mM Tris-HCl, pH 7.5) at room temperature for 1 h, followed by incubation in a 37°C shaker for 10 min in a buffer containing 1.0% Triton X-100 and 50 mM Tris-HCl (pH 7.5). The gel was incubated in the fresh buffer for 20 h to induce digestion of the embedded gelatin by the renatured MMPs in the sample. The gel was stained with 0.5% Coomassie blue G-250 followed by destaining until white bands could be visualized against the blue background.

For gelatin zymography for the cell lysate, after transfection of FLAG-TBC1D7 plasmid in WM-115 cells for 24 h, or transfection of siTBC1D7 in WM-266-4 cells for 72 h, cells were harvested. Cells were lysed using Cellytic M (Sigma-Aldrich) supplemented with a protease inhibitor cocktail. 10 μg of total protein was separated, under non-reducing conditions, using a 7.5% polyacrylamide gel including 1 mg/mL gelatin. The subsequent protocol was the same as we described above for gelatin zymography for conditioned media.

2.2.6 MTT proliferation assay

After transfection of WM-115 cells with TBC1D7 overexpression plasmid for 24 h, and that of WM-266-4 cells with siTBC1D7 for 48h, cells were seeded to a 96 well plate with 2000 cells in each well. After cell attachment, 90 μL of DMEM medium without FBS and 10 μL of 3-(4, 5-Dimethylthiazol-2-yl)-2, 5-diphenyltetrazolium bromide (MTT) solution

(5 mg/mL) were added. After incubation at 37°C for 4 h, the insoluble purple formazan was formed. 100 µL of lysis buffer (50% DMF, 20% SDS, 2.5% acetic acid, 20 mM HCl, pH 4.7) was added. After overnight incubation at 37°C, absorbance of the solubilized formazan product was measured at 570 nm using the Perkin Elmer Wallac 1420 Victor2 Microplate Reader.

2.2.7 Chromatin immunoprecipitation and Real-time Quantitative PCR (RT-qPCR)

WM-115 and WM-266-4 cells (1×10^7 cells) were treated with 1% formaldehyde at room temperature for 10 min to induce protein-DNA cross-linking. The cells were lysed and the cellular DNA was sonicated to produce 300-500 bp fragments. Chromatin was subsequently immunoprecipitated using anti-MITF (ab12039, Abcam) or IgG (2729S, Cell Signaling Technology). RT-qPCR was carried out after the purification of the precipitated DNA. The primers for *TBC1D16* gene were 5′-GGCCACATACAAAGGGATCG-3′ (forward) and 5′-CTCGCGGAGGCAATCTGA-3′ (reverse), and the primers for *TBC1D7* gene were 5′-TCCTAGAGGACGCCTTTGTC-3′ and 5′-ACAGCTGCATGACGATTTGG-3′.

2.2.8 Bioinformatic Analysis

The mRNA expression data of *TBC1D7* and survival information for 479 skin cutaneous melanoma (SKCM) patients were retrieved from the TCGA dataset using the cBioPortal for Cancer Genomics (http://www.cbioportal.org/) (22). The mRNA expression data of *TBC1D7* and survival information for a cohort of 150 melanoma patients were retrieved from GSE65904, and MedCalc (MedCalc Software, Ostend, Belgium; https://www.medcalc.org/) was used for the generation of Kaplan-Meier survival curve.

The mRNA expression of *TBC1D7* in SKCM and uveal melanoma (UVM) patients from the TCGA dataset was compared with six other representative cancer types. We also interrogated The Cancer Cell Line Encyclopedia (CCLE) database and NCI-60 cell line database, through which box-and-whisker plots were generated for TBC1D7 mRNA expression in melanoma cell lines versus cell lines derived from other types of cancer.

2.3 Results and Discussion

2.3.1. Up-regulation of TBC1D7 in WM-266-4 compared to WM-115 cells

To identify potential drivers or suppressors of melanoma metastasis, we utilized an unbiased quantitative proteomic approach to examine, at the entire proteome scale, the differential protein expression in paired WM-115 and WM-266-4 cells, which are melanoma cells derived from the primary and metastatic melanoma sites of the same patient (15). In this respect, the use of paired melanoma cell lines derived from the same patient allows for the minimization of inter-patient heterogeneity. In this vein, while skin and subcutaneous tissue are the most common sites for regional melanoma metastasis, melanoma is also known to metastasize to other organs including lung and liver (23); hence, drivers or suppressors for skin metastasis may vary from those for lung or liver metastasis owing to differences in tumor microenvironment.

To reduce sample complexity, we conducted an off-line SCX fractionation prior to LC-MS/MS analysis. By employing SILAC together with off-line 2D LC-MS/MS analysis, we were able to quantify a total of 7387 proteins (Shown in Figure 2.1A is a flowchart of the SILAC experiments. Data not shown). Among these proteins, 5955 (81%) were quantified in at least two replicates of SILAC-labeling experiments, and 1551 (21%) displayed at least

1.5-fold differences between these two cell lines (Data not shown). For those displayed at least 1.5-fold changes, we carried out pathway analyses using the Database for Annotation, Visualization and Integrated Discovery (DAVID) (24). Many of them have known functions in the pathway of signal transduction, oxidation-reduction process, and positive regulation of GTPases activity (Figure 2.2). These quantitative proteomic data provided an important basis for identifying putative proteins that play roles in metastatic transformation of melanoma.

Our laboratory recently examined the roles of aberrant expression of small GTPase proteins in melanoma and colorectal cancer metastasis (25-27). We are also interested in how regulatory proteins of small GTPases, i.e., GAPs and GEFs, modulate melanoma metastasis. In addition, many TBC domain-containing proteins are GAP proteins for small GTPases, and one of the them, TBC1D16, is a driver for melanoma metastasis (14). Therefore, we placed the emphasis of the present study on understanding whether differential expression of other TBC proteins affects melanoma metastasis. The above proteomic data also led to the quantification of 24 TBC domain-containing proteins, which accounts for 55% of this protein family, in WM-115 and WM-266-4 cells (Figure 2.1B, Table 2.1).

Fold-change has been widely used as a criterion for selection of altered proteins in quantitative proteomics. On the ground that the average relative standard deviation for the quantification results of TBC proteins was 12% (Table 2.1), we chose 1.5-fold as a cutoff for identifying up- and down-regulated proteins. Among the quantified TBC proteins, TBC1D4, TBC1D16, TBC1D7, and TBC1D10A were up-regulated, whereas TBC1D24 was down-regulated, by more than 1.5-fold in WM-266-4 relative to WM-115 cells. TBC1D7 was up-regulated by 2.5-fold in WM-266-4 metastatic melanoma cells relative to WM-115 primary melanoma cells. Figure 2.3A shows the MS and MS/MS of a tryptic peptide derived from TBC1D7, detected in both forward and reverse SILAC labeling experiments. We also verified the augmented expression of TBC1D7 protein in WM-266-4 over WM-115 cells by Western blot analysis (Figure 2.3B).

We next assessed if differential expression of these TBC family proteins is associated with survival of melanoma patients. To this end, we performed Kaplan-Meier survival analyses based on the gene expression data of TBC family proteins for the skin cutaneous melanoma (SKCM) patients in the TCGA database. We found that patients with higher expression levels of *TBC1D16* or *TBC1D7* gene exhibit poorer overall survival rates (Figure 2.4A), which is in line with our proteomic results. However, the survival analyses of TBC1D4, TBC1D25, TBC1D10A, and TBC1D24 do not support our proteomic data of WM-115 and WM-266-4 cells derived from a single patient, which is likely due to patient heterogeneity, or discrepancy between mRNA and protein expression. Notably, TBC1D16 was reported to enhance melanoma progression (14). Hence, these results suggest that TBC1D7 may play an important role in melanoma metastasis.

We also carried out distant metastasis-free survival (DMFS) analysis on the basis of data for a cohort of 150 melanoma patients in the GSE65904 dataset. We found that higher levels of expression of TBC1D7 are significantly correlated with lower DMFS of melanoma patients (Figure 2.4B). This result again supports our notion that TBC1D7 may play a crucial role in melanoma metastasis.

Interrogation of the TCGA data also unveiled significantly higher levels of mRNA expression of *TBC1D7* gene in SKCM and UVM patients relative to other types of cancer (Figure 2.4C). Likewise, analyses of the publicly available gene expression data for the Cancer Cell Line Encyclopedia (CCLE) and NCI-60 human tumor cell lines showed that the *TBC1D7* mRNA expression levels in melanoma cell lines were up-regulated relative to other types of cancer (Figure 2.4D-E). Therefore, results from quantitative proteomic and bioinformatic analyses suggest that TBC1D7 could be a potential driver for melanoma metastasis.

2.3.2. TBC1D7 promotes invasion of melanoma cells *in vitro*

We next asked if elevated TBC1D7 expression promotes melanoma cell invasion *in vitro*. Using transwell assays, we observed that siRNA-mediated knockdown of TBC1D7 led to significantly decreased migration and invasion of WM-266-4 cells (Figure 2.5A, and Figure 2.6C shows the Western blot results for the validation of knockdown efficiency of TBC1D7 in WM-266-4 cells). Reciprocally, overexpression of TBC1D7 in WM-115 cells results in elevated invasion of these cells, though no increase in migratory capacity was observed (Figure 2.5B).

Matrix metalloproteinases 2 (MMP2) and 9 (MMP9) are type-IV collagenases secreted by cells (28). Type-IV collagen is the main component of basement membranes (29), whose degradation is crucial for the metastatic transformation of cancer (30). Therefore, the altered expression or activities of MMP2 and MMP9 may contribute to cancer metastasis. We next explored if MMP2 and MMP9 play a role in the invasive phenotype of melanoma cells modulated by TBC1D7. Gelatin zymography assay results revealed substantially diminished activities of MMP2 and MMP9 in the conditioned media after siRNA-mediated knockdown of TBC1D7 in WM-266-4 cells compared to treatment with non-targeting control siRNA (Figure 2.5C). However, the enzymatic activities of MMP2 and MMP9 were not modulated by ectopic overexpression of TBC1D7 in WM-115 primary melanoma cells (Figure 2.5D).

We also performed gel zymography assays using cell lysates, and our results showed that genetic depletion of TBC1D7 in WM-266-4 cells did not alter the enzymatic activity of MMP2 in the cell lysate (Figure 2.6A), and similar findings were made for WM-115 cells upon ectopic overexpression of TBC1D7 (Figure 2.6B). Additionally, no MMP9 activity was detectable in the lysate of WM-115 or WM-266-4 cells. Because knockdown of TBC1D7 did not alter the MMP2 activity in lysate of WM-266-4 cells, but diminished the MMP2 activity in the secreted proteome, our results suggest that TBC1D7 may promote the transport of MMP2 secretory vesicles.

We further preformed the MTT assay and found that genetic depletion of *TBC1D7* led to drastically diminished proliferation of WM-266-4 cells (Figure 2.7A). Along this line, it is worth noting that lung cancer cell growth was shown to be suppressed by siRNA-mediated knockdown of TBC1D7 (21). Overexpression of *TBC1D7* in WM-115 cells suppressed proliferation prior to day four, though no significant change was found at day six (Figure 2.7B).

2.3.3 TBC1D7 mRNA expression is regulated by MITF

We further explored the mechanisms through which TBC1D7 drives melanoma cell invasion. MITF is a transcription factor regulating many genes involved in melanocyte development (31). Tirosh *et al.* (32) reported 100 genes exhibiting the highest correlations with MITF, including *TBC1D16* and *TBC1D7*. Mining of the TCGA and CCLE databases revealed a strong positive correlation between the mRNA expression of *TBC1D7* or *TBC1D16* and that of *MITF* in patients of the TCGA-SKCM cohort and melanoma cell lines in the CCLE database (Figure 2.8A-D).

To further explore if TBC1D7 is directly regulated by MITF, we carried out chromatin immunoprecipitation (ChIP) followed by RT-qPCR assay to evaluate the interaction between MITF and the promoter regions of *TBC1D16* and *TBC1D7* genes. Indeed, we found elevated enrichment of MITF in the promoter regions of *TBC1D16* and *TBC1D7* genes in WM-266-4 cells than WM-115 cells, suggesting that TBC1D7 is directly regulated by MITF in WM-266-4 cells (Figure 2.8E).

To substantiate the above findings, we assessed how the expression level of *TBC1D7* is modulated by knocking down the expression of *MITF* gene in the metastatic melanoma WM-266-4 cell lines. We found significantly decreased expression level of TBC1D7 protein in WM-266-4 cells after RNAi-mediated knock-down of *MITF* (Figure 2.8F). Together, these results reveal TBC1D7 as a transcriptional target of MITF.

2.3.4 Potential mechanism of TBC1D7 in driving melanoma cell invasion

Several TBC proteins (e.g., TBC1D3, TBC1D8, and TBC1D16) were reported to be involved in tumorigenesis or tumor progression (12-14). In particular, TBC1D16 is known to be a driver of melanoma metastasis. By using an unbiased quantitative proteomic method, we aim to identify other TBC proteins that may drive or suppress melanoma metastasis. We discovered that TBC1D7 promotes melanoma cell invasion in the WM pair of cultured melanoma cells, at least in part, through modulating the activities of MMP2 and MMP9.

We also studied the potential mechanism of upstream regulation of TBC1D7. In this vein, TBC1D16 was shown to be a transcriptional target of MITF, and TBC1D16 was found to be down-regulated in WM-115 relative to WM-266-4 cells, which was found to be correlated with the methylation status of cytosine residues in its promoter region (14). Because both TBC1D16 and TBC1D7 were up-regulated in metastatic melanoma cells based on our proteomic results, we next assessed whether elevated expression of *TBC1D7* gene arises from hypomethylation in its promoter region. Interrogation of the TCGA database revealed that, similar as *TBC1D16*, elevated levels of expression of *TBC1D7* are associated with promoter hypomethylation in the SKCM dataset (Figure 2.9A-B). *TBC1D7* hypomethylation is also correlated with poorer melanoma patient survival in the TCGA-SKCM cohort (Figure 2.9C). These support a model where epigenetic activation of TBC1D7 promotes the metastatic transformation of a large fraction of melanoma patients.

2.4 Conclusion

In conclusion, our SILAC-based quantitative proteomic experiment led to the quantification of 7387 proteins, including 55% of known TBC proteins, in WM-115 and WM-266-4 cells. We also discovered that TBC1D7, a MITF target, promotes cell invasion in this pair of cultured melanoma cells partly through modulating MMP2 and MMP9 activities. Bioinformatic analyses of clinical data of melanoma patients support that elevated expression of TBC1D7 is significantly associated with poorer overall patient survival and distant metastasis-free survival, suggesting TBC1D7 as a potential driver for melanoma metastasis, at least for a large majority of melanoma patients. In addition, the differentially expressed proteins in primary/metastatic melanoma cells revealed from this study may allow for uncovering other potential modulators of melanoma metastasis.

Figure 2.1 SILAC-based quantitative proteomic experiment revealed differential expression of TBC domain-containing proteins in WM-115/WM-266-4 paired primary/metastatic melanoma cells.

(A) A flowchart showing the SILAC-based quantification of the global proteome of WM-115 (primary melanoma) and WM-266-4 (metastatic melanoma) cells. (B) A bar graph showing the differential expression of TBC domain-containing proteins in WM-115 and WM-266-4 cells.

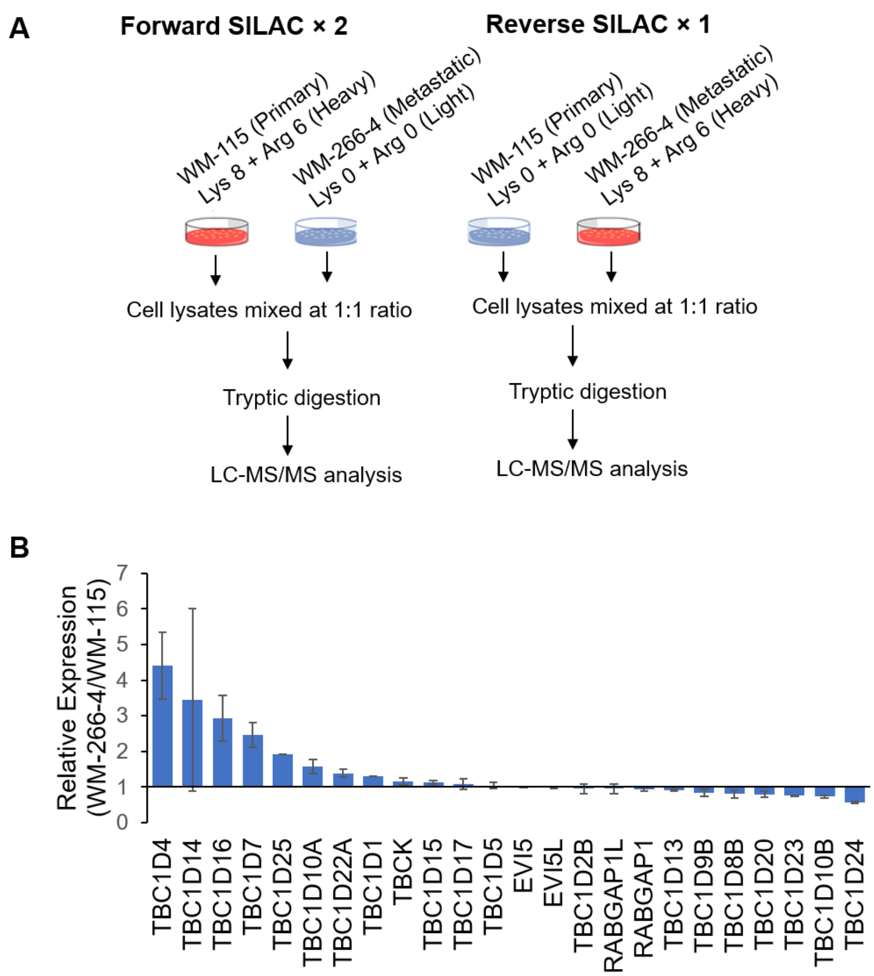


Figure 2.2 Gene Ontology (GO) analysis using DAVID for 1551 proteins that were differentially expressed by more than 1.5-fold in WM-115 and WM-266-4 cells. Top 10 pathways were displayed.

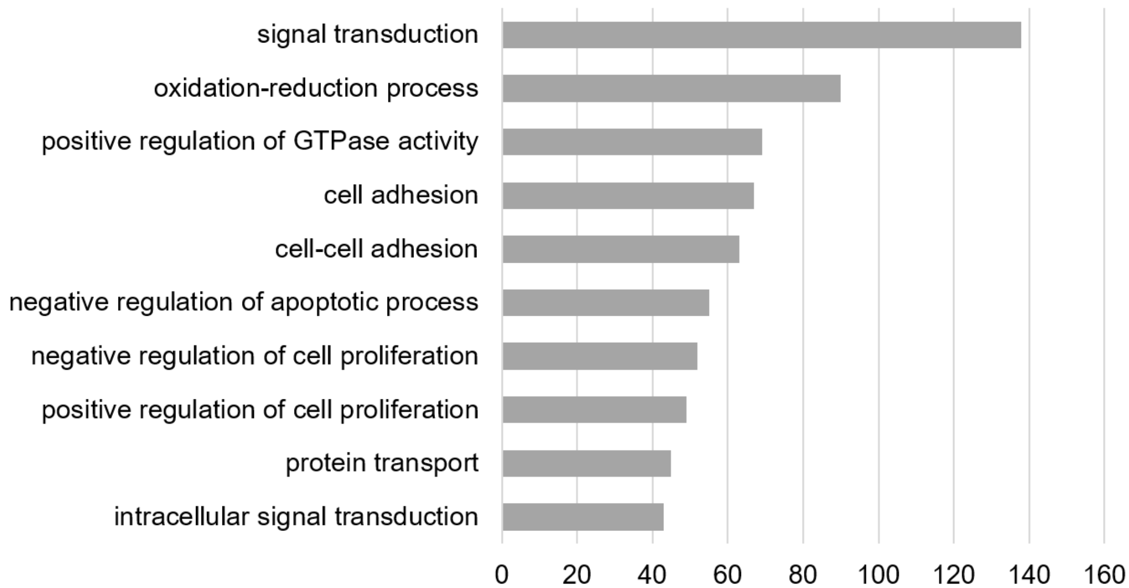
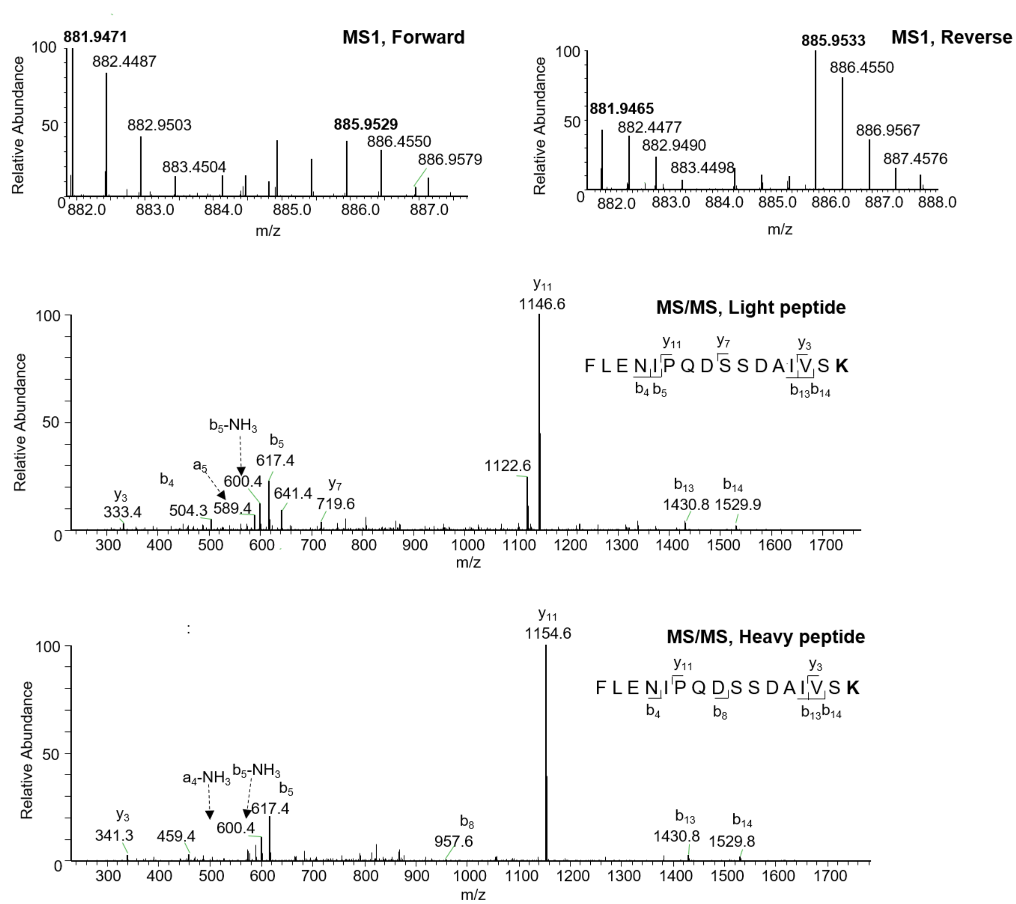


Figure 2.3 TBC1D7 exhibits elevated expression in metastatic melanoma cells relative to paired primary melanoma cells.

(A) Positive-ion ESI-MS for a representative tryptic peptide from TBC1D7, FLENIPQDSSDAIVSK, acquired from forward and reverse SILAC experiments, and the MS/MS for the $[M + 2H]^{2+}$ ions of light and heavy lysine-containing peptide. (B) Western blot for monitoring the expression of TBC1D7 protein in the two pairs of primary/metastatic melanoma cell lines. The data represent the mean ± S.D. (n = 3 or 4).

A


MS1, Forward
881.9471
882.4487
882.9503
883.4504
885.9529
886.4550
886.9579
MS1, Reverse
885.9533
886.4550
881.9465
882.4477
882.9490
883.4498
886.9567
887.4576
Relative Abundance
m/z
MS/MS, Light peptide
y11
1146.6
FLENIPQDSSSDAIVSK
b5-NH3
b5
617.4
a5
600.4
b4
641.4
y7
719.6
y3
333.4
504.3
589.4
1122.6
b13
b14
1430.8
1529.9
MS/MS, Heavy peptide
y11
1154.6
FLENIPQDSSSDAIVSK
a4-NH3
b5-NH3
b5
617.4
600.4
y3
341.3
459.4
b8
957.6
b13
b14
1430.8
1529.8

B

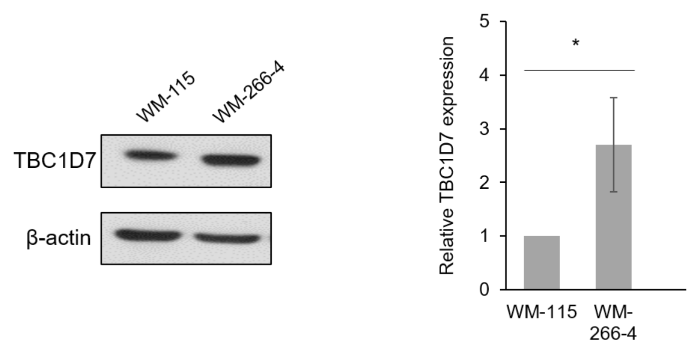

WM-115
WM-266-4
TBC1D7
β-actin
Relative TBC1D7 expression
*
WM-115
WM-266-4

Figure 2.4 Bioinformatic analysis suggests that TBC1D7 plays a role in melanoma progression.

(A) Kaplan-Meier plot for overall survival of skin cutaneous melanoma (SKCM) patients in the TCGA database. High and low expression levels refer to those patients with TBC1D7 expression being among the top and bottom quartiles of the TCGA-SKCM dataset, respectively (n = 404). (B) Kaplan-Meier plot for the distant metastasis-free survival of melanoma patients in the GSE65904 cohort (n = 150). Patients were stratified by the median mRNA expression level of *TBC1D7*. (C-E) Box-Whisker plot showed *TBC1D7* mRNA expression in SKCM and UVM patients in the TCGA database (C), in 60 melanoma cell lines in The Cancer Cell Line Encyclopedia (CCLE) database (D), and in 10 melanoma cell lines from the NCI-60 human tumor cell lines database (E). The horizontal edges of the box and line inside the box represent the top/bottom quartiles and median, respectively. The ends of the whisker denote the highest and lowest values. The survival analyses and *p* values in (A) and (B) were obtained using MedCalc, and all *p* values of (C-E) were calculated using the unpaired, two-tailed *t*-test.

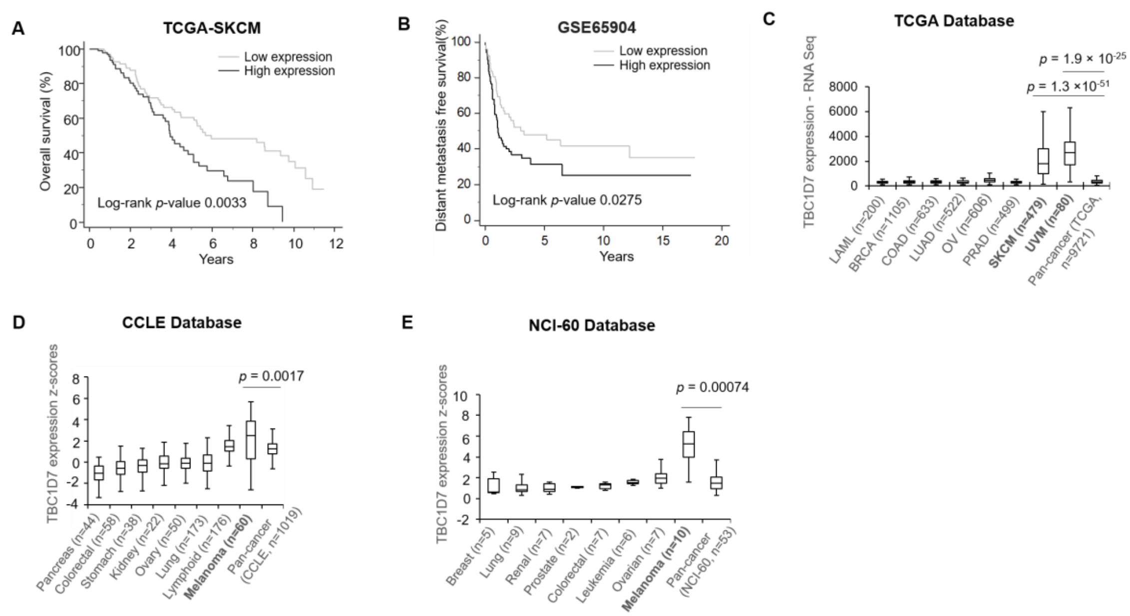


Figure 2.5 TBC1D7 enhances melanoma cell invasion *in vitro*, and it involves the alterations of enzymatic activities of secreted MMP2 and MMP9.

(A-B) Images and quantification results about the alterations in migration and invasion rates of WM-266-4 cells upon treatment with siTBC1D7 and non-targeting siRNA control (NTsiCtrl) (A), and WM-115 cells with ectopic overexpression of FLAG-TBC1D7 or empty vector control (B). (C-D) Gelatin zymography assays and quantification results for the enzymatic activities of secreted MMP2 and MMP9 in WM-115 and WM-266-4 cells upon modulation of TBC1D7 expression levels. "NS", $p > 0.05$; "∗∗", $0.001 \leq p < 0.01$; "∗∗∗", $p < 0.001$. The data represent the mean ± S.D. of results from three independent experiments. All p values were calculated using the unpaired, two-tailed t-test.

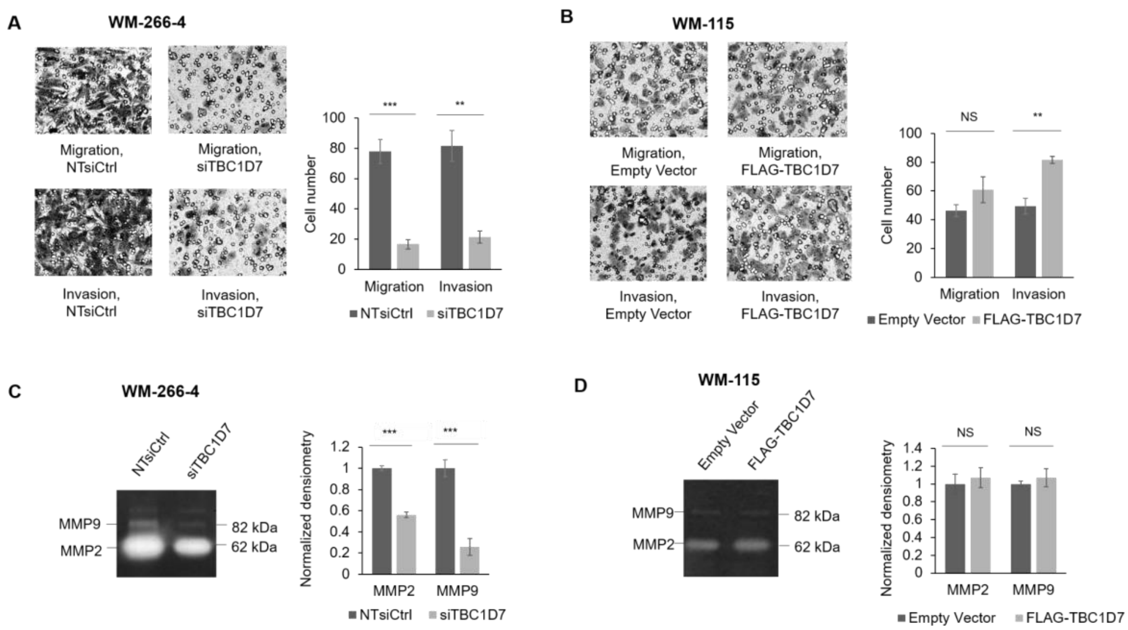


Figure 2.6 (A) Gelatin zymography assays for the cell lysates of WM-266-4 cells upon genetic depletion of TBC1D7 compared with NT siRNA control. (B) Gelatin zymography assays for the cell lysates of WM-115 cells upon ectopic overexpression of TBC1D7 compared with empty vector.

(C) Western blot results showed the validation of the knockdown efficiency of *TBC1D7* in WM-266-4. The data represent the mean ± S.D. of results from three independent experiments. “NS”, $p \geq 0.05$.

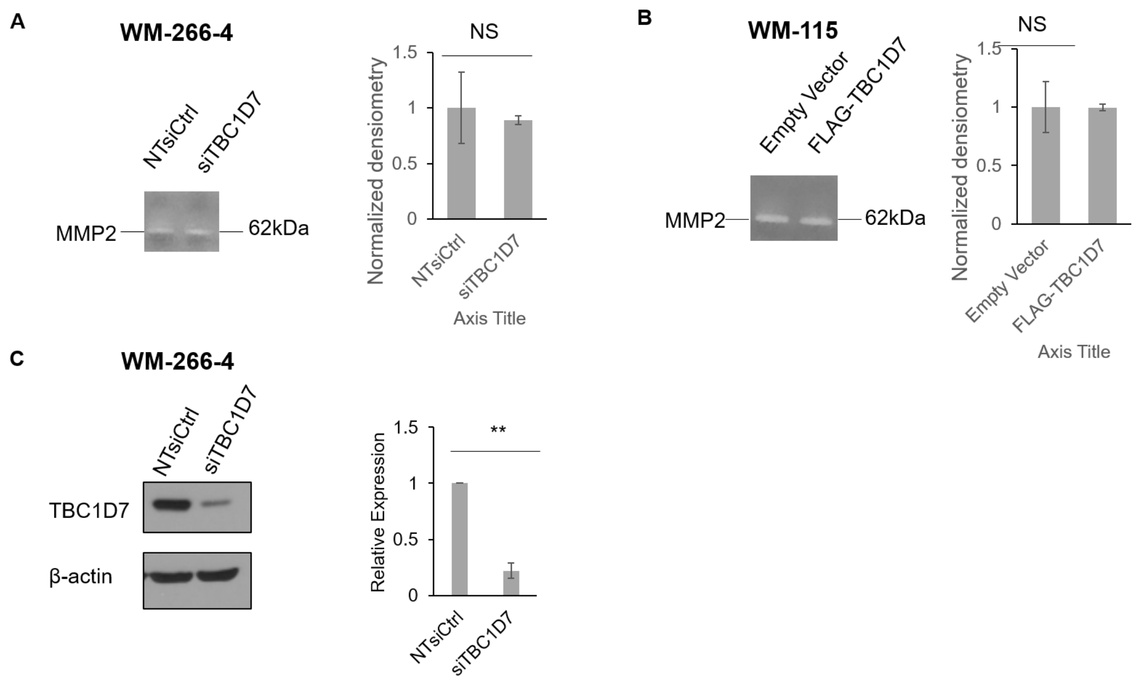


Figure 2.7 (A) Cell proliferation determined by MTT proliferation assay in siTBC1D7 treated WM-266-4 cells as compared to non-targeting siRNA control. (B) Cell proliferation determined by MTT proliferation assay in FLAG-TBC1D7 overexpressed WM-115 cells as compare to FLAG control.

Error bars represent means ± s.e.m in 4 replicates. "NS", $p \geq 0.05$. "∗∗", $0.001 \leq p < 0.01$; "∗∗∗", $p < 0.001$.

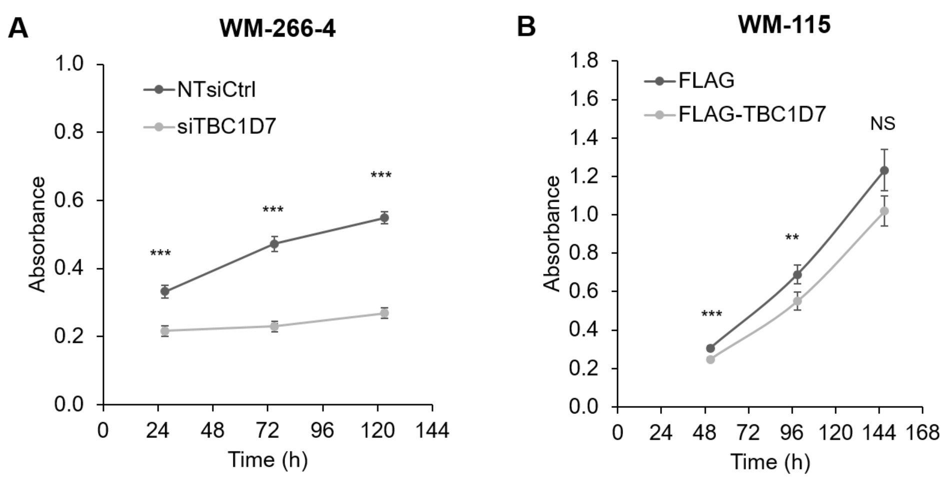
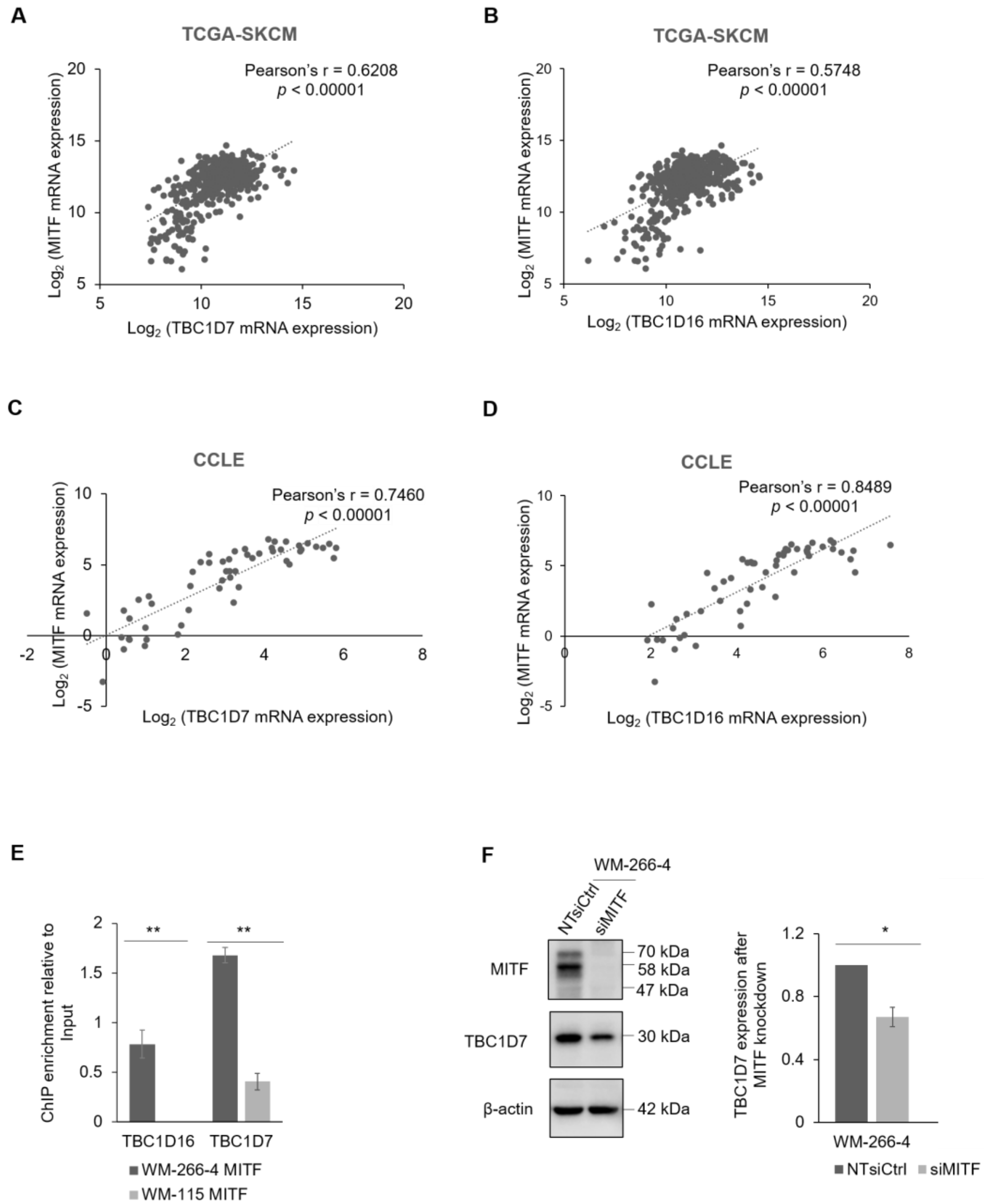


Figure 2.8 TBC1D7 expression is regulated by MITF.

(A-D) A scatter plot showing a positive correlation between mRNA expressions for *MITF* and *TBC1D7* genes (A), or *MITF* and *TBC1D16* genes (B) in the TCGA SKCM patients (n = 464), between mRNA expression levels of *MITF* and *TBC1D7* (C) or *TBC1D16* (D) gene in the CCLE melanoma cell lines (n = 54). Pearson's r and *p* values of A - D are calculated using the online calculator (www.socscistatistics.com/tests). (E) ChIP-qPCR showing the elevated enrichment of MITF in the promoter region of *TBC1D16* and *TBC1D7* in WM-266-4 cells relative to WM-115 cells. (F) Western blot for monitoring the expression level of TBC1D7 in WM-266-4 cells after siRNA-mediated knockdown of MITF. "∗", $p < 0.05$; "∗∗", $0.001 \leq p < 0.01$. The data represent the mean ± S.D. of results from three independent experiments. The *p* values in E and G were calculated using the unpaired, two-tailed *t*-test.



A
TCGA-SKCM
Pearson's r = 0.6208
p < 0.00001
Log2 (MITF mRNA expression)
Log2 (TBC1D7 mRNA expression)
B
TCGA-SKCM
Pearson's r = 0.5748
p < 0.00001
Log2 (MITF mRNA expression)
Log2 (TBC1D16 mRNA expression)
C
CCLE
Pearson's r = 0.7460
p < 0.00001
Log2 (MITF mRNA expression)
Log2 (TBC1D7 mRNA expression)
D
CCLE
Pearson's r = 0.8489
p < 0.00001
Log2 (MITF mRNA expression)
Log2 (TBC1D16 mRNA expression)
E
ChIP enrichment relative to Input
TBC1D16
TBC1D7
WM-266-4 MITF
WM-115 MITF
F
WM-266-4
NTsiCtrl
siMITF
MITF
TBC1D7
β-actin
70 kDa
58 kDa
47 kDa
30 kDa
42 kDa
TBC1D7 expression after MITF knockdown
WM-266-4
NTsiCtrl
siMITF

Figure 2.9 (A) A scatter plot shows a negative correlation between *TBC1D7* mRNA expression and promoter methylation in the TCGA-SKCM (n = 467) cohort.

(B) A scatter plot displays a negative correlation between *TBC1D16* mRNA expression and its promoter methylation in the TCGA-SKCM (n = 467) cohort. Spearman's rho and *p* values of A, B were calculated using the online calculator (www.socscistatistics.com/tests). (C) Kaplan-Meier survival analysis in the TCGA-SKCM. Patients were stratified by the methylation status of *TBC1D7*, where hypermethylation and hypomethylation refer to those patients with TBC1D7 methylation status being among the top and bottom quartile of the TCGA-SKCM dataset (n = 404, which consists of patients with less than 12 years of survival during follow-up). The survival curve and *p* values were generated using MedCalc.

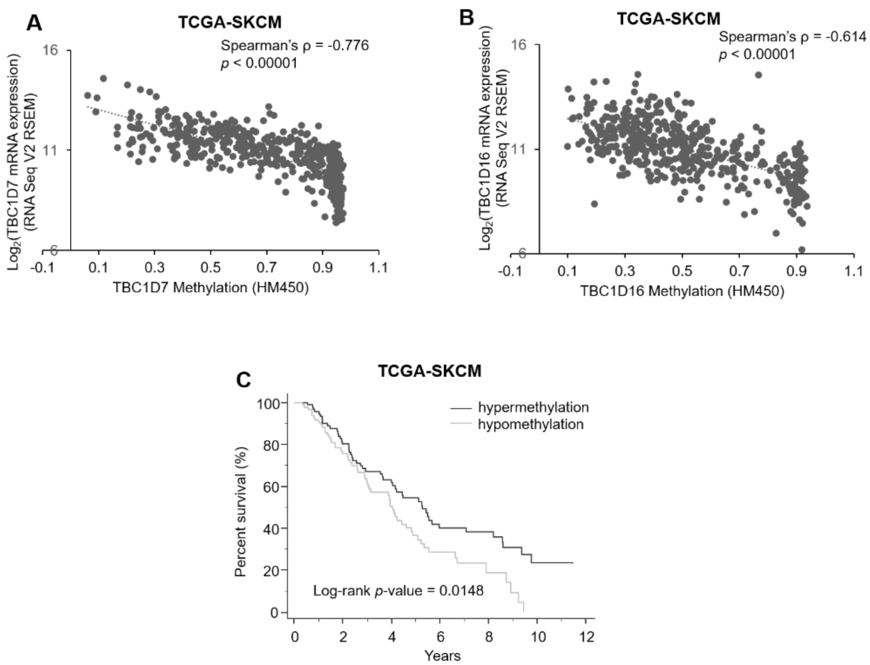


Table 2.1 A list of quantified TBC proteins and their relative expression ratios in WM-115 and WM-266-4 cells.

The results were obtained from the ProteinGroups txt file of MaxQuant searching results.

Gene Name	Protein Ratio (WM-115/WM-266-4), Forward-1 SILAC (F1)	Protein Ratio (WM-115/WM-266-4), Forward-2 SILAC (F2)	1/F1	1/F2	Protein Ratio (WM-266-4/WM-115), Reverse-1 (R1)	**Average Protein Ratio (WM-266-4/WM-115)**	S.D.	**R.S.D.**	Sequence Coverage (%)
EVI5	0.99	0.984	1.01	1.02	N/A	**1.01**	0.00	**0%**	6.1
EVI5L	1.03	1	0.97	1.00	N/A	**0.99**	0.02	**2%**	4.1
RABGAP1 (TBC1D11)	1.07	1.02	0.93	0.98	0.88	**0.93**	0.05	**5%**	30
RABGAP1L (TBC1D18)	0.94	1	1.06	1.00	0.8	**0.95**	0.14	**14%**	18
TBC1D1	0.77	N/A	1.30	N/A	N/A	**1.30**	N/A	**N/A**	3.8
TBC1D2B	N/A	0.95	N/A	1.05	0.86	**0.96**	0.14	**14%**	N/A
TBC1D4	0.22	0.19	4.55	5.26	3.41	**4.41**	0.93	**21%**	18
TBC1D5	0.9	0.94	1.11	1.06	0.94	**1.04**	0.09	**9%**	35
TBC1D7	0.37	0.45	2.70	2.22	N/A	**2.46**	0.34	**14%**	15.4
TBC1D8B	1.13	1.19	0.88	0.84	0.68	**0.80**	0.11	**13%**	7.8
TBC1D9B	1.17	1.1	0.85	0.91	0.73	**0.83**	0.09	**11%**	13.2
TBC1D10A	0.61	0.58	1.64	1.72	1.36	**1.57**	0.19	**12%**	33.1
TBC1D10B	1.45	1.33	0.69	0.75	0.76	**0.73**	0.04	**5%**	17.1
TBC1D13	1.09	1.09	0.92	0.92	0.87	**0.90**	0.03	**3%**	29.8
TBC1D14	0.19	N/A	5.26	N/A	1.63	**3.45**	2.57	**75%**	5.5
TBC1D15	0.88	0.86	1.14	1.16	1.08	**1.13**	0.04	**4%**	21.6
TBC1D16	0.34	0.28	2.94	3.57	2.28	**2.93**	0.65	**22%**	6.8
TBC1D17	0.83	0.91	1.20	1.10	0.92	**1.07**	0.14	**13%**	9.3
TBC1D20	1.17	1.36	0.85	0.74	N/A	**0.79**	0.08	**11%**	10
TBC1D22A	0.79	0.68	1.27	1.47	1.43	**1.39**	0.11	**8%**	15.7
TBC1D23	1.33	1.33	0.75	0.75	0.76	**0.75**	0.00	**1%**	14.6
TBC1D24	1.73	1.82	0.58	0.55	0.56	**0.56**	0.01	**3%**	13.8
TBC1D25	N/A	0.52	N/A	1.92	N/A	**1.92**	N/A	**N/A**	N/A
TBCK	0.93	0.82	1.08	1.22	N/A	**1.15**	0.10	**9%**	3.7
							Mean RSD	**12%**	

Reference

1. Enninga EAL, Moser JC, Weaver AL, Markovic SN, Brewer JD, Leontovich AA, et al. Survival of cutaneous melanoma based on sex, age, and stage in the United States, 1992-2011. Cancer Med. 2017;6:2203-12.

2. Chaffer CL, Weinberg RA. A perspective on cancer cell metastasis. Science. 2011;331:1559-64.

3. Hutagalung AH, Novick PJ. Role of Rab GTPases in membrane traffic and cell physiology. Physiol Rev. 2011;91:119-49.

4. Stenmark H, Olkkonen VM. The Rab GTPase family. Genome Biol. 2001;2:30071-5.

5. Bos JL, Rehmann H, Wittinghofer A. GEFs and GAPs: Critical elements in the control of small G proteins. Cell. 2007;129:865-77.

6. Richardson PM, Zon LI. Molecular-cloning of a cDNA with a novel domain present in the TRE-2 oncogene and the yeast-cell cycle regulators BUB2 and CDC16. Oncogene. 1995;11:1139-48.

7. Neuwald AF. A shared domain between a spindle assembly checkpoint protein and Ypt/Rab-specific GTPase-activators. Trends Biochem Sci. 1997;22:243-4.

8. Frasa MA, Koessmeier KT, Ahmadian MR, Braga VM. Illuminating the functional and structural repertoire of human TBC/RABGAPs. Nat Rev Mol Cell Biol. 2012;13:67-73.

9. Behrends C, Sowa ME, Gygi SP, Harper JW. Network organization of the human autophagy system. Nature. 2010;466:68-76.

10. Yoshimura S, Egerer J, Fuchs E, Haas AK, Barr FA. Functional dissection of Rab GTPases involved in primary cilium formation. J Cell Biol. 2007;178:363-9.

11. Hsu C, Morohashi Y, Yoshimura S, Manrique-Hoyos N, Jung S, Lauterbach MA, et al. Regulation of exosome secretion by Rab35 and its GTPase-activating proteins TBC1D10A-C. J Cell Biol. 2010;189:223-32.

12. Pei L, Peng Y, Yang Y, Ling XB, Van Eyndhoven WG, Nguyen KC, et al. PRC17, a novel oncogene encoding a Rab GTPase-activating protein, is amplified in prostate cancer. Cancer Res. 2002;62:5420-4.

13. Chen M, Sheng XJ, Qin YY, Zhu S, Wu QX, Jia L, et al. TBC1D8 Amplification Drives Tumorigenesis through Metabolism Reprogramming in Ovarian Cancer. Theranostics. 2019;9:676-90.

14. Vizoso M, Ferreira HJ, Lopez-Serra P, Carmona FJ, Martinez-Cardus A, Girotti MR, et al. Epigenetic activation of a cryptic TBC1D16 transcript enhances melanoma progression by targeting EGFR. Nat Med. 2015;21:741-50.

15. 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;11:429-39.

16. Ong SE, 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;1:376-86.

17. Wisniewski JR, Zougman A, Nagaraj N, Mann M. Universal sample preparation method for proteome analysis. Nat Meth. 2009;6:359-62.

18. Guo L, Xiao YS, Wang YS. Hexavalent chromium-induced alteration of proteomic landscape in human skin fibroblast cells. J Proteome Res. 2013;12:3511-8.

19. 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;26:1367-72.

20. Alfaiz AA, Micale L, Mandriani B, Augello B, Pellico MT, Chrast J, et al. TBC1D7 mutations are associated with intellectual disability, macrocrania, patellar dislocation, and celiac disease. Hum Mutat. 2014;35:447-51.

21. Sato N, Koinuma J, Ito T, Tsuchiya E, Kondo S, Nakamura Y, et al. Activation of an oncogenic TBC1D7 (TBC1 domain family, member 7) protein in pulmonary carcinogenesis. Genes Chromosomes Cancer. 2010;49:353-67.

22. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012;2:401-4.

23. Damsky WE, Rosenbaum LE, Bosenberg MJC. Decoding melanoma metastasis. 2010;3:126-63.

24. Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 2009;37:1-13.

25. Huang M, Qi TYF, Li L, Zhang G, Wang YS. A Targeted Quantitative Proteomic Approach Assesses the Reprogramming of Small GTPases during Melanoma Metastasis. Cancer Res. 2018;78:5431-45.

26. Huang M, Wang YS. Targeted Quantitative Proteomic Approach for Probing Altered Protein Expression of Small GTPases Associated with Colorectal Cancer Metastasis. Anal Chem. 2019;91:6233-41.

27. Cai R, Huang M, Wang YS. Targeted Quantitative Profiling of GTP-Binding Proteins in Cancer Cells Using Isotope-Coded GTP Probes. Anal Chem. 2018;90:14339-46.

28. Liotta LA, Stetlerstevenson WG. Tumor invasion and metastasis: an imbalance of positive and negative regulation. Cancer Res. 1991;51:S5054-S9.

29. Kalluri R. Basement membranes: Structure, assembly and role in tumour angiogenesis. Nat Rev Cancer. 2003;3:422-33.

30. Yee C, Shiu RPC. Degradation of endothelial basement membrane by human breast cancer cell lines. Cancer Res. 1986;46:1835-9.

31. Levy C, Khaled M, Fisher DE. MITF: master regulator of melanocyte development and melanoma oncogene. Trends Mol Med. 2006;12:406-14.

32. Tirosh I, Izar B, Prakadan SM, Wadsworth MH, Treacy D, Trombetta JJ, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science. 2016;352:189-96.

3. Chapter 3. Targeted Profiling of Epitranscriptomic Reader, Writer and Eraser Proteins Accompanied with Radioresistance in Breast Cancer Cells

3.1 Introduction

Unlike the extensively studied DNA methylation and histone post-translational modifications, the investigations about RNA modifications did not gain wide attention in the scientific community until the availability of high-throughput sequencing method rendered transcriptome-wide profiling of N^6-methyladenosine (m^6A) in 2012 (1). RNA is known to contain more than 170 types of modifications, among which m^6A is the most abundant internal modification in mRNA (2). m^6A-modifying enzymes ("writers" and "erasers") install or remove m^6A, whereas m^6A-binding proteins ("readers") recognize m^6A to confer downstream effects. m^6A is involved in regulating various cellular processes, including mRNA stability, splicing, translation, and decay (3-6). Aside from m^6A, other RNA modifications also regulate biological processes through their reader, writer, and eraser (RWE) proteins. For instance, ALYREF and YTHDF2, which are 5-methylcytidine (m^5C) reader proteins, modulate mRNA export and rRNA maturation, respectively (7, 8). In addition, NSUN2 (m^5C writer) and YBX1 (m^5C reader) drive the pathogenesis of human bladder urothelial carcinoma by targeting the m^5C site in the mRNA of *HDGF* gene (9).

Breast cancer represents the second most common cancer among women in the United States. Radiation therapy harnesses ionizing radiation to eliminate local malignant cells and prevent cancer recurrence. It delivers high-energy X-rays to target tissues and elicits

DNA damage in rapidly dividing cancer cells. Although more than 83% of breast cancer patients benefit from radiation therapy (10), some patients suffer from tumor recurrence due to the development of resistance to radiation therapy (11). Many genes involved in DNA damage repair and cell cycle checkpoints have been documented to modulate radioresistance, including *AKT*, *HER2*, *BRCA2*, *CDK1*, and *CHK1* (12-15).

Several studies also unveiled the functions of m^6A RWE proteins in modulating radioresistance of cancer cells. METTL3, the catalytic subunit of the major m^6A writer complex, promotes radioresistance in glioblastoma by regulating m^6A modification of *SOX2* mRNA and enhancing its stability (16). m^6A eraser ALKBH5 augments radioresistance by modulating homologous recombination in glioblastoma (17). m^6A reader YTHDC2 promotes radioresistance of nasopharyngeal carcinoma via enhancing *IGF1R* mRNA and activating the IGF1R-AKT/S6 signaling pathway (18). Little, however, is known about the roles of other epitranscriptomic RWE proteins, such as those for *N*1-methyladenosine (m^1A), m^5C, and pseudouridine (Ψ) in RNA, in modulating the sensitivity of cancer cells to radiation therapy.

Parallel-reaction monitoring (PRM)-based targeted proteomics, which can be performed on hybrid quadrupole-Orbitrap or quadrupole time-of-flight (TOF) mass spectrometers, can be used to quantify hundreds of peptides in complex sample matrices in a single LC-MS/MS run (19). Since the MS/MS are acquired on a high-resolution mass analyzer, PRM offers highly selective and reliable identification and quantification of target peptides. Moreover, the mass spectrometer can be programed to collect MS/MS of precursor ions in predefined retention time windows with the use of normalized retention time (iRT), which provides improved throughput of the LC-PRM method (20).

3.2 Materials and Methods

3.2.1 Cell Culture and SILAC

The radioresistant clones (C5 and C6) of MDA-MB-231 and MCF-7 cells were generated previously (21, 22). MDA-MB-231/C5 and MCF-7/C5 paired cells were cultured in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% fetal bovine serum (Invitrogen-Gibco) and 1% penicillin/streptomycin. Cells were kept at 37 °C in a humidified atmosphere containing 5% CO_2. For the SILAC labelling experiments, the MDA-MB-231/C5 and MCF-7/C6 paired cells were cultured in the light DMEM media (i.e., normal lysine and arginine), or the heavy DMEM media i.e., [$^{13}C_6$, $^{15}N_2$]-L-lysine and [$^{13}C_6$]-L-arginine, with the addition of 10% dialyzed FBS (Invitrogen) and 1% penicillin/streptomycin, for at least 21 days to complete incorporate isotope-labeled amino acids.

3.2.2 Tryptic digestion of whole cell lysates

After the complete SILAC labelling, MDA-MB-231/C5 and MCF-7/C6 paired cells were lysed on ice for 30-min using CelLytic M cell lysis reagent (Sigma) supplemented with 1% protease inhibitor cocktail, and centrifuged at 9000 g for 30 min at 4°C. The supernatants were collected, and the protein concentrations in the supernatants were quantified using the Bradford assay. In the forward SILAC labelling experiments, light-isotope-labelled C5 and C6 cell lysates were mixed at 1:1 ratio (by mass) with heavy-isotope-labelled MDA-MB-231 and MCF-7 cell lysates, respectively. In the reverse

SILAC labelling experiments, light-isotope-labelled MDA-MB-231 and MCF-7 cell lysates were mixed at 1:1 ratio (by mass) with heavy-isotope-labelled C5 and C6 cell lysates, respectively. Two forward and two reverse labelling experiments were carried out for MDA-MB-231/C5 and MCF-7/C6 pairs of cells.

Following the filter-aided sample preparation (FASP) protocol (23), 50 µg of protein samples (25 µg of heavy- and light-isotope-labelled cell lysate combined) were denatured twice in 8 M urea in 50 mM NH_4HCO_3 in the polyethersulfone (PES) Membrane 30 kDa centrifugal filter unit (VWR) by centrifuging at 11,000g for 25 min. The denatured samples were reduced with dithiothreitol at 37 °C for 1 hr, alkylated with iodoacetamide at room temperature for 30 min, followed by washing twice with 50 mM NH_4HCO_3. The samples were digested with MS-grade trypsin (Pierce) at 1:50 ratio (trypsin: protein, by mass) in 50 mM NH_4HCO_3 at 37 °C overnight. The tryptic peptides were collected by centrifugation, dried in a Speed-vac, desalted using OMIX C18 pipet tips (Agilent Technologies), and redissolved in 0.1% formic acid for LC-PRM analysis.

3.2.3 Establishment of PRM library

A PRM library containing unique tryptic peptides from 152 epitranscriptomic RWE proteins was established in Skyline (24). Two or three unique peptides exhibiting high intensities in previously published shotgun proteomic data were selected to represent each RWE protein in the PRM library (25), and the MS/MS of these peptides were deposited into the PRM library (25). Additionally, iRT of each peptide was derived from the linear regression of RT with iRT of tryptic peptides of BSA (with defined iRT) analyzed under the same chromatographic conditions.

3.2.4 LC-PRM data acquisition

Samples were subjected to LC-PRM analysis on a Q Exactive Plus quadrupole-Orbitrap mass spectrometer coupled with a Dionex UltiMate 3000 RSLCnano UPLC system. The analytical column was packed in-house using 3 μm Reprosil-Pur C18-AQ resin (Dr. Maisch GmbH HPLC) in a ~ 25-cm long, 75 μm i.d. fused silica column. The trapping column was also prepared in-house using 5 μm Reprosil-Pur C18-AQ resin (Dr. Maisch GmbH HPLC) in a 4-cm long, 150 μm i.d. fused silica column. SILAC samples (500 ng) were separated with a 125-min linear gradient from 6 – 43% mobile phase B (80% acetonitrile in 0.1% formic acid) at a flow rate of 300 nL/min. The spray voltage was 1.8 kV.

Before sample analysis, the tryptic digestion mixture of BSA was analyzed using the LC-PRM method under the same experimental settings, but with *m/z* values of ten tryptic peptides of BSA in the inclusion list. After importing the acquired data for BSA peptides to Skyline, three inclusion lists with *m/z* and a 7-min RT window of each precursor ion were generated and exported from Skyline with the maximum number of concurrent precursor ions being set at 40. Those inclusion lists were imported for LC-PRM analysis on the Q Exactive Plus mass spectrometer, where the precursor ions were distributed in three separate LC-PRM runs. The precursor ions were isolated in the quadrupole at an isolation window of 1.0 *m/z*, fragmented in the HCD collision cell at a normalized collision energy (NCE) of 28. Other settings were: MS/MS resolution, 17,500; automated gain control (AGC) target, 1×10^5; maximum accumulation time: 50 ms.

3.2.5 LC-PRM data processing

The acquired LC-PRM data were imported to Skyline. In Skyline, the acquired MS/MS of each precursor ion was compared with that in the spectral library, where similarity is measured by dot product (dotp) value (26). A dotp value of > 0.7 is imposed for positive peptide identification. In addition, 4-6 fragment ions in the light and heavy forms should share the same retention time. The potential interfering fragment ions that do not overlay with other fragment ions were manually excluded (i.e., processed data). The SILAC ratios of each precursor ion were calculated automatically in Skyline.

3.2.6 Western blots

MDA-MB-231/C5 and MCF-7/C6 pairs of breast cancer cells were lysed with CelLytic M cell lysis reagent (Sigma) supplemented with 1% protease inhibitor cocktail, and denatured at 95 °C for 5-min with Laemmli loading buffer. The same amount of proteins (10-20 μg) of denatured lysates were separated using SDS-PAGE, and transferred onto a nitrocellulose membrane at 90 V for 60-min at 4 °C. The membrane was blocked with 5% milk in PBS-T (PBS with 0.1% Tween 20) for 45-min, and incubated separately with primary antibodies that recognize human FTO (Abclonal, A1438, 1:1000), TRMT1 (Abclonal, A7116, 1:1000), and GAPDH (Santa Cruz, sc-32233, 1:10,000) at 4 °C overnight. After several thorough washes with PBS-T, the membrane was incubated with donkey anti-rabbit secondary antibody (Sigma, A0545, 1:5,000), or anti-mouse secondary antibody (Santa Cruz, m-IgGκ BP-HRP, 1:5,000), followed by several thorough washes with PBS-T. The protein bands were visualized using Amersham ECL™ Western Blot Detecting Reagent (GE Healthcare).

3.2.7 Bioinformatic analyses

All Kaplan-Meier survival analyses of the TCGA and METABRIC cohorts were carried out in the MedCal software (https://www.medcalc.org/). GSEA enrichment plots were generated in GSEA 4.1.0 software (http://www.broad.mit.edu/gsea) TCGA-BRAC dataset, downloaded from the Xenahubs database (https://gdc.xenahubs.net/download/TCGA-BRCA.htseq_fpkm.tsv.gz), was sorted by mRNA expression level of *TRMT1* from high to low with median value as cutoff. Patients (n = 1,217) were therefore categorized into high- and low-TRMT1-expression group. Gene set enrichment analysis (GSEA) of the stratified TCGA dataset was carried out against the hallmark gene sets (h.all.v7.4.symbols.gmt) which were downloaded from GSEA Molecular Signatures Database (http://www.gsea-msigdb.org/gsea/msigdb/collections.jsp#H). The number of permutations was set at 1000. A gene set with a false discovery rate less than 0.25 was considered significantly enriched.

3.3 Results and Discussion

3.3.1 Development of LC-PRM Method for Profiling a Total of 152 Epitranscriptomic RWE Proteins and Application to Radioresistance Breast Cancer Cells

To investigate systematically the roles of epitranscriptomic RWE proteins in modulating radioresistance in breast cancer, we established an LC-PRM method, coupled with stable isotope labelling by amino acids in cell culture (SILAC), to examine the differences in expression levels of the proteins in MDA-MB-231 and MCF-7 breast cancer cells relative to their corresponding radioresistant C5 and C6 clones (Figure 3.1a). We first developed a Skyline (24) PRM library, which includes all the 68 human epitranscriptomic RWE proteins deposited in the Modomics database (2), and another 84 RWE proteins retrieved from several recent review articles (Figure 3.1b, Table 3.1) (27-32). Each RWE protein is represented by two or three unique peptides, whose MS/MS were acquired from previously published shotgun proteomic analyses and imported into the Skyline library (25).

3.3.2 Quantification of Around 100 RWE Proteins in the MDA-MB-231/C5 and the MCF-7/C6 Pairs using LC-PRM

To achieve high-throughput analysis of these proteins, we employed scheduled LC-PRM with a 7-min retention time window and a maximum of 40 concurrent precursor ions. In this vein, iRT of each peptide was derived from the linear regression of RT with iRT by analyzing a tryptic digestion mixture of bovine serum albumin (BSA) under the same chromatographic conditions. With this method, the 152 epitranscriptomic RWE proteins (i.e., 444 unique peptides, and 888 precursor ions for SILAC) could be monitored in three LC-MS/MS runs with a 125-min gradient. The LC-PRM analysis enabled the quantifications of 106 and 99 epitranscriptomic RWE proteins from two forward and two reverse SILAC experiments in the MDA-MB-231/C5 and the MCF-7/C6 pairs, respectively, accounting for approximately 70% and 65% of proteins in the PRM library (Figure 3.1c). The quantification result of each RWE protein was calculated from the average ratios of all detected tryptic peptides of the protein, where the ratio of each peptide was calculated in Skyline based on LC-PRM results from the four replicates of SILAC experiments. A total of 96 epitranscriptomic RWE proteins were commonly quantified in the two pairs of cell lines. We also performed hierarchical clustering analysis to illustrate the differential expression of the quantified epitranscriptomic RWE proteins in the radioresistant C5 and C6 lines relative to the corresponding parental MDA-MB-231 and

MCF-7 lines (Figure 3.2). Such analysis revealed similarities and differences in alterations in expression of epitranscriptomic RWE proteins accompanied with the development of radioresistance in the two breast cancer cell lines (Figure 3.2).

Our LC-PRM data revealed that 8 and 11 epitranscriptomic RWE proteins were down-regulated by more than 1.5-fold, and 18 and 27 epitranscriptomic RWE proteins were up-regulated by over 1.5-fold in the radioresistant C5 and C6 lines relative to their corresponding parental lines, respectively (Figure 3.3a, b). Gene Ontology (GO) analysis of these differentially expressed proteins showed that the up-regulated epitranscriptomic RWE proteins are mainly involved in tRNA modification, tRNA processing, and rRNA base methylation (Figure 3.4a). The down-regulated epitranscriptomic RWE proteins play roles in tRNA methylation, oxidation-reduction process, and tRNA dihydrouridine synthesis (Figure 3.4a). In this context, it is worth noting that over 100 types of modifications have been detected in tRNA (2), including Ψ, m^1A, *N*1-methylguanosine (m^1G), and N^6-threonyl-carbamoyl-adenosine (t^6A), where many tRNA modifications regulate the stabilities of tRNA (33, 34).

3.3.3 Eight RWE Proteins Commonly Altered by Over 1.5-fold in Both Pairs of Breast Cancer Cell Lines

Among the epitranscriptomic RWE proteins that are up- or down-regulated by at least 1.5-fold, eight were commonly altered in both pairs of breast cancer cell lines (Figure 3.3c, d). For instance, MRM1 was pronouncedly down-regulated, whereas FTO and CTU1 were markedly up-regulated in radioresistant lines compared to parental lines in both pairs of breast cancer cells (Figure 3.2). Figure 3.5a illustrates the PRM traces of representative peptides from TRMT1 and FTO, two of the eight commonly altered proteins, in two pairs of matched radioresistant/parental breast cancer cells. The up-regulations of TRMT1 and FTO in the radioresistant cells were further validated by Western blot analysis (Figure 3.5b).

3.3.4 LC-PRM Enabling Highly Efficient, Selective, Sensitive, and Reproducible Peptide Quantification

Our proteomic results showed that the established LC-PRM method coupled with SILAC affords highly efficient, selective, sensitive, and reproducible peptide quantification. The efficiency of the method is manifested by its high throughput, where 888 precursor ions of 444 tryptic peptides derived from the 152 epitranscriptomic RWE proteins could be monitored in three LC-MS/MS runs. Additionally, the high consistency of quantification results of TRMT1 and FTO obtained from PRM and Western blot analyses underscores the high accuracy of the method. Moreover, the relatively high coverage (i.e., 70% and 65%) of the epitranscriptomic RWE proteins in the library indicates the high sensitivity of the PRM method. The PRM method is also highly reproducible, as reflected by the small mean relative standard deviations of the quantification results obtained from two forward and two reverse SILAC experiments, i.e., 11.7% and 9.1% in the MDA-MB-231/C5 and MCF-7/C6 and pairs of breast cancer cells, respectively. In this context, it is worth noting that our PRM method does not take into account post-translational modifications (PTMs) in the peptides employed for the quantifications of the epitranscriptomic RWE proteins. Hence, differences in PTMs between the radioresistant and parental breast cancer cells may contribute, in part, to variations in quantification results obtained from different peptides of the same protein.

3.3.5 Correlation of *TRMT1* mRNA Expression with Breast Cancer Patient Survival Who Received Radiation Therapy and DNA Repair Gene Sets

Considering that the above-mentioned proteomic results were acquired from breast cancer cell lines derived from two patients, we next asked if the findings could be extended to breast cancer patients in general. To this end, we performed Kaplan-Meier survival analyses in two breast cancer patient cohorts, i.e., The Cancer Genome Atlas-Breast Invasive Carcinoma (TCGA-BRCA) and Molecular Taxonomy of Breast Cancer International Consortium (METABRIC). We placed our emphasis on patients who received radiation therapy and explored the correlation between the mRNA expression level of each commonly altered epitranscriptomic RWE protein and patient survival. Our results showed that a higher level of mRNA expression of *TRMT1* is significantly correlated with poorer survival of breast cancer patients who received radiation therapy in both TCGA-BRCA and METABRIC cohorts (Figure 3.5c). This result is in keeping with our proteomic data showing that TRMT1 is up-regulated in C5 and C6 cells compared with parental MDA-MB-231 and MCF-7 cells. For the other commonly altered epitranscriptomic RWE proteins in both pairs, only the Kaplan-Meier survival analysis of *CTU1* gene in the METABRIC cohorts who received radiation therapy corroborates with proteomics results (Figure 3.4b). The lack of correlation for other proteins may be due to the differences in the mRNA and protein expression levels of epitranscriptomic RWE proteins, and/or the heterogeneity of breast cancer (35).

To explore the potential mechanism of TRMT1 in radioresistant breast cancer, we carried out gene set enrichment analysis (GSEA). TCGA-BRCA dataset was stratified by the high and low mRNA expression of *TRMT1* using its median value as a cutoff. Upon performing GSEA between the stratified TCGA dataset and the hallmark gene sets downloaded from the GSEA Molecular Signatures Database (36), we observed that, among 23 gene sets, four are significantly (at FDR < 25%) up-regulated in the high-TRMT1-expression group. DNA repair gene set is the most significantly enriched (Figure 3.5d). Since radioresistance is known to be associated with the enhanced ability to repair radiation-induced DNA damage (11), this finding again suggests a role of TRMT1 in promoting radioresistance. Additionally, two other hallmark gene sets, i.e., Myc_target_V2 (Figure 3.5d) and Myc_target_V1 (Figure 3.4c), were also enriched significantly with the high-TRMT1-expression group; hence, TRMT1 may be associated with Myc target genes. Moreover, the hallmark gene set UV_response_up, i.e., up-regulated in response to ultraviolet (UV) radiation, was also associated with high expression of TRMT1 (Figure 3.4d).

Radiation therapy is known to enhance cancer metastasis through activating epithelial-mesenchymal transition (EMT) transcription factors, including Snail, Slug, ZEB1, and ZEB (37). Additionally, radioresistant breast cancer cells exhibit increased metastatic potential (38), and breast cancer distant metastasis was shown to promote resistance to radiation therapy (39). Based on the observed co-occurrence between metastasis and radioresistance, several reports interrogated their cross-regulation and revealed several common pathways, including PI3K/AKT/mTOR, MAPK, Wnt/β-catenin, NF-κB, EMT, and reactive oxygen species scavenging (40-44).

TRMT1 dimethylates the N^2 position of guanosine 26 in most tRNAs to give $m^{2,2}G$. It was documented that urinary level of $m^{2,2}G$ was elevated in 35.1% or 57% in two cohorts of metastatic breast cancer patients (45, 46). TRMT1 is the only known writer of $m^{2,2}G$ in humans (47); thus, the augmented levels of $m^{2,2}G$ in metastatic breast cancer patients also suggest a role of TRMT1 in the metastatic transformation of breast cancer.

3.4 Conclusion

In summary, we established, for the first time, a high-throughput scheduled LC-PRM method for profiling simultaneously a total of 152 epitranscriptomic RWE proteins. We also employed this method to explore the roles of these proteins in radioresistance in breast cancer cells, we found that eight epitranscriptomic RWE proteins were commonly altered by over 1.5-fold in the MDA-MB-231/C5 and MCF-7/C6 pairs of breast cancer cells. Among them, TRMT1 may play a role in promoting radioresistance in breast cancer and be involved in breast cancer metastatic transformation. Thus, TRMT1 could be a target for overcoming radioresistance in breast cancer therapy. In addition, other differentially expressed epitranscriptomic RWE proteins in matched radioresistant/parental breast cancer cell lines revealed from this study may provide a comprehensive understanding of epitranscriptomic RWE proteins in modulating radiation sensitivity in breast cancer. Moreover, we envision that the LC-PRM method developed in this study can also be employed to examine, in the future, the roles of epitranscriptomic RWE proteins in the metastatic transformation of cancer and therapeutic resistance of other types of cancer.

Table 3.1 A list of epitranscriptomic RWE proteins included in the PRM library. ALKHB2, 4, 6, and 7, with unknown functions in RNA modifications, were also listed.

Ensembl Gene ID	**Protein Name**	**Description**	**Known Functions in RNA modifications**	**References**
ENSG00000160710	ADAR	adenosine deaminase RNA specific	A-to-I writer	(48)
ENSG00000065457	ADAT1	adenosine deaminase tRNA specific 1	A-to-I writer	(48)
ENSG00000189007	ADAT2	adenosine deaminase tRNA specific 2	A-to-I writer	(48)
ENSG00000213638	ADAT3	adenosine deaminase tRNA specific 3	A-to-I writer	(48)
ENSG00000100601	ALKBH1	alkB homolog 1, histone H2A dioxygenase	m^1A, m^5C eraser	(49, 50)
ENSG00000189046	ALKBH2	alkB homolog 2, alpha-ketoglutarate dependent dioxygenase	potential eraser	
ENSG00000166199	ALKBH3	alkB homolog 3, alpha-ketoglutarate dependent dioxygenase	m^1A eraser	(51)
ENSG00000160993	ALKBH4	alkB homolog 4, lysine demethylase	potential eraser	
ENSG00000091542	ALKBH5	alkB homolog 5, RNA demethylase	m^6A eraser	(52)
ENSG00000239382	ALKBH6	alkB homolog 6	potential eraser	

ENSG00000125652	ALKBH7	alkB homolog 7	potential eraser	
ENSG00000137760	ALKBH8	alkB homolog 8, tRNA methyltransferase	mcm^5U, mcm^5Um, $mchm^5U$, and mcm^5s^2U writer	(53)
ENSG00000183684	ALYREF	Aly/REF export factor	m^5C reader	(7)
ENSG00000244509	APOBEC3C	apolipoprotein B mRNA editing enzyme catalytic subunit 3C	C-to-U writer	(54)
ENSG00000239713	APOBEC3G	apolipoprotein B mRNA editing enzyme catalytic subunit 3G	C-to-U writer	(55)
ENSG00000186666	BCDIN3D	BCDIN3 domain containing RNA methyltransferase	5′ monophosphate methylation writer	(56)
ENSG00000071462	BUD23	BUD23 rRNA methyltransferase and ribosome maturation factor	m^7G writer	(57)
ENSG00000105879	CBLL1	Cbl proto-oncogene like 1	m^6A writer complex	(58)
ENSG00000101391	CDK5RAP1	CDK5 regulatory subunit associated protein 1	ms^2i^6A writer	(59)
ENSG00000145996	CDKAL1	CDK5 regulatory subunit associated protein 1 like 1	ms^2t^6A writer	(60)
ENSG00000144021	CIAO1	cytosolic iron-sulfur assembly component 1	s^2U, mcm^5s^2U writer	(61)

ENSG00000137200	CMTR1	cap methyltransferase 1	N_m writer	(62)
ENSG00000180917	CMTR2	cap methyltransferase 2	N_m writer	(62)
ENSG00000142544	CTU1	cytosolic thiouridylase subunit 1	s^2U, mcm^5s^2U writer	(63)
ENSG00000174177	CTU2	cytosolic thiouridylase subunit 2	s^2U, mcm^5s^2U writer	(63)
ENSG00000172795	DCP2	decapping mRNA 2	m^7GpppN eraser	(64)
ENSG00000110063	DCPS	decapping enzyme, scavenger	m^7GpppN eraser	(65)
ENSG00000086189	DIMT1	DIMT1 rRNA methyltransferase and ribosome maturation factor	$m_2^{6,6}A$ writer	(66)
ENSG00000144535	DIS3L2	DIS3 like 3'-5' exoribonuclease 2	uridylation reader	(67)
ENSG00000130826	DKC1	dyskerin pseudouridine synthase 1	Ψ writer	(68)
ENSG00000169718	DUS1L	dihydrouridine synthase 1 like	D writer	(69)
ENSG00000167264	DUS2	dihydrouridine synthase 2	D writer	(69)
ENSG00000141994	DUS3L	dihydrouridine synthase 3 like	D writer	(70)
ENSG00000107581	EIF3A	eukaryotic translation initiation factor 3 subunit A	m^6A reader	(71)

ENSG00000106263	EIF3B	eukaryotic translation initiation factor 3 subunit B	m^6A reader	(71)
ENSG00000066044	ELAVL1	ELAV like RNA binding protein 1	m^6A reader	(72)
ENSG00000070061	ELP1	elongator acetyltransferase complex subunit 1	cm^5U, ncm^5U, mcm^5U, mcm^5s^2U writer	(61, 73)
ENSG00000134014	ELP3	elongator acetyltransferase complex subunit 3	cm^5U, ncm^5U, mcm^5U, mcm^5s^2U writer	(61, 73)
ENSG00000109911	ELP4	elongator acetyltransferase complex subunit 4	cm^5U, ncm^5U, mcm^5U, mcm^5s^2U writer	(61, 73)
ENSG00000170291	ELP5	elongator acetyltransferase complex subunit 5	cm^5U, ncm^5U, mcm^5U, mcm^5s^2U writer	(61, 73)
ENSG00000126749	EMG1	EMG1 N1-specific pseudouridine methyltransferase	m^1acp3-Psi writer	(74)
ENSG00000105202	FBL	fibrillarin	N_m writer	(75)
ENSG00000102081	FMR1	FMRP translational regulator 1	m^6A reader	(76)
ENSG00000140718	FTO	FTO alpha-ketoglutarate dependent dioxygenase	m^6A eraser	(77)
ENSG00000068438	FTSJ1	FtsJ RNA 2'-O-methyltransferase 1	C_m, U_m, G_m, f^5C_m, hm^5C_m, mcm^5U_m writer	(61, 78)

ENSG000 00108592	FTSJ3	FtsJ RNA 2'-O-methyltransferase 3	C_m, U_m, G_m writer	(79)
ENSG000 00170270	GON7	GON7 subunit of KEOPS complex	t^6A writer	(80)
ENSG000 00130299	GTPBP3	GTP binding protein 3, mitochondrial	tm^5U writer	(81)
ENSG000 00162639	HENMT 1	HEN methyltransferase 1	N_m writer	(82)
ENSG000 00122566	HNRNP A2B1	heterogeneous nuclear ribonucleoprotein A2/B1	m^6A reader	(83)
ENSG000 00092199	HNRNP C	heterogeneous nuclear ribonucleoprotein C	m^6A reader	(84)
ENSG000 00072506	HSD17B 10	hydroxysteroid 17-beta dehydrogenase 10	m^1G, m^1A writer subunit	(85)
ENSG000 00159217	IGF2BP 1	insulin like growth factor 2 mRNA binding protein 1	m^6A reader	(86)
ENSG000 00073792	IGF2BP 2	insulin like growth factor 2 mRNA binding protein 2	m^6A reader	(86)
ENSG000 00136231	IGF2BP 3	insulin like growth factor 2 mRNA binding protein 3	m^6A reader	(86)
ENSG000 00136003	ISCU	iron-sulfur cluster assembly enzyme	s^2U, mcm^5s^2U writer	(61)
ENSG000 00196976	LAGE3	L antigen family member 3	t^6A writer	(87)
ENSG000 00168806	LCMT2	leucine carboxyl methyltransferase 2	o2Yw, yW writer	(61)

ENSG00000138095	LRPPRC	leucine rich pentatricopeptide repeat containing	m^6A reader	(88)
ENSG00000146834	MEPCE	methylphosphate capping enzyme	5′ monophosphate methylation writer	(89)
ENSG00000037897	METTL1	methyltransferase like 1	m^7G writer	(90)
ENSG00000145388	METTL14	methyltransferase like 14	m^6A writer complex	(91)
ENSG00000169519	METTL15	methyltransferase like 15	m^4C writer	(92)
ENSG00000127804	METTL16	methyltransferase like 16	m^6A writer	(93)
ENSG00000165792	METTL17	methyltransferase like 17	m^4C, m^5C writer	(94)
ENSG00000165055	METTL2B	methyltransferase like 2b	m^3C writer	(95)
ENSG00000165819	METTL3	methyltransferase like 3	m^6A writer complex	(91)
ENSG00000138382	METTL5	methyltransferase like 5	m^6A writer	(96)
ENSG00000206562	METTL6	methyltransferase like 6	m^3C writer	(95)
ENSG00000123600	METTL8	methyltransferase like 8	m^3C writer	(95)
ENSG00000197006	METTL9	methyltransferase like 9	1-methylhistidine writer	(97)

ENSG000 00124217	MOCS3	molybdenum cofactor synthesis 3	s^2U, mcm^5s^2U writer	(98)
ENSG000 00128309	MPST	mercaptopyruvate sulfurtransferase	s^2U, mcm^5s^2U writer	(61)
ENSG000 00278619	MRM1	mitochondrial rRNA methyltransferase 1	G_m writer	(99)
ENSG000 00122687	MRM2	mitochondrial rRNA methyltransferase 2	U_m writer	(100)
ENSG000 00171861	MRM3	mitochondrial rRNA methyltransferase 3	G_m writer	(101)
ENSG000 00135297	MTO1	mitochondrial tRNA translation optimization 1	tm^5U writer	(61, 102)
ENSG000 00135372	NAT10	N-acetyltransferase 10	ac^4C writer	(103)
ENSG000 00244005	NFS1	NFS1 cysteine desulfurase	s^2U, mcm^5s^2U writer	(98)
ENSG000 00111641	NOP2	NOP2 nucleolar protein	m^5C writer	(104)
ENSG000 00037474	NSUN2	NOP2/Sun RNA methyltransferase 2	m^5C writer	(105)
ENSG000 00117481	NSUN4	NOP2/Sun RNA methyltransferase 4	m^5C writer	(106)
ENSG000 00130305	NSUN5	NOP2/Sun RNA methyltransferase 5	m^5C writer	(107)
ENSG000 00241058	NSUN6	NOP2/Sun RNA methyltransferase 6	m^5C writer	(108)
ENSG000 00103274	NUBP1	nucleotide binding protein 1	s^2U, mcm^5s^2U writer	(61)

ENSG00000198585	NUDT16	nudix hydrolase 16	m^7GpppN eraser	(109)
ENSG00000092094	OSGEP	O-sialoglycoprotein endopeptidase	t^6A writer	(110)
ENSG00000100982	PCIF1	phosphorylated CTD interacting factor 1	m^6A_m writer	(111)
ENSG00000204469	PRRC2A	proline rich coiled-coil 2A	m^6A reader	(112)
ENSG00000177192	PUS1	pseudouridine synthase 1	Ψ writer	(113)
ENSG00000162927	PUS10	pseudouridine synthase 10	Ψ writer	(114)
ENSG00000110060	PUS3	pseudouridine synthase 3	Ψ writer	(115)
ENSG00000091127	PUS7	pseudouridine synthase 7	Ψ writer	(116)
ENSG00000129317	PUS7L	pseudouridine synthase 7 like	Ψ writer	(117)
ENSG00000213339	QTRT1	queuine tRNA-ribosyltransferase catalytic subunit 1	Q writer	(118)
ENSG00000151576	QTRT2	queuine tRNA-ribosyltransferase accessory subunit 2	Q writer	(119)
ENSG00000162775	RBM15	RNA binding motif protein 15	m^6A writer complex	(120)
ENSG00000259956	RBM15B	RNA binding motif protein 15B	m^6A writer complex	(120)

ENSG00000147274	RBMX	RNA binding motif protein X-linked	m^6A reader	(121)
ENSG00000111880	RNGTT	RNA guanylyltransferase and 5'-phosphatase	m^7GpppN writer	(122)
ENSG00000101654	RNMT	RNA guanine-7 methyltransferase	m^7GpppN writer	(122)
ENSG00000007376	RPUSD1	RNA pseudouridine synthase domain containing 1	ψ (probable) writer	(123)
ENSG00000166133	RPUSD2	RNA pseudouridine synthase domain containing 2	ψ (probable) writer	(123)
ENSG00000156990	RPUSD3	RNA pseudouridine synthase D3	ψ writer	(123)
ENSG00000165526	RPUSD4	RNA pseudouridine synthase D4	ψ writer	(123)
ENSG00000132275	RRP8	ribosomal RNA processing 8	m^1A writer	(124)
ENSG00000129158	SERGEF	secretion regulating guanine nucleotide exchange factor	s^2U, mcm^5s^2U writer	(61)
ENSG00000197157	SND1	staphylococcal nuclease and tudor domain containing 1	m^6A reader	(125)
ENSG00000100138	SNU13	small nuclear ribonucleoprotein 13	methylation writer complex	(126)
ENSG00000059588	TARBP1	TAR (HIV-1) RNA binding protein 1	G_m writer	(127)

ENSG00000029639	TFB1M	transcription factor B1, mitochondrial	$m^6{}_2A$ writer	(128)
ENSG00000162851	TFB2M	transcription factor B2, mitochondrial	$m^6{}_2A$ writer	(128)
ENSG00000137574	TGS1	trimethylguanosine synthase 1	$m^{2,2,7}G$ writer	(129)
ENSG00000113272	THG1L	tRNA-histidine guanylyltransferase 1 like	xG writer	(61)
ENSG00000066654	THUMPD1	THUMP domain containing 1	ac^4C writer unit	(130)
ENSG00000172315	TP53RK	TP53 regulating kinase	t^6A writer	(61, 131)
ENSG00000144034	TPRKB	TP53RK binding protein	t^6A writer	(61, 131)
ENSG00000107614	TRDMT1	tRNA aspartic acid methyltransferase 1	m^5C writer	(132)
ENSG00000043514	TRIT1	tRNA isopentenyltransferase 1	i^6A writer	(133)
ENSG00000104907	TRMT1	tRNA methyltransferase 1	$m^{2,2}G$ writer	(134)
ENSG00000145331	TRMT10A	tRNA methyltransferase 10A	m^1G writer	(135)
ENSG00000174173	TRMT10C	tRNA methyltransferase 10C, mitochondrial RNase P subunit	m^1A, m^1G writer	(61, 136)
ENSG00000066651	TRMT11	tRNA methyltransferase 11 homolog	m^2G writer	(137)

ENSG000 00173113	TRMT1 12	tRNA methyltransferase subunit 11-2	m^7G writer	(138)
ENSG000 00122435	TRMT1 3	tRNA methyltransferase 13 homolog	C_m, A_m writer	(139)
ENSG000 00099899	TRMT2 A	tRNA methyltransferase 2 homolog A	m^5U writer	(140)
ENSG000 00155275	TRMT4 4	tRNA methyltransferase 44 homolog	U_m writer	(61)
ENSG000 00126814	TRMT5	tRNA methyltransferase 5	m^1G, m^1I writer	(61, 141)
ENSG000 00089195	TRMT6	tRNA methyltransferase 6	m^1A writer	(136)
ENSG000 00166166	TRMT6 1A	tRNA methyltransferase 61A	m^1A writer	(136)
ENSG000 00171103	TRMT6 1B	tRNA methyltransferase 61B	m^1A writer	(142)
ENSG000 00100416	TRMU	tRNA mitochondrial 2-thiouridylase	mnm^5s^2U writer	(143)
ENSG000 00165832	TRUB1	TruB pseudouridine synthase family member 1	Ψ writer	(144)
ENSG000 00167112	TRUB2	TruB pseudouridine synthase family member 2	Ψ writer	(145)
ENSG000 00134744	TUT4	terminal uridylyl transferase 4	uridylation writer	(146)
ENSG000 00083223	TUT7	terminal uridylyl transferase 7	uridylation writer	(146)

ENSG00000198874	TYW1	tRNA-yW synthesizing protein 1 homolog	4-demethylwyosine writer	(147)
ENSG00000162623	TYW3	tRNA-yW synthesizing protein 3 homolog	7-aminocarboxypropylwyosine writer	(148)
ENSG00000167118	URM1	ubiquitin related modifier 1	mcm^5s^2U writer	(148)
ENSG00000164944	VIRMA	vir like m6A methyltransferase associated	m^6A writer complex	(149)
ENSG00000160193	WDR4	WD repeat domain 4	m^7G writer	(90)
ENSG00000178252	WDR6	WD repeat domain 6	C_m, G_m, f^5C_m, hm^5C_m writer	(61, 78)
ENSG00000146457	WTAP	WT1 associated protein	m^6A writer complex	(150)
ENSG00000065978	YBX1	Y-box binding protein 1	m^5C reader	(9)
ENSG00000196449	YRDC	yrdC N6-threonylcarbamoyltransferase domain containing	t^6A writer	(87)
ENSG00000083896	YTHDC1	YTH domain containing 1	m^6A reader	(151)
ENSG00000047188	YTHDC2	YTH domain containing 2	m^6A reader	(152)
ENSG00000149658	YTHDF1	YTH N6-methyladenosine RNA binding protein 1	m^6A, m^1A reader	(151)

ENSG000 00198492	YTHDF 2	YTH N6-methyladenosine RNA binding protein 2	m^6A, m^1A, m^5C reader	(8, 151, 153)
ENSG000 00185728	YTHDF 3	YTH N6-methyladenosine RNA binding protein 3	m^6A, m^1A reader	(151, 153)
ENSG000 00123200	ZC3H13	zinc finger CCCH-type containing 13	m^6A writer complex	(154)
ENSG000 00168228	ZCCHC 4	zinc finger CCHC-type containing 4	m^6A writer	(155)

RNA modifications and their abbreviations:

I, inosine; m^1A, 1-methyladenosine; m^5C, 5-methylcytidine; m^6A, *N^6*-methyladenosine; mcm^5U, 5-methoxycarbonylmethyluridine; mcm^5U_m, 5-methoxycarbonylmethyl-2′-*O*-methyluridine; $mchm^5U$, 5-(carboxyhydroxymethyl)uridine methyl ester; mcm^5s^2U, 5-methoxycarbonylmethyl-2-thiouridine; m^7G, 7-methylguanosine; ms^2i^6A, 2-methylthio-*N^6*-isopentenyladenosine; ms^2t^6A, 2-methylthio-*N^6*-threonylcarbamoyladenosine; s^2U, 2-thiouridine; N_m, 2′-*O*-methylation; $m_2^{6,6}A$, *$N^{6,6}$*-dimethyladenosine; Ψ, pseudouridine; D, dihydrouridine; cm^5U, 5-carboxymethyluridine; ncm^5U, 5-carbamoylmethyluridine; m1acp3-Psi, *N^1*-methyl-*N^3*-(3-amino-3-carboxypropyl) pseudouridine; C_m, 2′-*O*-methylcytidine; Um, 2′-*O*-methyluridine; Gm, 2’-*O*-methylguanosine; f^5Cm, 5-formyl-2’-*O*-methylcytidine; hm^5Cm, 2′-O-methyl-5-hydroxymethylcytidine; t^6A, *N^6*-threonylcarbamoyladenosine; tm^5U, 5-taurinomethyluridine; m^1G, 1-methylguanosine; o2Yw, peroxywybutosine; yW, wybutosine; m^4C, *N^4*-methylcytidine; m^3C, 3-methylcytidine; ac^4C, *N^4*-acetylcytidine; m^6Am, *N^6*,2′-*O*-dimethyladenosine; Q, queuosine; $m^6{}_2A$, *N^6*,*N^6*-dimethyladenosine; $m^{2,2,7}G$, *N^2*,*N^2*,7-trimethylguanosine; xG, unknown modified guanosine; $m^{2,2}G$, *N^2*,*N^2*-dimethylguanosine; m^2G, *N^2*-methylguanosine; m^5U, 5-methyluridine; m^1I, 1-methylinosine; mnm^5s^2U, 5-methylaminomethyl-2-thiouridine

References

1. 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;485:201-6.

2. Boccaletto P, Machnicka MA, Purta E, Piatkowski P, Baginski B, Wirecki TK, et al. MODOMICS: a database of RNA modification pathways. 2017 update. Nucleic Acids Res. 2018;46:D303-d7.

3. Wang X, Lu Z, Gomez A, Hon GC, Yue Y, Han D, et al. N6-methyladenosine-dependent regulation of messenger RNA stability. Nature. 2014;505:117-20.

4. Xiao W, Adhikari S, Dahal U, Chen YS, Hao YJ, Sun BF, et al. Nuclear m(6)A Reader YTHDC1 Regulates mRNA Splicing. Mol Cell. 2016;61:507-19.

5. Liu T, Wei Q, Jin J, Luo Q, Liu Y, Yang Y, et al. The m6A reader YTHDF1 promotes ovarian cancer progression via augmenting EIF3C translation. Nucleic Acids Res. 2020;48:3816-31.

6. Shi H, Wang X, Lu Z, Zhao BS, Ma H, Hsu PJ, et al. YTHDF3 facilitates translation and decay of N(6)-methyladenosine-modified RNA. Cell Res. 2017;27:315-28.

7. Yang X, Yang Y, Sun BF, Chen YS, Xu JW, Lai WY, et al. 5-methylcytosine promotes mRNA export - NSUN2 as the methyltransferase and ALYREF as an m(5)C reader. Cell Res. 2017;27:606-25.

8. Dai X, Gonzalez G, Li L, Li J, You C, Miao W, et al. YTHDF2 Binds to 5-Methylcytosine in RNA and Modulates the Maturation of Ribosomal RNA. Anal Chem. 2020;92:1346-54.

9. Chen X, Li A, Sun BF, Yang Y, Han YN, Yuan X, et al. 5-methylcytosine promotes pathogenesis of bladder cancer through stabilizing mRNAs. Nat Cell Biol. 2019;21:978-90.

10. Delaney G, Jacob S, Featherstone C, Barton M. The role of radiotherapy in cancer treatment: estimating optimal utilization from a review of evidence-based clinical guidelines. Cancer. 2005;104:1129-37.

11. Boreham DR, Mitchel REJ. DNA Repair in Chlamydomonas reinhardtii Induced by Heat Shock and Gamma Radiation. Radiat Res. 1993;135:365-71.

12. 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.

13. Xia F, Taghian DG, DeFrank JS, Zeng ZC, Willers H, Iliakis G, et al. Deficiency of human BRCA2 leads to impaired homologous recombination but maintains normal nonhomologous end joining. Proc Natl Acad Sci U S A. 2001;98:8644-9.

14. Zhou H, Kim YS, Peletier A, McCall W, Earp HS, Sartor CI. Effects of the EGFR/HER2 kinase inhibitor GW572016 on EGFR- and HER2-overexpressing breast cancer cell line proliferation, radiosensitization, and resistance. Int J Radiat Oncol Biol Phys. 2004;58:344-52.

15. Liang K, Jin W, Knuefermann C, Schmidt M, Mills GB, Ang KK, et al. Targeting the phosphatidylinositol 3-kinase/Akt pathway for enhancing breast cancer cells to radiotherapy. Mol Cancer Ther. 2003;2:353-60.

16. Visvanathan A, Patil V, Arora A, Hegde AS, Arivazhagan A, Santosh V, et al. Essential role of METTL3-mediated m(6)A modification in glioma stem-like cells maintenance and radioresistance. Oncogene. 2018;37:522-33.

17. Kowalski-Chauvel A, Lacore MG, Arnauduc F, Delmas C, Toulas C, Cohen-Jonathan-Moyal E, et al. The m6A RNA Demethylase ALKBH5 Promotes Radioresistance and Invasion Capability of Glioma Stem Cells. Cancers. 2020;13.

18. He JJ, Li Z, Rong ZX, Gao J, Mu Y, Guan YD, et al. m(6)A Reader YTHDC2 Promotes Radiotherapy Resistance of Nasopharyngeal Carcinoma via Activating IGF1R/AKT/S6 Signaling Axis. Frontiers in oncology. 2020;10:1166.

19. Ronsein GE, Pamir N, von Haller PD, Kim DS, Oda MN, Jarvik GP, et al. Parallel reaction monitoring (PRM) and selected reaction monitoring (SRM) exhibit comparable linearity, dynamic range and precision for targeted quantitative HDL proteomics. Journal of proteomics. 2015;113:388-99.

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. Ahmed KM, Dong S, Fan M, Li JJ. Nuclear Factor-κB p65 Inhibits Mitogen-Activated Protein Kinase Signaling Pathway in Radioresistant Breast Cancer Cells. 2006;4:945-55.

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

23. Wisniewski JR, Zougman A, Nagaraj N, Mann M. Universal sample preparation method for proteome analysis. Nat Meth. 2009;6:359-62.

24. 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 (Oxford, England). 2010;26:966-8.

25. Miao WL, Li L, Wang YS. A Targeted Proteomic Approach for Heat Shock Proteins Reveals DNAJB4 as a Suppressor for Melanoma Metastasis. Anal Chem. 2018;90:6835-42.

26. Kawahara R, Bollinger JG, Rivera C, Ribeiro ACP, Brandão TB, Paes Leme AF, et al. A targeted proteomic strategy for the measurement of oral cancer candidate biomarkers in human saliva. Proteomics. 2016;16:159-73.

27. de Crécy-Lagard V, Boccaletto P, Mangleburg CG, Sharma P, Lowe TM, Leidel SA, et al. Matching tRNA modifications in humans to their known and predicted enzymes. Nucleic Acids Res. 2019;47:2143-59.

28. Mathlin J, Le Pera L, Colombo T. A Census and Categorization Method of Epitranscriptomic Marks. Int J Mol Sci. 2020;21.

29. Kadumuri RV, Janga SC. Epitranscriptomic Code and Its Alterations in Human Disease. Trends Mol Med. 2018;24:886-903.

30. Suzuki T, Suzuki T. A complete landscape of post-transcriptional modifications in mammalian mitochondrial tRNAs. Nucleic Acids Res. 2014;42:7346-57.

31. Boriack-Sjodin PA, Ribich S, Copeland RA. RNA-modifying proteins as anticancer drug targets. Nature reviews Drug discovery. 2018;17:435-53.

32. Begik O, Lucas MC, Liu H, Ramirez JM, Mattick JS, Novoa EM. Integrative analyses of the RNA modification machinery reveal tissue- and cancer-specific signatures. Genome Biol. 2020;21:97.

33. Liu F, Clark W, Luo G, Wang X, Fu Y, Wei J, et al. ALKBH1-Mediated tRNA Demethylation Regulates Translation. Cell. 2016;167:1897.

34. Yarian CS, Basti MM, Cain RJ, Ansari G, Guenther RH, Sochacka E, et al. Structural and functional roles of the N1- and N3-protons of Ψ at tRNA's position 39. Nucleic Acids Res. 1999;27:3543-9.

35. Holm J, Eriksson L, Ploner A, Eriksson M, Rantalainen M, Li J, et al. Assessment of Breast Cancer Risk Factors Reveals Subtype Heterogeneity. Cancer Res. 2017;77:3708-17.

36. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102:15545-50.

37. Zhang P, Sun Y, Ma L. ZEB1: at the crossroads of epithelial-mesenchymal transition, metastasis and therapy resistance. Cell cycle (Georgetown, Tex). 2015;14:481-7.

38. Gray M, Turnbull AK, Ward C, Meehan J, Martínez-Pérez C, Bonello M, et al. Development and characterisation of acquired radioresistant breast cancer cell lines. Radiat Oncol. 2019;14:64.

39. Hara T, Iwadate M, Tachibana K, Waguri S, Takenoshita S, Hamada N. Metastasis of breast cancer cells to the bone, lung, and lymph nodes promotes resistance to ionizing radiation. Strahlenther Onkol. 2017;193:848-55.

40. Che Y, Li Y, Zheng F, Zou K, Li Z, Chen M, et al. TRIP4 promotes tumor growth and metastasis and regulates radiosensitivity of cervical cancer by activating MAPK, PI3K/AKT, and hTERT signaling. Cancer Lett. 2019;452:1-13.

41. You G-R, Chang JT, Li Y-L, Chen Y-J, Huang Y-C, Fan K-H, et al. Molecular Interplays Between Cell Invasion and Radioresistance That Lead to Poor Prognosis in Head-Neck Cancer. Frontiers in oncology. 2021;11:681717-.

42. Xu T, Zeng Y, Shi L, Yang Q, Chen Y, Wu G, et al. Targeting NEK2 impairs oncogenesis and radioresistance via inhibiting the Wnt1/β-catenin signaling pathway in cervical cancer. Journal of experimental & clinical cancer research : CR. 2020;39:183.

43. Hou Y, Liang H, Rao E, Zheng W, Huang X, Deng L, et al. Non-canonical NF-κB Antagonizes STING Sensor-Mediated DNA Sensing in Radiotherapy. Immunity. 2018;49:490-503.e4.

44. Tsao T, Beretov J, Ni J, Bai X, Bucci J, Graham P, et al. Cancer stem cells in prostate cancer radioresistance. Cancer Lett. 2019;465:94-104.

45. Tormey DC, Waalkes TP, Gehrke CW. Biological markers in breast carcinoma--clinical correlations with pseudouridine, N2,N2-dimethylguanosine, and 1-methylinosine. Journal of surgical oncology. 1980;14:267-73.

46. Tormey DC, Waalkes TP, Ahmann D, Gehrke CW, Zumwatt RW, Snyder J, et al. Biological markers in breast carcinoma. I. Incidence of abnormalities of CEA, HCG, three polyamines, and three minor nucleosides. Cancer. 1975;35:1095-100.

47. Boccaletto P, Machnicka MA, Purta E, Piątkowski P, Bagiński B, Wirecki TK, et al. MODOMICS: a database of RNA modification pathways. 2017 update. Nucleic Acids Res. 2017;46:D303-D7.

48. Nishikura K. Functions and regulation of RNA editing by ADAR deaminases. Annu Rev Biochem. 2010;79:321-49.

49. Liu F, Clark W, Luo G, Wang X, Fu Y, Wei J, et al. ALKBH1-Mediated tRNA Demethylation Regulates Translation. Cell. 2016;167:816-28.e16.

50. Chen YS, Yang WL, Zhao YL, Yang YG. Dynamic transcriptomic m(5) C and its regulatory role in RNA processing. Wiley interdisciplinary reviews RNA. 2021;12:e1639.

51. Li X, Xiong X, Wang K, Wang L, Shu X, Ma S, et al. Transcriptome-wide mapping reveals reversible and dynamic N(1)-methyladenosine methylome. Nat Chem Biol. 2016;12:311-6.

52. Zheng G, Dahl JA, Niu Y, Fedorcsak P, Huang CM, Li CJ, et al. ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. Mol Cell. 2013;49:18-29.

53. van den Born E, Vågbø CB, Songe-Møller L, Leihne V, Lien GF, Leszczynska G, et al. ALKBH8-mediated formation of a novel diastereomeric pair of wobble nucleosides in mammalian tRNA. Nature Communications. 2011;2:172.

54. Chen Z, Eggerman TL, Bocharov AV, Baranova IN, Vishnyakova TG, Kurlander RJ, et al. Hypermutation of ApoB mRNA by rat APOBEC-1 overexpression mimics APOBEC-3 hypermutation. J Mol Biol. 2012;418:65-81.

55. Sharma S, Wang J, Alqassim E, Portwood S, Cortes Gomez E, Maguire O, et al. Mitochondrial hypoxic stress induces widespread RNA editing by APOBEC3G in natural killer cells. Genome Biol. 2019;20:37.

56. Xhemalce B, Robson SC, Kouzarides T. Human RNA methyltransferase BCDIN3D regulates microRNA processing. Cell. 2012;151:278-88.

57. Létoquart J, Huvelle E, Wacheul L, Bourgeois G, Zorbas C, Graille M, et al. Structural and functional studies of Bud23-Trm112 reveal 18S rRNA N7-G1575 methylation occurs on late 40S precursor ribosomes. Proc Natl Acad Sci U S A. 2014;111:E5518-26.

58. Zheng F, Du F, Qian H, Zhao J, Wang X, Yue J, et al. Expression and clinical prognostic value of m6A RNA methylation modification in breast cancer. Biomarker research. 2021;9:28.

59. Fakruddin M, Wei FY, Emura S, Matsuda S, Yasukawa T, Kang D, et al. Cdk5rap1-mediated 2-methylthio-N6-isopentenyladenosine modification is absent from nuclear-derived RNA species. Nucleic Acids Res. 2017;45:11954-61.

60. Wei FY, Suzuki T, Watanabe S, Kimura S, Kaitsuka T, Fujimura A, et al. Deficit of tRNA(Lys) modification by Cdkal1 causes the development of type 2 diabetes in mice. The Journal of clinical investigation. 2011;121:3598-608.

61. de Crécy-Lagard V, Boccaletto P, Mangleburg CG, Sharma P, Lowe TM, Leidel SA, et al. Matching tRNA modifications in humans to their known and predicted enzymes. Nucleic Acids Res. 2019;47:2143-59.

62. Smietanski M, Werner M, Purta E, Kaminska KH, Stepinski J, Darzynkiewicz E, et al. Structural analysis of human 2'-O-ribose methyltransferases involved in mRNA cap structure formation. Nat Commun. 2014;5:3004.

63. Dewez M, Bauer F, Dieu M, Raes M, Vandenhaute J, Hermand D. The conserved Wobble uridine tRNA thiolase Ctu1-Ctu2 is required to maintain genome integrity. Proc Natl Acad Sci U S A. 2008;105:5459-64.

64. Wang Z, Jiao X, Carr-Schmid A, Kiledjian M. The hDcp2 protein is a mammalian mRNA decapping enzyme. Proc Natl Acad Sci U S A. 2002;99:12663-8.

65. van Dijk E, Le Hir H, Séraphin B. DcpS can act in the 5'-3' mRNA decay pathway in addition to the 3'-5' pathway. Proc Natl Acad Sci U S A. 2003;100:12081-6.

66. Shen H, Stoute J, Liu KF. Structural and catalytic roles of the human 18S rRNA methyltransferases DIMT1 in ribosome assembly and translation. J Biol Chem. 2020;295:12058-70.

67. Ustianenko D, Hrossova D, Potesil D, Chalupnikova K, Hrazdilova K, Pachernik J, et al. Mammalian DIS3L2 exoribonuclease targets the uridylated precursors of let-7 miRNAs. RNA (New York, NY). 2013;19:1632-8.

68. Heiss NS, Knight SW, Vulliamy TJ, Klauck SM, Wiemann S, Mason PJ, et al. X-linked dyskeratosis congenita is caused by mutations in a highly conserved gene with putative nucleolar functions. Nat Genet. 1998;19:32-8.

69. Xing F, Martzen MR, Phizicky EM. A conserved family of Saccharomyces cerevisiae synthases effects dihydrouridine modification of tRNA. RNA (New York, NY). 2002;8:370-81.

70. Dai W, Li A, Yu NJ, Nguyen T, Leach RW, Wühr M, et al. Activity-based RNA-modifying enzyme probing reveals DUS3L-mediated dihydrouridylation. Nat Chem Biol. 2021;17:1178-87.

71. Meyer KD, Patil DP, Zhou J, Zinoviev A, Skabkin MA, Elemento O, et al. 5' UTR m(6)A Promotes Cap-Independent Translation. Cell. 2015;163:999-1010.

72. Chen Y, Peng C, Chen J, Chen D, Yang B, He B, et al. WTAP facilitates progression of hepatocellular carcinoma via m6A-HuR-dependent epigenetic silencing of ETS1. Molecular cancer. 2019;18:127.

73. Chen C, Huang B, Anderson JT, Byström AS. Unexpected accumulation of ncm(5)U and ncm(5)S(2) (U) in a trm9 mutant suggests an additional step in the synthesis of mcm(5)U and mcm(5)S(2)U. PLoS One. 2011;6:e20783.

74. Wurm JP, Meyer B, Bahr U, Held M, Frolow O, Kötter P, et al. The ribosome assembly factor Nep1 responsible for Bowen-Conradi syndrome is a pseudouridine-N1-specific methyltransferase. Nucleic Acids Res. 2010;38:2387-98.

75. Elliott BA, Ho H-T, Ranganathan SV, Vangaveti S, Ilkayeva O, Abou Assi H, et al. Modification of messenger RNA by 2′-O-methylation regulates gene expression in vivo. Nature Communications. 2019;10:3401.

76. Worpenberg L, Paolantoni C, Longhi S, Mulorz MM, Lence T, Wessels HH, et al. Ythdf is a N6-methyladenosine reader that modulates Fmr1 target mRNA selection and restricts axonal growth in Drosophila. EMBO J. 2021;40:e104975.

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;7:885-7.

78. Li J, Wang YN, Xu BS, Liu YP, Zhou M, Long T, et al. Intellectual disability-associated gene ftsj1 is responsible for 2'-O-methylation of specific tRNAs. EMBO Rep. 2020;21:e50095.

79. Ringeard M, Marchand V, Decroly E, Motorin Y, Bennasser Y. FTSJ3 is an RNA 2'-O-methyltransferase recruited by HIV to avoid innate immune sensing. Nature. 2019;565:500-4.

80. Zhang W, Collinet B, Graille M, Daugeron MC, Lazar N, Libri D, et al. Crystal structures of the Gon7/Pcc1 and Bud32/Cgi121 complexes provide a model for the complete yeast KEOPS complex. Nucleic Acids Res. 2015;43:3358-72.

81. Kopajtich R, Nicholls TJ, Rorbach J, Metodiev MD, Freisinger P, Mandel H, et al. Mutations in GTPBP3 cause a mitochondrial translation defect associated with hypertrophic cardiomyopathy, lactic acidosis, and encephalopathy. Am J Hum Genet. 2014;95:708-20.

82. Lim SL, Qu ZP, Kortschak RD, Lawrence DM, Geoghegan J, Hempfling AL, et al. HENMT1 and piRNA Stability Are Required for Adult Male Germ Cell Transposon Repression and to Define the Spermatogenic Program in the Mouse. PLoS genetics. 2015;11:e1005620.

83. Alarcón CR, Goodarzi H, Lee H, Liu X, Tavazoie S, Tavazoie SF. HNRNPA2B1 Is a Mediator of m(6)A-Dependent Nuclear RNA Processing Events. Cell. 2015;162:1299-308.

84. Liu N, Dai Q, Zheng G, He C, Parisien M, Pan T. N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. Nature. 2015;518:560-4.

85. Oerum S, Roovers M, Rambo RP, Kopec J, Bailey HJ, Fitzpatrick F, et al. Structural insight into the human mitochondrial tRNA purine N1-methyltransferase and ribonuclease P complexes. J Biol Chem. 2018;293:12862-76.

86. Huang H, Weng H, Sun W, Qin X, Shi H, Wu H, et al. Recognition of RNA N(6)-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. Nat Cell Biol. 2018;20:285-95.

87. Arrondel C, Missoury S, Snoek R, Patat J, Menara G, Collinet B, et al. Defects in t(6)A tRNA modification due to GON7 and YRDC mutations lead to Galloway-Mowat syndrome. Nat Commun. 2019;10:3967.

88. Arguello AE, DeLiberto AN, Kleiner RE. RNA Chemical Proteomics Reveals the N(6)-Methyladenosine (m(6)A)-Regulated Protein-RNA Interactome. J Am Chem Soc. 2017;139:17249-52.

89. Yang Y, Eichhorn CD, Wang Y, Cascio D, Feigon J. Structural basis of 7SK RNA 5′-γ-phosphate methylation and retention by MePCE. Nat Chem Biol. 2019;15:132-40.

90. Lin S, Liu Q, Lelyveld VS, Choe J, Szostak JW, Gregory RI. Mettl1/Wdr4-Mediated m(7)G tRNA Methylome Is Required for Normal mRNA Translation and Embryonic Stem Cell Self-Renewal and Differentiation. Mol Cell. 2018;71:244-55.e5.

91. Liu J, Yue Y, Han D, Wang X, Fu Y, Zhang L, et al. A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. Nat Chem Biol. 2014;10:93-5.

92. Laptev I, Shvetsova E, Levitskii S, Serebryakova M, Rubtsova M, Zgoda V, et al. METTL15 interacts with the assembly intermediate of murine mitochondrial small ribosomal subunit to form m4C840 12S rRNA residue. Nucleic Acids Res. 2020;48:8022-34.

93. Pendleton KE, Chen B, Liu K, Hunter OV, Xie Y, Tu BP, et al. The U6 snRNA m(6)A Methyltransferase METTL16 Regulates SAM Synthetase Intron Retention. Cell. 2017;169:824-35.e14.

94. Shi Z, Xu S, Xing S, Yao K, Zhang L, Xue L, et al. Mettl17, a regulator of mitochondrial ribosomal RNA modifications, is required for the translation of mitochondrial coding genes. FASEB J. 2019;33:13040-50.

95. Xu L, Liu X, Sheng N, Oo KS, Liang J, Chionh YH, et al. Three distinct 3-methylcytidine (m(3)C) methyltransferases modify tRNA and mRNA in mice and humans. J Biol Chem. 2017;292:14695-703.

96. van Tran N, Ernst FGM, Hawley BR, Zorbas C, Ulryck N, Hackert P, et al. The human 18S rRNA m6A methyltransferase METTL5 is stabilized by TRMT112. Nucleic Acids Res. 2019;47:7719-33.

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

98. Neukranz Y, Kotter A, Beilschmidt L, Marelja Z, Helm M, Gräf R, et al. Analysis of the Cellular Roles of MOCS3 Identifies a MOCS3-Independent Localization of NFS1 at the Tips of the Centrosome. Biochemistry. 2019;58:1786-98.

99. Lee K-W, Bogenhagen DF. Assignment of 2'-O-methyltransferases to modification sites on the mammalian mitochondrial large subunit 16 S ribosomal RNA (rRNA). The Journal of biological chemistry. 2014;289:24936-42.

100. Pintard L, Bujnicki JM, Lapeyre B, Bonnerot C. MRM2 encodes a novel yeast mitochondrial 21S rRNA methyltransferase. EMBO J. 2002;21:1139-47.

101. Rorbach J, Boesch P, Gammage PA, Nicholls TJJ, Pearce SF, Patel D, et al. MRM2 and MRM3 are involved in biogenesis of the large subunit of the mitochondrial ribosome. Mol Biol Cell. 2014;25:2542-55.

102. Kirino Y, Yasukawa T, Marjavaara SK, Jacobs HT, Holt IJ, Watanabe K, et al. Acquisition of the wobble modification in mitochondrial tRNALeu(CUN) bearing the G12300A mutation suppresses the MELAS molecular defect. Hum Mol Genet. 2006;15:897-904.

103. Ito S, Horikawa S, Suzuki T, Kawauchi H, Tanaka Y, Suzuki T, et al. Human NAT10 is an ATP-dependent RNA acetyltransferase responsible for N4-acetylcytidine formation in 18 S ribosomal RNA (rRNA). J Biol Chem. 2014;289:35724-30.

104. King MY, Redman KL. RNA methyltransferases utilize two cysteine residues in the formation of 5-methylcytosine. Biochemistry. 2002;41:11218-25.

105. Squires JE, Patel HR, Nousch M, Sibbritt T, Humphreys DT, Parker BJ, et al. Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. Nucleic Acids Res. 2012;40:5023-33.

106. Metodiev MD, Spahr H, Loguercio Polosa P, Meharg C, Becker C, Altmueller J, et al. NSUN4 is a dual function mitochondrial protein required for both methylation of 12S rRNA and coordination of mitoribosomal assembly. PLoS genetics. 2014;10:e1004110.

107. Schosserer M, Minois N, Angerer TB, Amring M, Dellago H, Harreither E, et al. Methylation of ribosomal RNA by NSUN5 is a conserved mechanism modulating organismal lifespan. Nat Commun. 2015;6:6158.

108. Haag S, Warda AS, Kretschmer J, Gunnigmann MA, Hobartner C, Bohnsack MT. NSUN6 is a human RNA methyltransferase that catalyzes formation of m5C72 in specific tRNAs. RNA (New York, NY). 2015;21:1532-43.

109. Anadón C, van Tetering G, Ferreira HJ, Moutinho C, Martínez-Cardús A, Villanueva A, et al. Epigenetic loss of the RNA decapping enzyme NUDT16 mediates C-MYC activation in T-cell acute lymphoblastic leukemia. Leukemia. 2017;31:1622-5.

110. Edvardson S, Prunetti L, Arraf A, Haas D, Bacusmo JM, Hu JF, et al. tRNA N6-adenosine threonylcarbamoyltransferase defect due to KAE1/TCS3 (OSGEP) mutation manifest by neurodegeneration and renal tubulopathy. Eur J Hum Genet. 2017;25:545-51.

111. Sun H, Zhang M, Li K, Bai D, Yi C. Cap-specific, terminal N(6)-methylation by a mammalian m(6)Am methyltransferase. Cell Res. 2019;29:80-2.

112. Wu R, Li A, Sun B, Sun JG, Zhang J, Zhang T, et al. A novel m(6)A reader Prrc2a controls oligodendroglial specification and myelination. Cell Res. 2019;29:23-41.

113. Carlile TM, Rojas-Duran MF, Zinshteyn B, Shin H, Bartoli KM, Gilbert WV. Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. Nature. 2014;515:143-6.

114. Roovers M, Hale C, Tricot C, Terns MP, Terns RM, Grosjean H, et al. Formation of the conserved pseudouridine at position 55 in archaeal tRNA. Nucleic Acids Res. 2006;34:4293-301.

115. Lecointe F, Simos G, Sauer A, Hurt EC, Motorin Y, Grosjean H. Characterization of yeast protein Deg1 as pseudouridine synthase (Pus3) catalyzing the formation of psi 38 and psi 39 in tRNA anticodon loop. J Biol Chem. 1998;273:1316-23.

116. Guzzi N, Ciesla M, Ngoc PCT, Lang S, Arora S, Dimitriou M, et al. Pseudouridylation of tRNA-Derived Fragments Steers Translational Control in Stem Cells. Cell. 2018;173:1204-16.e26.

117. Spenkuch F, Motorin Y, Helm M. Pseudouridine: still mysterious, but never a fake (uridine)! RNA biology. 2014;11:1540-54.

118. Zhang J, Lu R, Zhang Y, Matuszek Ż, Zhang W, Xia Y, et al. tRNA Queuosine Modification Enzyme Modulates the Growth and Microbiome Recruitment to Breast Tumors. Cancers. 2020;12.

119. Zhang W, Xu R, Matuszek Ż, Cai Z, Pan T. Detection and quantification of glycosylated queuosine modified tRNAs by acid denaturing and APB gels. RNA (New York, NY). 2020;26:1291-8.

120. Patil DP, Chen CK, Pickering BF, Chow A, Jackson C, Guttman M, et al. m(6)A RNA methylation promotes XIST-mediated transcriptional repression. Nature. 2016;537:369-73.

121. Wu H, Dong H, Fu Y, Tang Y, Dai M, Chen Y, et al. Expressions of m6A RNA methylation regulators and their clinical predictive value in cervical squamous cell carcinoma and endometrial adenocarcinoma. Clin Exp Pharmacol Physiol. 2021;48:270-8.

122. Pillutla RC, Shimamoto A, Furuichi Y, Shatkin AJ. Human mRNA capping enzyme (RNGTT) and cap methyltransferase (RNMT) map to 6q16 and 18p11.22-p11.23, respectively. Genomics. 1998;54:351-3.

123. Uddin MB, Wang Z, Yang C. Dysregulations of Functional RNA Modifications in Cancer, Cancer Stemness and Cancer Therapeutics. Theranostics. 2020;10:3164-89.

124. Waku T, Nakajima Y, Yokoyama W, Nomura N, Kako K, Kobayashi A, et al. NML-mediated rRNA base methylation links ribosomal subunit formation to cell proliferation in a p53-dependent manner. J Cell Sci. 2016;129:2382-93.

125. Baquero-Perez B, Antanaviciute A, Yonchev ID, Carr IM, Wilson SA, Whitehouse A. The Tudor SND1 protein is an m(6)A RNA reader essential for replication of Kaposi's sarcoma-associated herpesvirus. Elife. 2019;8.

126. Ussowicz M, Marcel V, Long FNV, Kazanowska B, Diaz JJ, Wołowiec D. Analysis of the rRNA methylation complex components in pediatric B-cell precursor acute lymphoblastic leukemia: A pilot study. Advances in clinical and experimental medicine : official organ Wroclaw Medical University. 2020;29:107-13.

127. Freund I, Buhl DK, Boutin S, Kotter A, Pichot F, Marchand V, et al. 2'-O-methylation within prokaryotic and eukaryotic tRNA inhibits innate immune activation by endosomal Toll-like receptors but does not affect recognition of whole organisms. RNA (New York, NY). 2019;25:869-80.

128. Liu X, Shen S, Wu P, Li F, Liu X, Wang C, et al. Structural insights into dimethylation of 12S rRNA by TFB1M: indispensable role in translation of mitochondrial genes and mitochondrial function. Nucleic Acids Res. 2019;47:7648-65.

129. Mouaikel J, Bujnicki JM, Tazi J, Bordonné R. Sequence-structure-function relationships of Tgs1, the yeast snRNA/snoRNA cap hypermethylase. Nucleic Acids Res. 2003;31:4899-909.

130. Li K, Liu J, Yang X, Tu Z, Huang K, Zhu X. Pan-cancer analysis of N4-acetylcytidine adaptor THUMPD1 as a predictor for prognosis and immunotherapy. Biosci Rep. 2021.

131. Goswami MT, VanDenBerg KR, Han S, Wang LL, Singh B, Weiss T, et al. Identification of TP53RK-Binding Protein (TPRKB) Dependency in TP53-Deficient Cancers. Molecular cancer research : MCR. 2019;17:1652-64.

132. Goll MG, Kirpekar F, Maggert KA, Yoder JA, Hsieh CL, Zhang X, et al. Methylation of tRNAAsp by the DNA methyltransferase homolog Dnmt2. Science. 2006;311:395-8.

133. Spinola M, Galvan A, Pignatiello C, Conti B, Pastorino U, Nicander B, et al. Identification and functional characterization of the candidate tumor suppressor gene TRIT1 in human lung cancer. Oncogene. 2005;24:5502-9.

134. Dewe JM, Fuller BL, Lentini JM, Kellner SM, Fu D. TRMT1-Catalyzed tRNA Modifications Are Required for Redox Homeostasis To Ensure Proper Cellular Proliferation and Oxidative Stress Survival. Mol Cell Biol. 2017;37.

135. Ontiveros RJ, Shen H, Stoute J, Yanas A, Cui Y, Zhang Y, et al. Coordination of mRNA and tRNA methylations by TRMT10A. Proc Natl Acad Sci U S A. 2020;117:7782-91.

136. Safra M, Sas-Chen A, Nir R, Winkler R, Nachshon A, Bar-Yaacov D, et al. The m1A landscape on cytosolic and mitochondrial mRNA at single-base resolution. Nature. 2017;551:251-5.

137. Esteve-Puig R, Bueno-Costa A, Esteller M. Writers, readers and erasers of RNA modifications in cancer. Cancer Lett. 2020;474:127-37.

138. Zorbas C, Nicolas E, Wacheul L, Huvelle E, Heurgué-Hamard V, Lafontaine DLJ. The human 18S rRNA base methyltransferases DIMT1L and WBSCR22-TRMT112 but not rRNA modification are required for ribosome biogenesis. Mol Biol Cell. 2015;26:2080-95.

139. Begik O, Lucas MC, Liu H, Ramirez JM, Mattick JS, Novoa EM. Integrative analyses of the RNA modification machinery reveal tissue- and cancer-specific signatures. Genome Biol. 2020;21:97.

140. Carter JM, Emmett W, Mozos IR, Kotter A, Helm M, Ule J, et al. FICC-Seq: a method for enzyme-specified profiling of methyl-5-uridine in cellular RNA. Nucleic Acids Res. 2019;47:e113.

141. Powell CA, Kopajtich R, D'Souza AR, Rorbach J, Kremer LS, Husain RA, et al. TRMT5 Mutations Cause a Defect in Post-transcriptional Modification of Mitochondrial tRNA Associated with Multiple Respiratory-Chain Deficiencies. Am J Hum Genet. 2015;97:319-28.

142. Chujo T, Suzuki T. Trmt61B is a methyltransferase responsible for 1-methyladenosine at position 58 of human mitochondrial tRNAs. RNA (New York, NY). 2012;18:2269-76.

143. Yan Q, Li X, Faye G, Guan MX. Mutations in MTO2 related to tRNA modification impair mitochondrial gene expression and protein synthesis in the presence of a paromomycin resistance mutation in mitochondrial 15 S rRNA. J Biol Chem. 2005;280:29151-7.

144. Becker HF, Motorin Y, Planta RJ, Grosjean H. The yeast gene YNL292w encodes a pseudouridine synthase (Pus4) catalyzing the formation of psi55 in both mitochondrial and cytoplasmic tRNAs. Nucleic Acids Res. 1997;25:4493-9.

145. Zucchini C, Strippoli P, Biolchi A, Solmi R, Lenzi L, D'Addabbo P, et al. The human TruB family of pseudouridine synthase genes, including the Dyskeratosis Congenita 1 gene and the novel member TRUB1. Int J Mol Med. 2003;11:697-704.

146. Lim J, Ha M, Chang H, Kwon SC, Simanshu DK, Patel DJ, et al. Uridylation by TUT4 and TUT7 marks mRNA for degradation. Cell. 2014;159:1365-76.

147. Grell TAJ, Young AP, Drennan CL, Bandarian V. Biochemical and Structural Characterization of a Schiff Base in the Radical-Mediated Biosynthesis of 4-Demethylwyosine by TYW1. J Am Chem Soc. 2018;140:6842-52.

148. Currie MA, Brown G, Wong A, Ohira T, Sugiyama K, Suzuki T, et al. Structural and functional characterization of the TYW3/Taw3 class of SAM-dependent methyltransferases. RNA (New York, NY). 2017;23:346-54.

149. Yue Y, Liu J, Cui X, Cao J, Luo G, Zhang Z, et al. VIRMA mediates preferential m(6)A mRNA methylation in 3'UTR and near stop codon and associates with alternative polyadenylation. Cell discovery. 2018;4:10.

150. Ping XL, Sun BF, Wang L, Xiao W, Yang X, Wang WJ, et al. Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. Cell Res. 2014;24:177-89.

151. Wang X, Zhao BS, Roundtree IA, Lu Z, Han D, Ma H, et al. N(6)-methyladenosine Modulates Messenger RNA Translation Efficiency. Cell. 2015;161:1388-99.

152. Hsu PJ, Zhu YF, Ma HH, Guo YH, Shi XD, Liu YY, et al. Ythdc2 is an N-6-methyladenosine binding protein that regulates mammalian spermatogenesis. Cell Res. 2017;27:1115-27.

153. Dai XX, Wang TL, Gonzalez G, Wang YS. Identification of YTH Domain-Containing Proteins as the Readers for N1-Methyladenosine in RNA. Anal Chem. 2018;90:6380-4.

154. Wen J, Lv R, Ma H, Shen H, He C, Wang J, et al. Zc3h13 Regulates Nuclear RNA m(6)A Methylation and Mouse Embryonic Stem Cell Self-Renewal. Mol Cell. 2018;69:1028-38.e6.

155. Ma H, Wang X, Cai J, Dai Q, Natchiar SK, Lv R, et al. N(6-)Methyladenosine methyltransferase ZCCHC4 mediates ribosomal RNA methylation. Nat Chem Biol. 2019;15:88-94.

Figure 3.1 LC-PRM method for uncovering alterations in expression of epitranscriptomic RWE proteins associated with the development of radioresistance.

(a) A SILAC-based LC-PRM workflow. The parental cells (i.e., MDA-MB-231 and MCF-7) and their radioresistant counterparts (i.e., C5 and C6) were labeled in light- or heavy-amino acid-containing media for over six cell doubling times. In the forward SILAC labelling experiments, light-isotope-labelled C5 and C6 cell lysates were mixed at 1:1 ratio (by mass) with heavy-isotope-labelled MDA-MB-231 and MCF-7 cell lysates, respectively. In the reverse SILAC labelling experiments, light-isotope-labelled MDA-MB-231 and MCF-7 cell lysates were mixed at 1:1 ratio (by mass) with heavy-isotope-labelled C5 and C6 cell lysates, respectively. The mixed cell lysate was tryptic digested and subjected to LC-PRM analysis. Data were processed using Skyline. (b-c) Venn diagrams showing the number and percentage of human epitranscriptomic RWE proteins deposited in the Modomics database (purple) compared with those included in the PRM library of this study (yellow) (b), and illustrating the number and percentage of quantified epitranscriptomic RWE proteins in MDA-MB-231/C5 and MCF-7/C6 pairs of breast cancer cells from LC-PRM analyses, compared with those deposited in the PRM library (c). Blue and pink circles in (b) and (c) designate the numbers of quantified epitranscriptomic RWE proteins in MDA-MB-231/C5 and MCF-7/C6 pairs of breast cancer cells, respectively.

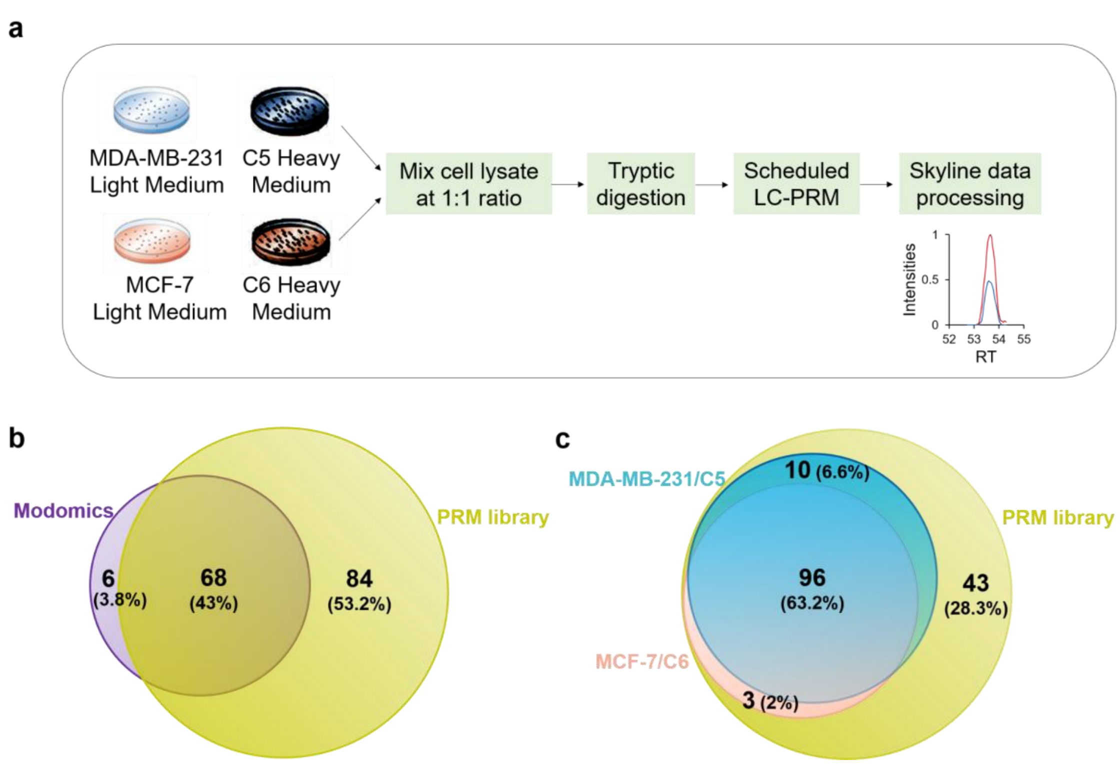


Figure 3.2 Hierarchical clustering displaying the Log_2 transformed expression fold differences of epitranscriptomic RWE proteins in C5/MDA-MB-231 and C6/MCF-7 cells.

The expression fold differences were averaged from two forward and two reverse SILAC experiments. Hierarchical clustering was generated using Perseus, where red and blue boxes designate proteins up- and down-regulated in radioresistant breast cancer cells compared with the corresponding parental lines, respectively; gray boxes represent missed data. Genes were clustered using Euclidean distance.

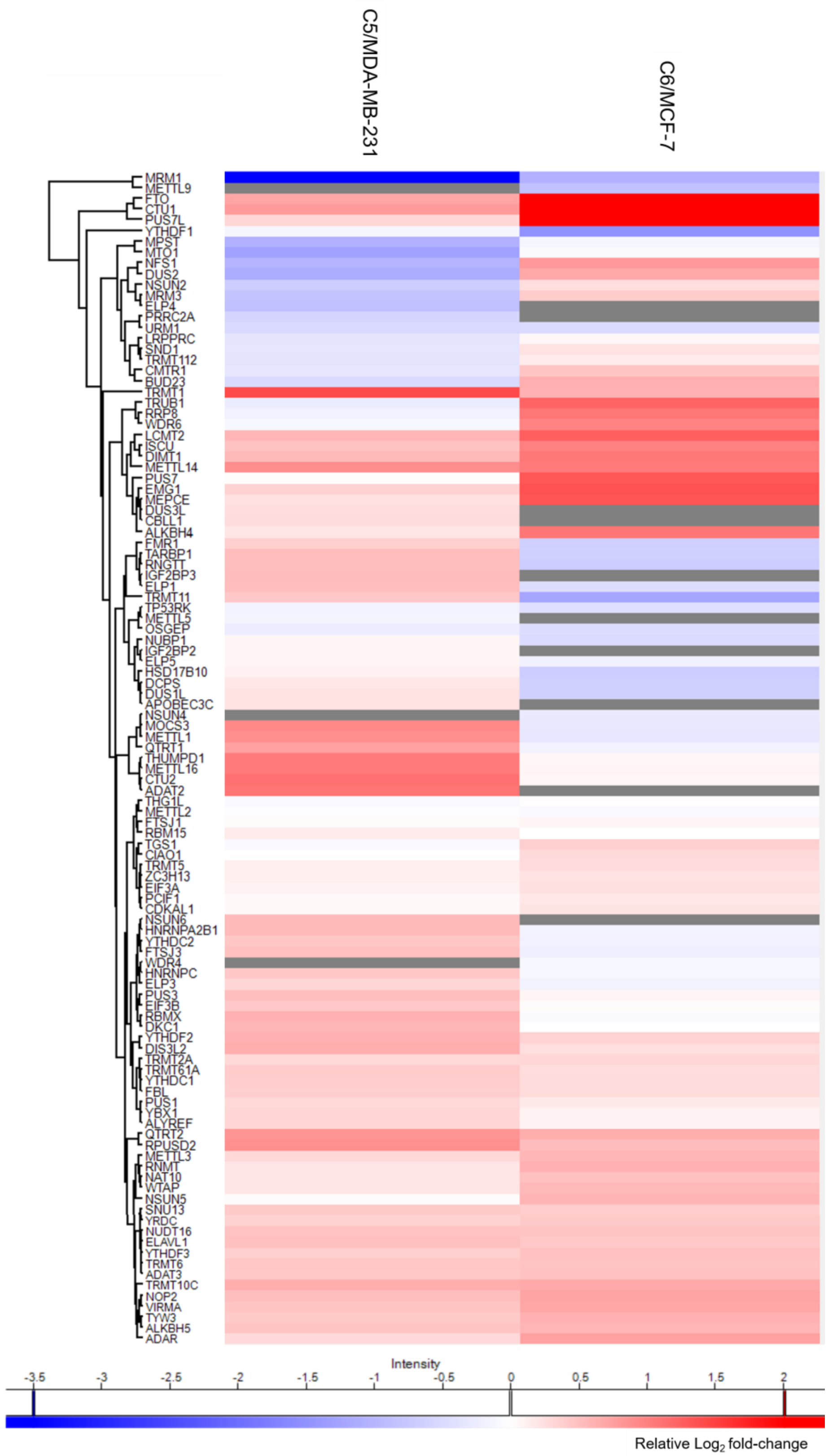

C5/MDA-MB-231
C6/MCF-7
MRM1
METTL9
FTO
CTU1
PUS7L
YTHDF1
MPST
MTO1
NFS1
DUS2
NSUN2
MRM3
ELP4
PRRC2A
URM1
LRPPRC
SND1
TRMT112
CMTR1
BUD23
TRMT1
TRUB1
RRP8
WDR6
LCMT2
ISCU
DIMT1
METTL14
PUS7
EMG1
MEPCE
DUS3L
CBLL1
ALKBH4
FMR1
TARBP1
RNGTT
IGF2BP3
ELP1
TRMT11
TP53RK
METTL5
OSGEP
NUBP1
IGF2BP2
ELP5
HSD17B10
DCPS
DUS1L
APOBEC3C
NSUN4
MOCS3
METTL1
QTRT1
THUMPD1
METTL16
CTU2
ADAT2
THG1L
METTL2
FTSJ1
RBM15
TGS1
CIAO1
TRMT5
ZC3H13
EIF3A
PCIF1
CDKAL1
NSUN6
HNRNPA2B1
YTHDC2
FTSJ3
WDR4
HNRNPC
ELP3
PUS3
EIF3B
RBMX
DKC1
YTHDF2
DIS3L2
TRMT2A
TRMT61A
YTHDC1
FBL
PUS1
YBX1
ALYREF
QTRT2
RPUSD2
METTL3
RNMT
NAT10
WTAP
NSUN5
SNU13
YRDC
NUDT16
ELAVL1
YTHDF3
TRMT6
ADAT3
TRMT10C
NOP2
VIRMA
TYW3
ALKBH5
ADAR
Intensity
-3.5 -3 -2.5 -2 -1.5 -1 -0.5 0 0.5 1 1.5 2
Relative Log2 fold-change

Figure 3.3 Bar graphs showing epitranscriptomic RWE proteins that altered over 1.5-fold in radioresistant cells relative to the corresponding parental cells.

(a-b) Bar graphs depicting the LC-PRM results for those epitranscriptomic RWE proteins with expression differences of over 1.5-fold or less than 0.67-fold in radioresistant cells relative to the corresponding parental cells. (c) Bar graphs illustrating epitranscriptomic RWE proteins that were commonly altered by over 1.5-fold in the two pairs of matched breast cancer cells. (d) A scatter plot displaying $\log_{10}$ transformed expression ratios of the quantified epitranscriptomic RWE proteins in the two pairs of matched breast cancer cells. Eight commonly altered RWE proteins from both pairs by over 1.5-fold were labeled in red dots. The data in (a-c) display the mean and standard deviation of the quantified ratios of different peptides representing a specific epitranscriptomic RWE protein, where the ratio of each peptide was averaged from the quantification results of two forward and two reverse SILAC experiments. Error bars were displayed for those epitranscriptomic RWE proteins with more than one peptide being quantified.

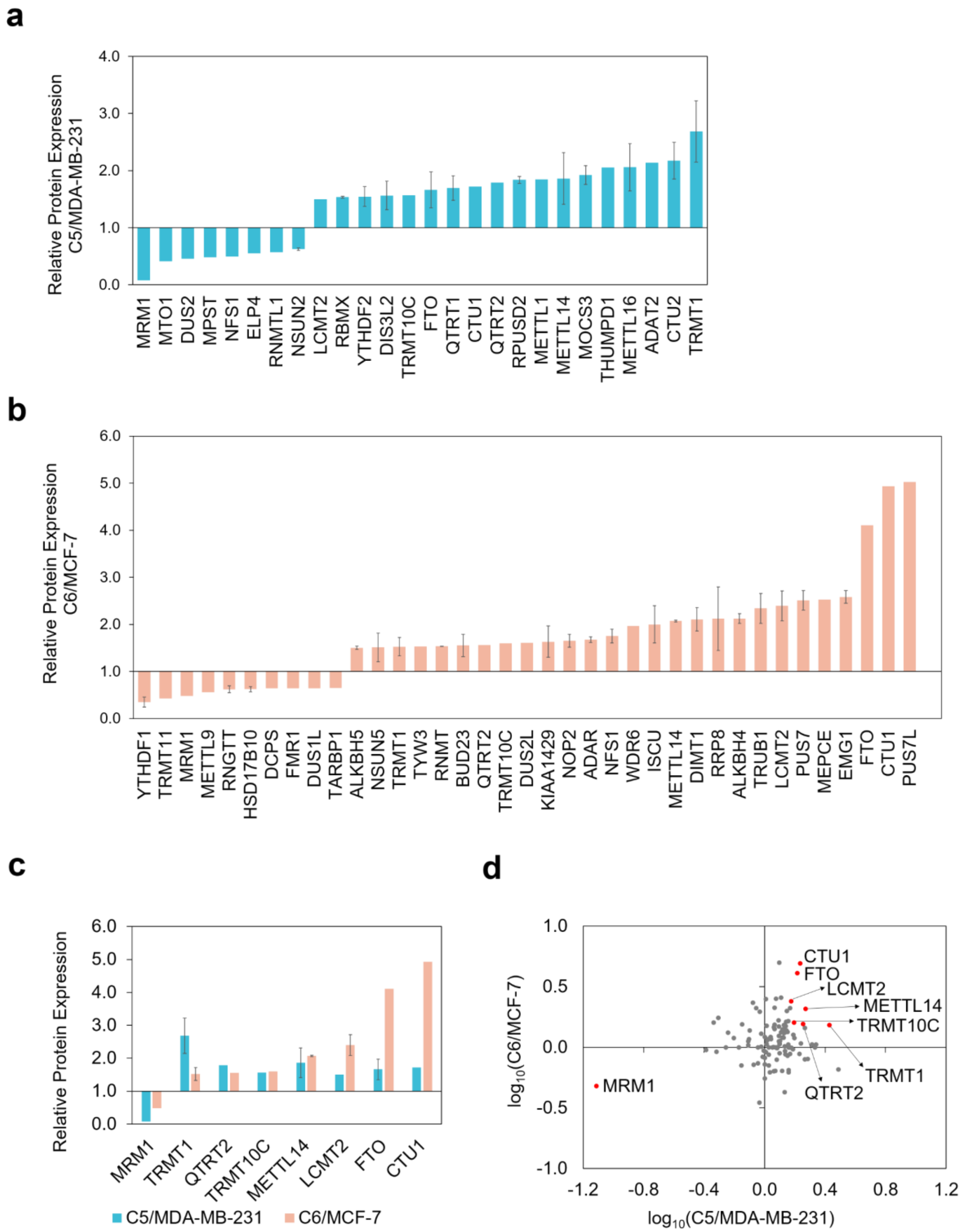

a
Relative Protein Expression C5/MDA-MB-231
MRM1 MTO1 DUS2 MPST NFS1 ELP4 RNMTL1 NSUN2 LCMT2 RBMX YTHDF2 DIS3L2 TRMT10C FTO QTRT1 CTU1 QTRT2 RPUSD2 METTL1 METTL14 MOCS3 THUMPD1 METTL16 ADAT2 CTU2 TRMT1
b
Relative Protein Expression C6/MCF-7
YTHDF1 TRMT11 MRM1 METTL9 RNGTT HSD17B10 DCPS FMR1 DUS1L TARBP1 ALKBH5 NSUN5 TRMT1 TYW3 RNMT BUD23 QTRT2 TRMT10C DUS2L KIAA1429 NOP2 ADAR NFS1 WDR6 ISCU METTL14 DIMT1 RRP8 ALKBH4 TRUB1 LCMT2 PUS7 MEPCE EMG1 FTO CTU1 PUS7L
c
Relative Protein Expression
MRM1 TRMT1 QTRT2 TRMT10C METTL14 LCMT2 FTO CTU1
C5/MDA-MB-231
C6/MCF-7
d
log10(C6/MCF-7)
log10(C5/MDA-MB-231)
CTU1
FTO
LCMT2
METTL14
TRMT10C
TRMT1
QTRT2
MRM1

Figure 3.4 Gene Ontology (GO) analysis of epitranscriptomic RWE proteins differentially expressed by at least 1.5-fold. Clinical relevance of CTU1 and TRMT1.

(a) GO Biological Pathway (BP) analysis of those epitranscriptomic RWE proteins differentially expressed by at least 1.5-fold in either of the two pairs of matched parental/radioresistant breast cancer cell lines. The analysis was carried out using Database for Annotation, Visualization and Integrated Discovery (DAVID). (b) Kaplan-Meier survival analysis of METABRIC cohort who received radiation therapy. Patients were stratified by mRNA expression level of *CTU1* with median value as a cutoff. (c-d) GSEA enrichment plots generated from GSEA 4.1.0 software showing significant enrichment of TRMT1 with Myc_targets_V1 (c) and UV-response_up (d).

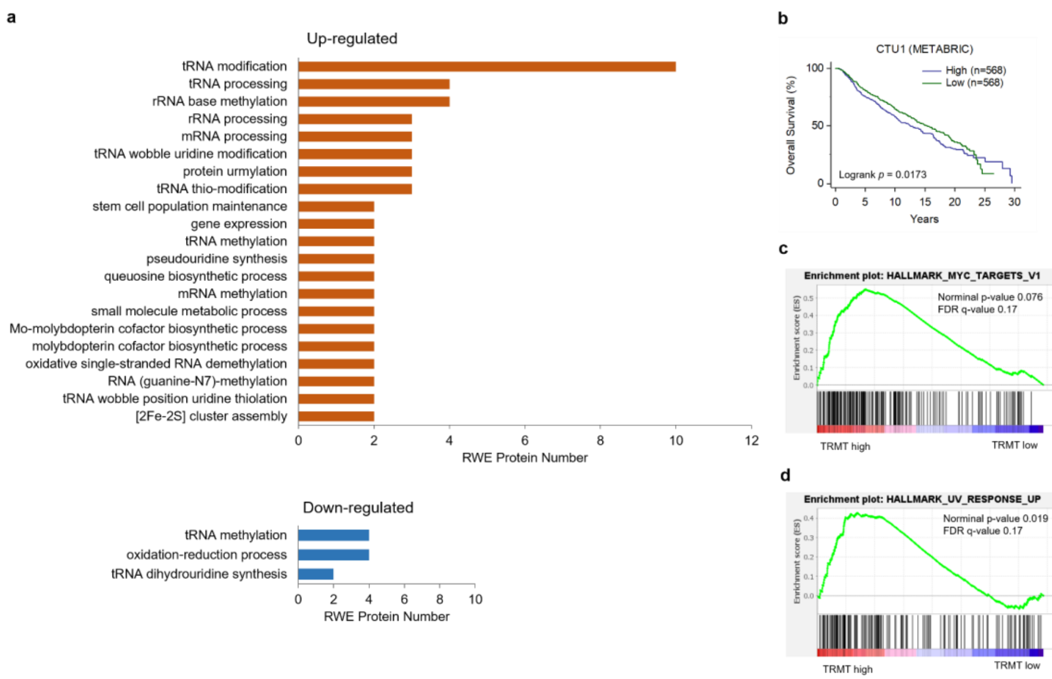
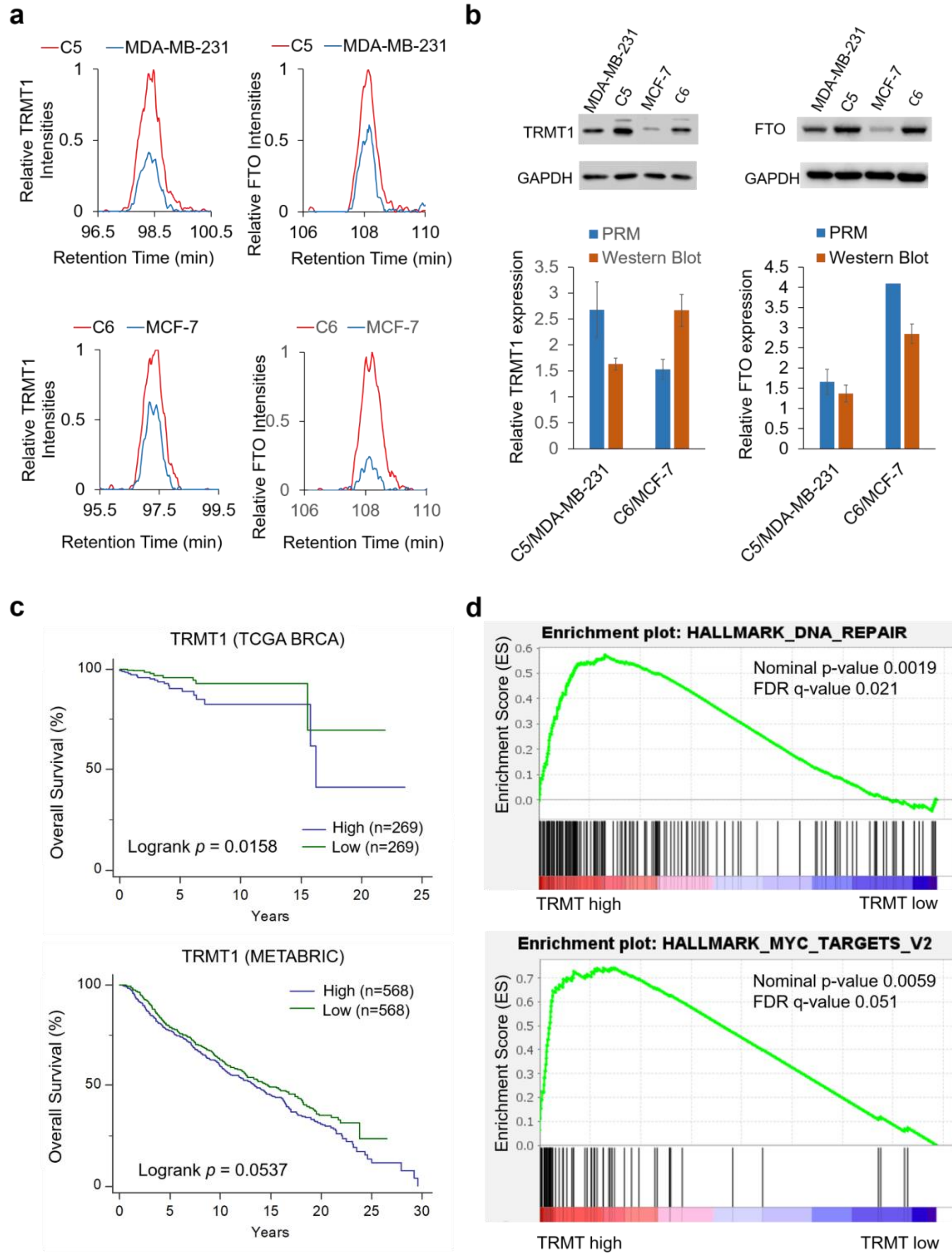


Figure 3.5 Western blot analyses for validating the protein expression levels of TRMT1 and FTO. Clinical relevance of TRMT1 and its correlation with DNA repair gene sets.

(a) PRM traces of representative peptides, FALEVPGLR from TRMT1 and FTVPWPVK from FTO, in C5/MDA-MB-231 and C6/MCF-7 pairs of breast cancer cells. (b) Western blots of TRMT1 and FTO proteins in MDA-MB-231 and MCF-7 pairs of radioresistant/parental breast cancer cells. Relative quantification results of TRMT1 and FTO obtained from PRM and Western blot analysis were shown. The PRM results represent the mean and standard deviation of quantification results of different peptides from a given epitranscriptomic RWE protein, where the ratio of each peptide in the radioresistant over parental cells was averaged from the quantification data of two forward and two reverse SILAC experiments. Western blot data represent the mean and standard deviation of results obtained from three separate experiments. (c) Kaplan-Meier survival analysis of *TRMT1* gene in the TCGA-BRCA and METABRIC cohort of patients who received radiation therapy. Breast cancer patients were stratified by the mRNA expression of *TRMT1* using its median value as a cutoff. The survival plots and log-rank *p*-values were generated and calculated by using MedCalc software. (d) GSEA enrichment plots were generated using GSEA 4.1.0, where the number of permutations was set at 1000.



a
C5 MDA-MB-231
Relative TRMT1 Intensities
Retention Time (min)
Relative FTO Intensities
C6 MCF-7
b
MDA-MB-231 C5 MCF-7 C6
TRMT1
GAPDH
FTO
PRM
Western Blot
Relative TRMT1 expression
Relative FTO expression
C5/MDA-MB-231
C6/MCF-7
c
TRMT1 (TCGA BRCA)
Overall Survival (%)
High (n=269)
Low (n=269)
Logrank p = 0.0158
Years
TRMT1 (METABRIC)
High (n=568)
Low (n=568)
Logrank p = 0.0537
d
Enrichment plot: HALLMARK_DNA_REPAIR
Enrichment Score (ES)
Nominal p-value 0.0019
FDR q-value 0.021
TRMT high
TRMT low
Enrichment plot: HALLMARK_MYC_TARGETS_V2
Nominal p-value 0.0059
FDR q-value 0.051

4. Chapter 4. Parallel-reaction Monitoring Revealed the Roles of Epitranscriptomic Reader, Writer and Eraser Proteins in Colorectal Cancer Metastasis

4.1 Introduction

Epigenetic modifications of DNA and histones are well studied; however, much less is investigated about the roles of RNA modifications in cellular processes. Over 170 types of chemical modifications exist in RNA (1); most of these modifications are found in tRNAs, which contain an average of 13 modifications per molecule (2). The most abundant internal modification in eukaryotic mRNA, N^6-methyladenosine (m^6A) (3), has drawn substantial attention after high-throughput sequencing revealed its widespread occurrence in the transcriptome (4). Cellular proteins have been uncovered for the installation ("writers") (5-9), recognition ("readers") (10-12), and removal ("erasers") (13-15) of m^6A in mRNA. Aside from m^6A, mRNA also contains *N*1-methyladenosine (m^1A) (16-18), 5-methylcytidine (m^5C) (19-21), *N*7-methylguanosine (m^7G) (22), pseudouridine (ψ) (23, 24), and 2′-*O*-methylated nucleosides (25, 26).

Recent studies showed that genetic depletions of some of the epitranscriptomic reader, writer and eraser (RWE) proteins confer embryonic lethality and/or other developmental abnormalities, and their mutations and/or aberrant expressions result in the initiation and progression of cancer, impaired anti-viral response, and defective neurogenesis (27-30). For instance, Li *et al.* (31) found that FTO facilitates leukemogenesis by modulating the expression of mRNAs of *ASB2* and *RARA* genes, which play crucial roles in leukocyte proliferation. Ma *et al.* (32) observed that METTL14 inhibits the metastatic transformation of hepatocellular carcinoma through promoting the binding of DGCR8 to pri-miR-126 and enhancing the maturation of miR-126. These studies provided insights into the roles of individual epitranscriptomic RWE proteins in cancer development; nonetheless, to our knowledge, there is no systematic investigation about how aberrant expression of epitranscriptomic RWE proteins modulates cancer progression.

CRC is well-known for its high occurrence and mortality, with approximately 1.93 million newly diagnosed cases and 0.94 million deaths in 2021 (33). The five-year survival rate for CRC patients with distant metastasis is as low as 5%, with an average of 13 months of survival after diagnosis (34).

Scheduled parallel-reaction monitoring (PRM) is a targeted proteomic method where predefined *m/z* values of precursor ions of peptides and their retention time information are incorporated into an inclusion list for MS/MS analyses (35). The PRM method capitalizes on the high-resolution, accurate-mass-measurement abilities of an Orbitrap or time-of-flight mass analyzer, which facilitate unambiguous identification and confident quantification of peptides in complex sample matrices (36). PRM, coupled with stable isotope labelling by amino acids in cell culture (SILAC) (37), also enables highly accurate, reproducible, and reliable quantification of proteins in cultured mammalian cells (38, 39).

In this study, we applied our recently established LC-PRM method (40), together with SILAC, to assess the expression differences in epitranscriptomic RWE proteins in a matched pair of primary/metastatic colorectal cancer cell lines (SW480/SW620) derived from the same patient (Figure 4.1a). We uncovered a number of differentially expressed epitranscriptomic RWE proteins and explored the roles of some of these proteins in CRC progression.

4.2 Materials and Methods

4.2.1 Cell culture

SW480 primary colorectal cancer (CRC) cells and SW620 lymph-node metastatic CRC cells derived from the same patient were purchased from American Type Culture Collection (ATCC) and cultured in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% fetal bovine serum (Invitrogen-Gibco) and 1% penicillin/streptomycin. For the SILAC experiments,(37) to the lysine, arginine-depleted SILAC medium were added dialyzed FBS (Invitrogen) and unlabeled lysine/arginine to yield the light DMEM media, or [$^{13}C_6$, $^{15}N_2$]-L-lysine/[$^{13}C_6$]-L-arginine to give the heavy DMEM media. SW480 or SW620 cells were cultured in the heavy media for at least five cell doublings to enable nearly complete heavy isotope labelling. Cells were kept at 37 °C in a humidified atmosphere containing 5% CO_2.

4.2.2 Tryptic digestion of whole-cell protein lysate

Details of cell culture procedures can be found in the Supporting Information. Light- or heavy-isotope labelled SW480 and SW620 cells were collected, and lysed on ice for 30 min with CelLytic M lysis buffer supplemented with 1% protease inhibitor cocktail. After centrifugation at 16,100 *g* for 30 min at 4 °C, Bradford assay was conducted to quantify total proteins in the supernatant. In forward SILAC experiments, the total protein lysate of SW480 cells cultured in the heavy medium and that of SW620 cells cultured in the light medium were combined at 1/1 ratio by mass. The reverse SILAC experiments were conducted in the opposite way. Proteins in the cell lysates were denatured, reduced, and alkylated before digestion with MS-grade trypsin (Pierce) in 50 mM NH_4HCO_3, pH 8.0, at 37 °C for 16 h according to the filter-aided sample preparation (FASP) procedure (41). The tryptic peptides were collected by centrifugation and desalted using OMIX C18 pipet tips prior to LC-PRM analysis.

4.2.3 LC-PRM data acquisition and analysis

The LC-PRM experiments were carried out on a Q Exactive Plus quadrupole-Orbitrap mass spectrometer coupled with a Dionex UltiMate 3000 RSLCnano UPLC system. Prior to analyzing the aforementioned SILAC samples, tryptic digestion mixture of BSA was analyzed under the same conditions to define the linear relationship between iRT and RT, since the iRT of each BSA peptide was pre-defined. The RTs of tryptic peptides for RWE proteins were predicted from their iRTs in the PRM library and the above-described iRT-RT relationship for BSA peptides. By using our recently developed Skyline PRM library (42), we were able to monitor the precursor ions representing the light and heavy forms of 444 unique tryptic peptides of 152 epitranscriptomic RWE proteins in three separate LC-MS/MS runs, where the inclusion lists encompassed the *m/z* value and RT window (7 min in width) for each precursor ion of interest.

The analytical column was packed in-house using 3 μm Reprosil-Pur C18-AQ resin (Dr. Maisch GmbH HPLC) stationary phase material in a 25-cm fused silica column (75 μm i.d.). The trapping column was also prepared in-house with 5 μm Reprosil-Pur C18-AQ resin (Dr. Maisch GmbH HPLC) in a 4-cm long fused silica column (150 μm i.d.). SILAC samples (500 ng) were separated using the analytical column with a 125-min linear gradient of 6-43% mobile phase B (80% acetonitrile in 0.1% formic acid) in mobile phase A (0.1% formic acid in Milli-Q water), where the flow rate was 300 nL/min. The spray voltage was 1.8 kV. The precursor ions were isolated in the quadrupole at an isolation window of 1.0 *m/z* and fragmented in the HCD collision cell at a normalized collision energy (NCE) of 28. MS/MS were acquired at a resolution of 17,500, an automated gain control (AGC) target 1×10^5, and a maximum accumulation time of 50 ms.

4.2.4 LC-PRM data analysis

After data acquisition, raw files were imported into Skyline for plotting the extracted-ion chromatograms for peak integration. For a positive detection of precursor ion of interest, 4-6 most abundant y ions from the same precursor ion in light and heavy forms should co-elute, and the relative abundances of fragment ions from the acquired MS/MS should match those in the MS/MS in the library, which is measured by dot product (dotp) value (43). A dotp value of > 0.7, ideally above 0.9, is considered to be highly similar. We manually excluded the potential interfering fragment ion if it does not overlay with other fragment ions. The sum of peak areas from the 4-6 fragment ions with overlaid elution profiles were employed for the quantification. The ratio of peak areas for the light and heavy forms of the peptide, provided by Skyline, reflects the relative abundance of that peptide, and by extension, the expression ratio of the corresponding protein, in SW620 over SW480 cells.

4.2.5 Western blot

After harvesting, SW480 and SW620 cells were lysed using CelLytic M (Sigma) lysis reagent supplemented with 1% protease inhibitor cocktail, and the ensuing proteins denatured in Laemmli loading buffer at 95 °C for 5 min. The same amount of total proteins from the paired CRC cells were resolved on an SDS-PAGE gel, and the proteins in the gel were then transferred onto a nitrocellulose membrane at 60 V for 90 min at 4 °C. The membrane was subsequently blocked using 5% non-fat dry milk in PBS-T (PBS with 0.1% Tween 20) for 45 min, and then incubated with PBS-T containing primary antibodies that recognize human FTO (Abclonal, A1438, 1:1000), hnRNPA2B1 (Santa Cruz, sc-53531, 1:1000), hnRNPC (Santa Cruz, sc-32308, 1:1000), and GAPDH (Santa Cruz, sc-32233, 1:10,000), at 4 °C overnight. After washing using PBS-T for five times, the membranes were incubated with donkey anti-rabbit (Sigma, A0545, 1:10,000) or anti-mouse (Santa Cruz, m-IgGκ BP-HRP, 1:10,000) secondary antibody in PBS-T at room temperature for 1 h. Prior to visualizing the protein bands using Amersham ECL™ Western Blot Detecting Reagent (GE Healthcare), the membranes were washed with PBS-T for five times.

4.2.6 Bioinformatic analyses

Gene Ontology (GO) analysis on biological process of genes encoding all up-regulated RWE proteins was carried out using Database for Annotation, Visualization and Integrated Discovery (DAVID, version 6.8; https://david.ncifcrf.gov/). Gene set enrichment analysis (GSEA) was performed in GSEA 4.1.0 software (http://www.broad.mit.edu/gsea). The Cancer Genome Atlas-Colon Adenocarcinoma (TCGA-COAD) dataset (n = 512) was downloaded from https://gdc.xenahubs.net/download/TCGA-COAD.htseq_fpkm.tsv.gz. The mRNA expression of each gene encoding the top 10 up-regulated RWE proteins obtained from the LC-PRM analysis was first stratified using its median value. GSEA was then conducted between the stratified TCGA-COAD and the hallmark gene sets (h.all.v7.4.symbols.gmt).

The relative expression levels of RWE proteins between colon cancer tissues and normal adjacent tissues of the Clinical Proteomic Tumor Analysis Consortium (CPTAC) samples (44) were retrieved using the UALCAN online tool (http://ualcan.path.uab.edu/) (45). The mRNA expression levels of genes encoding RWE proteins in normal, primary tumor, and metastatic tumor tissues of liver were retrieved from GSE41258 by using GEO2R (46). The outliers of each tissue group were identified and removed, where box-whisker plots were generated using an online tool (https://www.statskingdom.com/boxplot-maker.html) and re-plotted using Excel.

4.3 Results and Discussion

4.3.1 Scheduled LC-PRM Analysis Reveals Differentially Expressed Epitranscriptomic RWE Proteins in Metastatic SW620 Over Primary SW480 CRC cells

Our goal of this study was to interrogate systematically the contributions of epitranscriptomic RWE proteins in CRC metastasis. Toward this objective, we began with assessing differential expression of epitranscriptomic RWE proteins in a matched pair of primary/metastatic CRC cells derived from the same patient, i.e., the SW480/SW620 cells. By employing our recently developed LC-PRM method (40), in combination with SILAC, we were able to quantify 113 distinct epitranscriptomic RWE proteins in this pair of CRC cells; these proteins represent 74.3% of RWE proteins in the PRM library (Figure 4.1b). Positive identification was considered achieved when 4-6 transitions (i.e., product ions) from the same precursor ion exhibit the same retention time and a dot product (dotp) value (43) of > 0.7. In this regard, the dotp value gauges the similarities in relative abundances of fragment ions between the acquired MS/MS and the reference MS/MS in the PRM library that were acquired from previous shotgun proteomic analysis (47). Of the detected epitranscriptomic RWE proteins, 48 were up-regulated, and 5 were down-regulated by > 1.5-fold in the metastatic SW620 relative to the primary SW480 CRC cells. Among them, DUS2, DCP2, NAT10, and hnRNPC were markedly up-regulated by more than 4-fold in SW620 over SW480 cells (Figure 4.1c, and those proteins with expression ratios between 0.67 and 1.5-fold are shown in Figure 4.2a).

4.3.2 Scheduled LC-PRM Analysis Affords Highly Reproducible and Accurate Quantifications of Epitranscriptomic RWE Proteins

We next examined the reproducibility of the PRM method by comparing the data acquired from the four SILAC replicates. In this vein, we commonly detected 95 RWE proteins from the two forward and two reverse SILAC experiments (Figure 4.3a). The $\log_{10}$-transformed protein expression ratios obtained from the averaged ratios of two forward and those of two reverse SILAC labeling experiments exhibited an excellent linear fit (Figure 4.3b). Moreover, we calculated the replicate ratio based on LC-PRM results of all component peptides from each replicate, and we found that the mean relative standard deviation (RSD) of ratio of the RWE proteins from the four replicates of SILAC experiments was 21.8%. These results together support the consistency of quantification results obtained from four SILAC experiments. We also determined the ratios of all detected tryptic peptides of each RWE protein from the LC-PRM data of four SILAC replicates, and determined that the mean RSD of expression ratio for each protein from those of its component peptides was 9.1% (Table S1). This result reveals the relatively small variations among the different quantified peptides from the same protein. Together, the PRM method affords reproducible quantifications of RWE proteins.

LC-PRM also offers accurate quantification of peptides of target proteins since the quantification is based on their unique amino acid sequences derived from the proteins of interest. We further performed Western blot analyses for three proteins (i.e., FTO, hnRNPA2B1, and hnRNPC) (Figure 4.3c), and the results are in agreement with what we obtained from quantitative proteomic experiments, underscoring the accuracies of the PRM method. Figure 4.3d illustrates the extracted-ion chromatograms for the component peptides of FTO, hnRNPA2B1, and hnRNPC.

4.3.3 Analysis of the Up-Regulated RWE proteins from LC-PRM analysis using the CCLE database, GO, and GSEA

We queried the RNA-Seq data in the Cancer Cell Line Encyclopedia (CCLE) database for the mRNA expression levels of the top 10 up-regulated RWE proteins in SW620 over SW480 cells, as revealed from LC-PRM analysis. We uncovered that the mRNA levels of *DUS2*, *NAT10*, *ADAT2*, *DKC1*, *TARBP1*, *RBMX*, and *DKC1* genes were also up-regulated by over 1.5-fold in SW620 relative to SW480 cells. On the other hand, the mRNA levels of *DCP2*, *hnRNPC,* and *YRDC* genes differed slightly between SW480 and SW620 cells (Figure 4.4a, b). These results suggest that the augmented expressions of DCP2, hnRNPC, and YRDC in the metastatic CRC cells arise, at least in part, from post-transcriptional up-regulation.

We also conducted GO analysis on those RWE proteins that were up-regulated in metastatic over primary CRC cells by at least 1.5-fold. The result showed that these proteins are mainly associated with tRNA modification, methylation and processing, RNA methylation, and ψ synthesis (Figure 4.4c). GSEA identifies cumulative expression changes of multiple genes within a *priori* defined gene set displaying statistically significant difference in two phenotypes (48). GSEA of the top 10 up-regulated RWE proteins from the LC-PRM analysis was performed individually for the corresponding genes after stratifying the results to high- and low-mRNA expression groups according to their median values in the TCGA-COAD dataset, against the hallmark gene sets provided by the GSEA Molecular Signatures Database (48). Notably, we observed significantly ($p < 0.01$ and FDR < 0.25) enriched hallmark gene sets in the high-expression group of all the top 10 up-regulated RWE proteins except ADAT2 and TARBP1 (Table 4.1). Specifically, E2F targets, G2-M checkpoint, and MYC target V1 were among the most frequently enriched gene sets, where six out of eight RWE proteins display enrichment with those gene sets in the high-expression group of the protein (Figure 4.4d). Figure 4.2b depicts enrichment plots of DUS2. Those hallmark gene sets are related to cancer cell proliferation and tumor metastasis (49). Notably, E2F transcription factors 1 and 7 (E2F1 and E2F7) were found to modulate colon cancer metastasis and development (50). Some E2F targets, including EZH2 and BMP4, could mediate melanoma and breast cancer metastasis (51, 52). Moreover, it was shown that ER+/HER2- breast cancer with high activity in G2-M checkpoint pathway genes activity is more likely to metastasize (53). Furthermore, even though the association between c-Myc and CRC metastasis remained controversial (54), c-

Myc targets, such as YTHDF1 and AP4, were associated with CRC metastasis (55, 56). It is also worth noting that the mRNA expression levels of *DKC1* and *NAT10* are positively correlated with those of *MYC* in the TCGA-COAD dataset (Figure 4.2c).

Several differentially expressed RWE proteins identified in this study were found to be associated with cancer metastasis. For instance, DUS2, a tRNA-dihydrouridine synthase, is up-regulated in non-small cell lung carcinomas (NSCLC), and a higher level of DUS2 is accompanied with a poorer prognosis of lung cancer patients (57). In addition, NAT10, an acetyltransferase, promotes gastric cancer metastasis through inducing the formation of N^4-acetylcytidine in mRNA of *COL5A1* gene, and GSK-3β was found to promote the invasion of CRC cells through modulating the subcellular redistribution of NAT10 (58, 59). Moreover, hnRNPC, an m^6A reader protein, is involved in CRC progression since its elevated level in SW620 cells drives alternative cleavage and polyadenylation of *MTHFD1L* mRNA, a potential therapeutic target for CRC (60). Furthermore, DKC1, a ψ synthase, plays essential roles in angiogenesis and CRC metastasis by activating the transcription of HIF-1α, and may serve as a therapeutic target for CRC (61). Additionally, YRDC, an N^6-threonyl-carbamoyl-adenosine (t^6A) writer, promotes hepatocellular carcinoma by activating MEK/ERK signaling pathway (62). These studies of DUS2, NAT10, hnRNPC, DKC1, and YRDC support our proteomic results that these proteins may promote metastatic transformation of CRC.

4.3.4 NAT10, hnRNPC, DKC1, RBMX, and DUS1L May Also Be Accompanied with and Contribute to CRC Initiation

Vasaikar *et al.* (44) analyzed the global protein expression in matched tumor and tumor-adjacent normal tissues from 110 colon cancer patients from the CPTAC. For comparison, 75% of RWE proteins in the PRM library were quantified in our LC-PRM analysis, and CPTAC analysis allowed for the quantification of 66% of the RWE proteins included in the same library (Figure 4.5a).

We further investigated if the top 10 RWE proteins that are up-regulated in SW620 relative to SW480 cells also modulate colorectal cancer tumorigenesis. Our results showed that five differentially expressed RWE proteins, namely, NAT10, hnRNPC, DKC1, RBMX, and DUS1L, displayed pronounced differences in expression between tumor and tumor-adjacent normal tissues in the CPTAC samples (Figure 4.5b). This result indicates that elevated expressions of these proteins are positively correlated with both the initiation and metastatic transformation of CRC. On the other hand, DUS2, TARBP1, and YRDC were either not significantly correlated or slightly down-regulated in primary tumor tissues relative to tumor-adjacent normal tissues. These differences are not surprising, considering that tumor initiation and metastasis may also involve distinct molecular pathways, and that our PRM data were acquired from cell lines derived from a single patient. In addition, DCP2 and ADAT2 were quantified in SW480/SW620 cells in our PRM analysis, but not in CPTAC samples.

It is worth comparing our PRM method with the shotgun proteomic analysis employed in the CPTAC project. From the standpoint of sample preparation, tryptic digestion mixtures of CPTAC samples were first fractionated by reversed-phase LC into 96 fractions, concatenated into 12 fractions, and then subjected to LC-MS/MS analysis in the data-dependent acquisition mode. In contrast, the sample preparation of the paired SW480/SW620 cells for PRM analysis employed the whole-cell protein lysate without any pre-fractionation. With scheduled PRM, the entire library of tryptic peptides of RWE proteins (888 precursor ions, representing the light and heavy Lys- and Arg- labelled forms of 444 unique peptides) could be monitored in three LC-MS/MS runs, whereas 12 LC-MS/MS runs were employed in the CPTAC analysis. The lengths of the LC gradient for LC-MS/MS analyses employed in our PRM method (125 min) and CPTAC analysis (120 min) are very similar. Hence, aside from offering a better coverage of the epitranscriptomic RWE proteome, the PRM method affords a higher throughput than the shotgun proteomic method employed in the CPTAC study. Nevertheless, CPTAC analysis monitors the entire proteome, whereas the PRM method focuses selectively on the epitranscriptomic RWE proteome.

Apart from examining the protein expression of the top 10 up-regulated genes in the CPTAC samples, we explored GSE41258 (46), a microarray dataset for mRNA expression in normal, primary tumor, and liver metastasis tissues collected from colon cancer patients. Such analysis revealed markedly higher mRNA expressions of *NAT10*, *HNRNPC*, *DKC1*, *YRDC*, *TARBP1*, *RBMX*, and *DUS1L* genes in primary tumor tissues compared with normal tissues (Figure 4.5c), which is in keeping with the CPTAC analysis. Moreover, *DCP2*, *TARBP1*, and *DUS1L* genes display significantly higher mRNA expression in liver metastasis tissues relative to primary colon tumor tissues (Figure 4.5c), indicating that they may also modulate CRC metastasis.

4.4 Conclusion

We employed a scheduled LC-PRM method for highly sensitive, robust, and high-throughput profiling of epitranscriptomic RWE proteins accompanied with CRC metastasis. Our LC-PRM approach facilitated reproducible and accurate quantifications of 95 epitranscriptomic RWE proteins quantified in all four SILAC experiments, with 48 and 5 of these proteins being up- and down-regulated, respectively, by > 1.5-fold in SW620 metastatic CRC cells relative to SW480 primary CRC cells. NAT10, hnRNPC, and DKC1 exhibit pronounced up-regulations in the metastatic over primary CRC cells, and the roles of these proteins in CRC metastasis are known (59-61), which is in keeping with our LC-PRM analysis. Interrogation of publicly available data of CRC patients unveiled that the elevated expression of these and other epitranscriptomic RWE proteins are also accompanied with the initiation of CRC, suggesting their dual involvements in the initiation and metastatic transformation of CRC. We envision that the LC-PRM method developed herein can also be harnessed for future investigations about how epitranscriptomic modulators regulate the metastatic transformations of other types of cancer.

References

1. Kadumuri RV, Janga SC. Epitranscriptomic Code and Its Alterations in Human Disease. Trends in molecular medicine. 2018;24:886-903.

2. Pan T. Modifications and functional genomics of human transfer RNA. Cell Res. 2018;28:395-404.

3. Roundtree IA, Evans ME, Pan T, He C. Dynamic RNA modifications in gene expression regulation. Cell. 2017;169:1187-200.

4. 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;485:201-6.

5. Knuckles P, Lence T, Haussmann IU, Jacob D, Kreim N, Carl SH, et al. Zc3h13/Flacc is required for adenosine methylation by bridging the mRNA-binding factor Rbm15/Spenito to the m^6A machinery component Wtap/Fl(2)d. Genes Dev. 2018;32:415-29.

6. Liu J, Yue Y, Han D, Wang X, Fu Y, Zhang L, et al. A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. Nat Chem Biol. 2014;10:93-5.

7. Ping XL, Sun BF, Wang L, Xiao W, Yang X, Wang WJ, et al. Mammalian WTAP is a regulatory subunit of the RNA N^6-methyladenosine methyltransferase. Cell Res. 2014;24:177-89.

8. Yue Y, Liu J, Cui X, Cao J, Luo G, Zhang Z, et al. VIRMA mediates preferential m^6A mRNA methylation in 3'UTR and near stop codon and associates with alternative polyadenylation. Cell Discov. 2018;4:10.

9. Pendleton KE, Chen B, Liu K, Hunter OV, Xie Y, Tu BP, et al. The U6 snRNA m_6A methyltransferase METTL16 regulates SAM synthetase intron retention. Cell. 2017;169:824-35 e14.

10. Hsu PJ, Zhu Y, Ma H, Guo Y, Shi X, Liu Y, et al. Ythdc2 is an N^6-methyladenosine binding protein that regulates mammalian spermatogenesis. Cell Res. 2017;27:1115-27.

11. Li A, Chen YS, Ping XL, Yang X, Xiao W, Yang Y, et al. Cytoplasmic m^6A reader YTHDF3 promotes mRNA translation. Cell Res. 2017;27:444-7.

12. Wang X, Lu Z, Gomez A, Hon GC, Yue Y, Han D, et al. N6-methyladenosine-dependent regulation of messenger RNA stability. Nature. 2014;505:117-20.

13. Jia G, Fu Y, Zhao X, Dai Q, Zheng G, Yang Y, et al. N^6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. Nat Chem Biol. 2011;7:885-7.

14. Zheng G, Dahl JA, Niu Y, Fedorcsak P, Huang CM, Li CJ, et al. ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. Mol Cell. 2013;49:18-29.

15. Mauer J, Luo X, Blanjoie A, Jiao X, Grozhik AV, Patil DP, et al. Reversible methylation of m^6A_m in the 5' cap controls mRNA stability. Nature. 2017;541:371-5.

16. Dominissini D, Nachtergaele S, Moshitch-Moshkovitz S, Peer E, Kol N, Ben-Haim MS, et al. The dynamic N^1-methyladenosine methylome in eukaryotic messenger RNA. Nature. 2016;530:441-6.

17. Li X, Xiong X, Wang K, Wang L, Shu X, Ma S, et al. Transcriptome-wide mapping reveals reversible and dynamic N^1-methyladenosine methylome. Nat Chem Biol. 2016;12:311-6.

18. Safra M, Sas-Chen A, Nir R, Winkler R, Nachshon A, Bar-Yaacov D, et al. The m^1A landscape on cytosolic and mitochondrial mRNA at single-base resolution. Nature. 2017;551:251-5.

19. Squires JE, Patel HR, Nousch M, Sibbritt T, Humphreys DT, Parker BJ, et al. Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. Nucleic Acids Res. 2012;40:5023-33.

20. Yang X, Yang Y, Sun BF, Chen YS, Xu JW, Lai WY, et al. 5-methylcytosine promotes mRNA export - NSUN2 as the methyltransferase and ALYREF as an m(5)C reader. Cell Res. 2017;27:606-25.

21. Dai X, Gonzalez G, Li L, Li J, You C, Miao W, et al. YTHDF2 Binds to 5-Methylcytosine in RNA and Modulates the Maturation of Ribosomal RNA. Anal Chem. 2020;92:1346-54.

22. Lin S, Liu Q, Lelyveld VS, Choe J, Szostak JW, Gregory RI. Mettl1/Wdr4-mediated m^7G tRNA methylome is required for normal mRNA translation and embryonic stem cell self-renewal and differentiation. Mol Cell. 2018;71:244-55 e5.

23. Carlile TM, Rojas-Duran MF, Zinshteyn B, Shin H, Bartoli KM, Gilbert WV. Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. Nature. 2014;515:143-6.

24. Schwartz S, Bernstein DA, Mumbach MR, Jovanovic M, Herbst RH, Leon-Ricardo BX, et al. Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. Cell. 2014;159:148-62.

25. Darzacq X, Jady BE, Verheggen C, Kiss AM, Bertrand E, Kiss T. Cajal body-specific small nuclear RNAs: a novel class of 2 '-O-methylation and pseudouridylation guide RNAs. EMBO J. 2002;21:2746-56.

26. Rebane A, Roomere H, Metspalu A. Locations of several novel 2 '-O-methylated nucleotides in human 28S rRNA. BMC Mol Biol. 2002;3.

27. Frye M, Harada BT, Behm M, He C. RNA modifications modulate gene expression during development. Science. 2018;361:1346-9.

28. Tirumuru N, Zhao BS, Lu W, Lu Z, He C, Wu L. N^6-methyladenosine of HIV-1 RNA regulates viral infection and HIV-1 Gag protein expression. eLife. 2016;5.

29. Yoon KJ, Ringeling FR, Vissers C, Jacob F, Pokrass M, Jimenez-Cyrus D, et al. Temporal control of mammalian cortical neurogenesis by m^6A methylation. Cell. 2017;171:877-89 e17.

30. Lan Q, Liu PY, Haase J, Bell JL, Huttelmaier S, Liu T. The critical role of RNA m^6A methylation in cancer. Cancer Res. 2019;79:1285-92.

31. Li Z, Weng H, Su R, Weng X, Zuo Z, Li C, et al. FTO Plays an Oncogenic Role in Acute Myeloid Leukemia as a N(6)-Methyladenosine RNA Demethylase. Cancer Cell. 2017;31:127-41.

32. Ma JZ, Yang F, Zhou CC, Liu F, Yuan JH, Wang F, et al. METTL14 suppresses the metastatic potential of hepatocellular carcinoma by modulating N(6) -methyladenosine-dependent primary MicroRNA processing. Hepatology. 2017;65:529-43.

33. Xi Y, Xu P. Global colorectal cancer burden in 2020 and projections to 2040. Transl Oncol. 2021;14:101174.

34. Zacharakis M, Xynos ID, Lazaris A, Smaro T, Kosmas C, Dokou A, et al. Predictors of survival in stage IV metastatic colorectal cancer. Anticancer Res. 2010;30:653-60.

35. Peterson AC, Russell JD, Bailey DJ, Westphall MS, Coon JJ. Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics. Molecular & cellular proteomics : MCP. 2012;11:1475-88.

36. Weng LW, Guo LL, Vachani A, Mesaros C, Blai IA. Quantification of Serum High Mobility Group Box 1 by Liquid Chromatography/High-Resolution Mass Spectrometry: Implications for Its Role in Immunity, Inflammation, and Cancer. Anal Chem. 2018;90:7552-60.

37. Ong SE, 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;1:376-86.

38. Miao WL, Li L, Wang YS. Identification of Helicase Proteins as Clients for HSP90. Anal Chem. 2018;90:11751-5.

39. Miao WL, Guo L, Wang YS. Imatinib-Induced Changes in Protein Expression and ATP-Binding Affinities of Kinases in Chronic Myelocytic Leukemia Cells. Anal Chem. 2019;91:3209-14.

40. Qi TF, Miao W, Wang Y. Targeted Profiling of Epitranscriptomic Reader, Writer, and Eraser Proteins Accompanied with Radioresistance in Breast Cancer Cells. Anal Chem. 2022.

41. Wisniewski JR, Zougman A, Nagaraj N, Mann M. Universal sample preparation method for proteome analysis. Nat Meth. 2009;6:359-62.

42. Qi TF, Miao W, Wang Y. Targeted profiling of epitranscriptomic reader, writer, and eraser proteins accompanied with radioresistance in breast cancer cells. Anal Chem. 2022;94:1525-30.

43. de Graaf EL, Altelaar AFM, van Breukelen B, Mohammed S, Heck AJR. 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.

44. Vasaikar S, Huang C, Wang X, Petyuk VA, Savage SR, Wen B, et al. Proteogenomic Analysis of Human Colon Cancer Reveals New Therapeutic Opportunities. Cell. 2019;177:1035-49 e19.

45. Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi B, et al. UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. Neoplasia. 2017;19:649-58.

46. Sheffer M, Bacolod MD, Zuk O, Giardina SF, Pincas H, Barany F, et al. Association of survival and disease progression with chromosomal instability: a genomic exploration of colorectal cancer. Proc Natl Acad Sci U S A. 2009;106:7131-6.

47. Miao WL, Li L, Wang YS. A Targeted Proteomic Approach for Heat Shock Proteins Reveals DNAJB4 as a Suppressor for Melanoma Metastasis. Anal Chem. 2018;90:6835-42.

48. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005;102:15545-50.

49. Gong Y, Liu Y, Wang T, Li Z, Gao L, Chen H, et al. Age-Associated Proteomic Signatures and Potential Clinically Actionable Targets of Colorectal Cancer. Mol Cell Proteomics. 2021;20:100115.

50. Fang Z, Lin M, Li C, Liu H, Gong C. A comprehensive review of the roles of E2F1 in colon cancer. Am J Cancer Res. 2020;10:757-68.

51. White JR, Thompson DT, Koch KE, Kiriazov BS, Beck AC, van der Heide DM, et al. AP-2α-Mediated Activation of E2F and EZH2 Drives Melanoma Metastasis. Cancer Res. 2021;81:4455-70.

52. Hollern DP, Honeysett J, Cardiff RD, Andrechek ER. The E2F transcription factors regulate tumor development and metastasis in a mouse model of metastatic breast cancer. Mol Cell Biol. 2014;34:3229-43.

53. Oshi M, Takahashi H, Tokumaru Y, Yan L, Rashid OM, Matsuyama R, et al. G2M Cell Cycle Pathway Score as a Prognostic Biomarker of Metastasis in Estrogen Receptor (ER)-Positive Breast Cancer. Int J Mol Sci. 2020;21.

54. He W-L, Weng X-T, Wang J-L, Lin Y-K, Liu T-W, Zhou Q-Y, et al. Association Between c-Myc and Colorectal Cancer Prognosis: A Meta-Analysis. Front Physiol. 2018;9.

55. Nishizawa Y, Konno M, Asai A, Koseki J, Kawamoto K, Miyoshi N, et al. Oncogene c-Myc promotes epitranscriptome m(6)A reader YTHDF1 expression in colorectal cancer. Oncotarget. 2017;9:7476-86.

56. Jackstadt R, Röh S, Neumann J, Jung P, Hoffmann R, Horst D, et al. AP4 is a mediator of epithelial-mesenchymal transition and metastasis in colorectal cancer. J Exp Med. 2013;210:1331-50.

57. Kato T, Daigo Y, Hayama S, Ishikawa N, Yamabuki T, Ito T, et al. A novel human tRNA-dihydrouridine synthase involved in pulmonary carcinogenesis. Cancer Res. 2005;65:5638-46.

58. Zhang Y, Jing Y, Wang Y, Tang J, Zhu X, Jin WL, et al. NAT10 promotes gastric cancer metastasis via N4-acetylated COL5A1. Signal Transduct Target Ther. 2021;6:173.

59. Zhang H, Hou W, Wang HL, Liu HJ, Jia XY, Zheng XZ, et al. GSK-3β-regulated N-acetyltransferase 10 is involved in colorectal cancer invasion. Clin Cancer Res. 2014;20:4717-29.

60. Fischl H, Neve J, Wang ZQ, Patel R, Louey A, Tian B, et al. hnRNPC regulates cancer-specific alternative cleavage and polyadenylation profiles. Nucleic Acids Res. 2019;47:7580-91.

61. Hou PF, Shi PC, Jiang T, Yin H, Chu SF, Shi ML, et al. DKC1 enhances angiogenesis by promoting HIF-1 alpha transcription and facilitates metastasis in colorectal cancer. Br J Cancer. 2020;122:668-79.

62. Huang S, Zhu P, Sun B, Guo J, Zhou H, Shu Y, et al. Modulation of YrdC promotes hepatocellular carcinoma progression via MEK/ERK signaling pathway. Biomed Pharmacother. 2019;114:108859.

Figure 4.1 LC-PRM coupled with SILAC for quantifying epitranscriptomic RWE proteins in SW620 (metastatic) and SW480 (primary) CRC cells.

(a) A SILAC- and LC-PRM-based workflow for targeted quantifications of epitranscriptomic RWE proteins in SW620 (metastatic) and SW480 (primary) CRC cells. (b) A Venn diagram showing the number of quantified RWE proteins in SW480/SW620 cells in comparison with that of RWE proteins deposited in the PRM library. (c) A bar graph depicting the relative expression levels of differentially expressed RWE proteins in SW620 vs. SW480 CRC cells. Red and blue bars designate those proteins with expression ratios in SW620/SW480 cells being > 1.5 and < 0.67, respectively. Error bars represent S.D. of results obtained from a total of four SILAC experiments after determining the protein expression ratio in each replicate based on LC-PRM results of all component peptides of the protein.

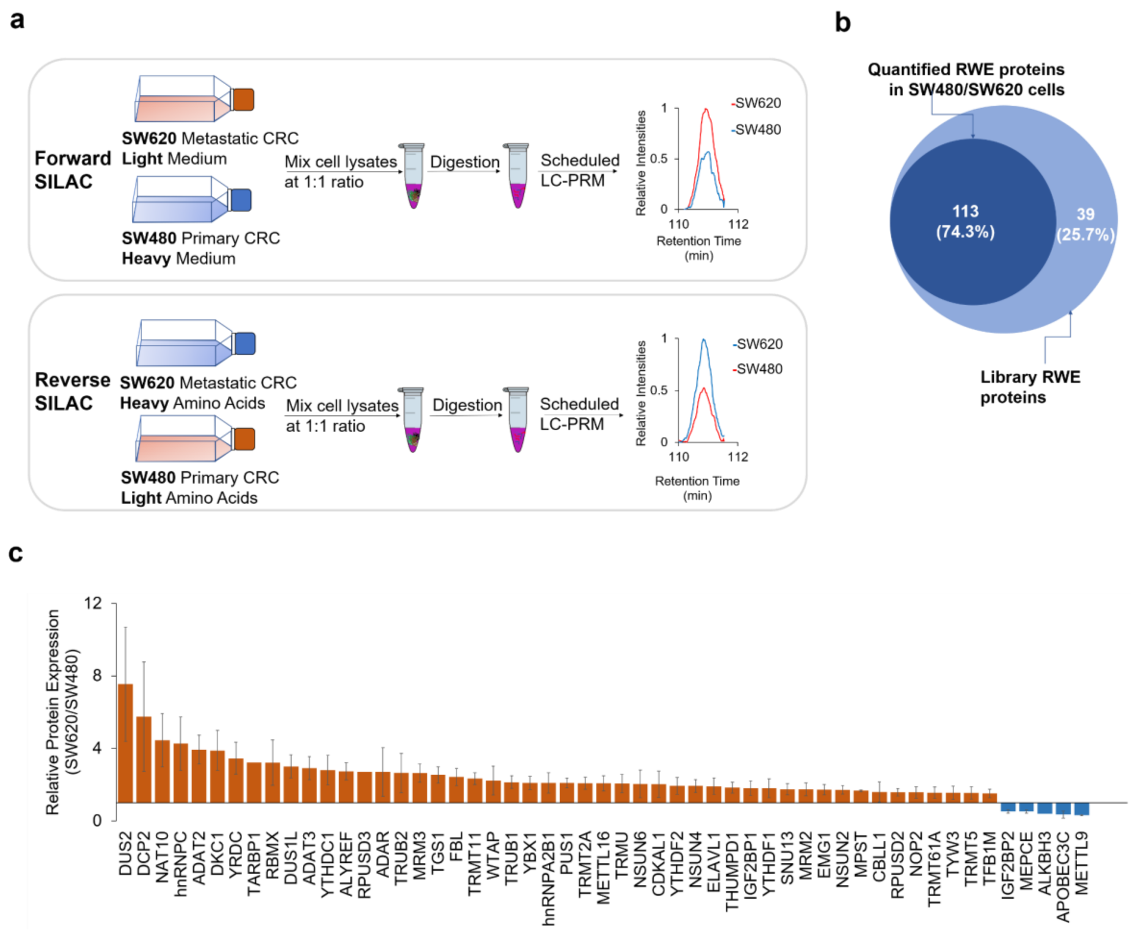
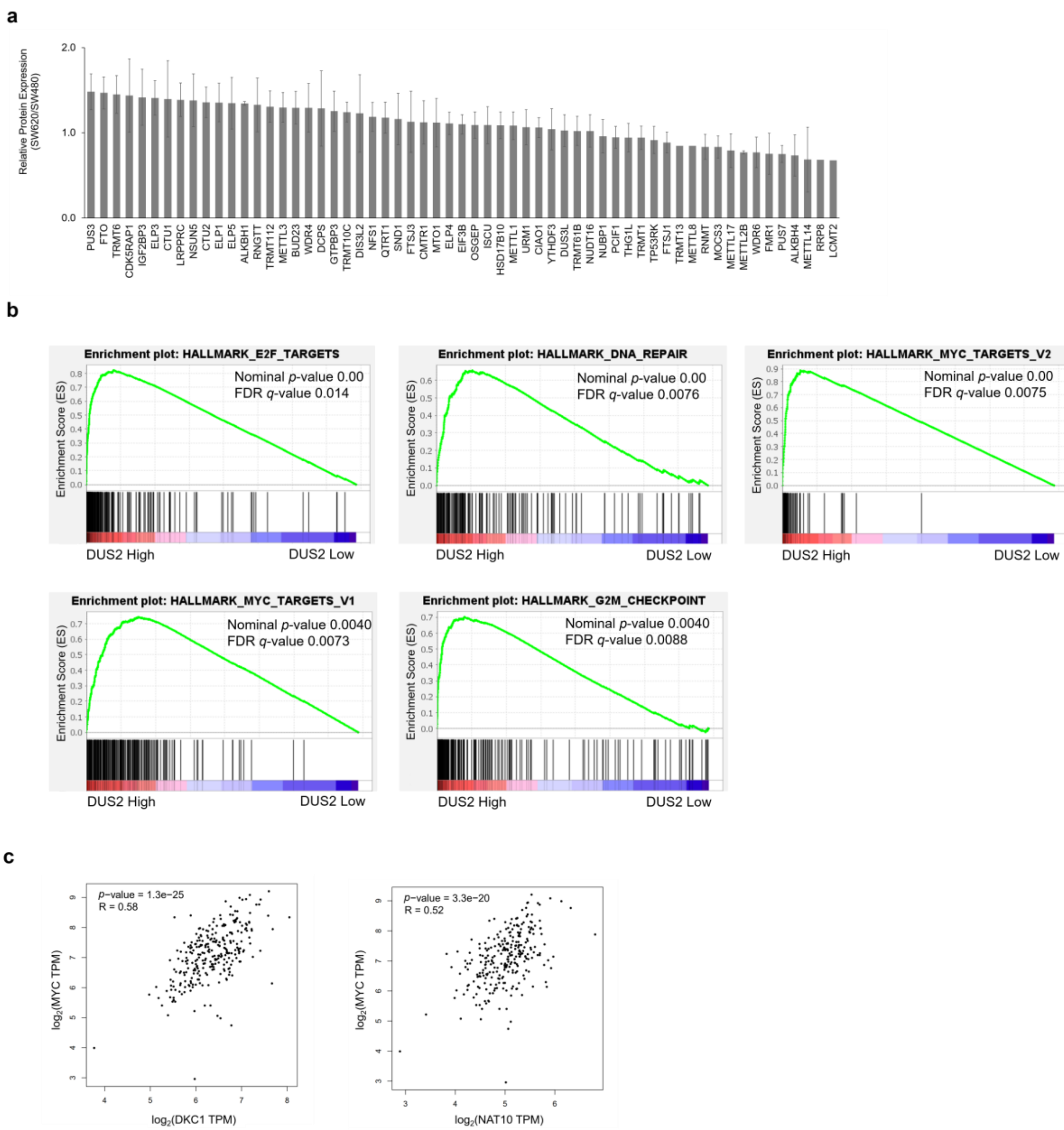


Figure 4.2 (a) A bar graph depicting the relative expression levels of those RWE proteins with expression ratios in SW620 vs. SW480 CRC cells being between 0.67 and 1.5.

Error bars represent S.D. of results obtained from two forward and two reverse SILAC experiments. (b) Enrichment plots of gene sets significantly enriched in the high-DUS2-expression group, generated by GSEA 4.1.0. (c) Scatter plots showing the correlation between mRNA expression levels of *MYC* and those of *DKC1* or *NAT10* in TCGA-COAD dataset. The plots were generated using gene expression profiling interactive analysis (GEPIA). Spearman correlation coefficients are displayed.

a



Relative Protein Expression (SW620/SW480)
PUS3
FTO
TRMT6
CDK5RAP1
IGF2BP3
ELP3
CTU1
LRPPRC
NSUN5
CTU2
ELP1
ELP5
ALKBH1
RNGTT
TRMT112
METTL3
BUD23
WDR4
DCPS
GTPBP3
TRMT10C
DIS3L2
NFS1
QTRT1
SND1
FTSJ3
CMTR1
MTO1
ELP4
EIF3B
OSGEP
ISCU
HSD17B10
METTL1
URM1
CIAO1
YTHDF3
DUS3L
TRMT61B
NUDT16
NUBP1
PCIF1
THG1L
TRMT1
TP53RK
FTSJ1
TRMT13
METTL8
RNMT
MOCS3
METTL17
METTL2B
WDR6
FMR1
PUS7
ALKBH4
METTL14
RRP8
LCMT2
b
Enrichment plot: HALLMARK_E2F_TARGETS
Nominal p-value 0.00
FDR q-value 0.014
DUS2 High
DUS2 Low
Enrichment plot: HALLMARK_DNA_REPAIR
Nominal p-value 0.00
FDR q-value 0.0076
Enrichment plot: HALLMARK_MYC_TARGETS_V2
Nominal p-value 0.00
FDR q-value 0.0075
Enrichment plot: HALLMARK_MYC_TARGETS_V1
Nominal p-value 0.0040
FDR q-value 0.0073
Enrichment plot: HALLMARK_G2M_CHECKPOINT
Nominal p-value 0.0040
FDR q-value 0.0088
Enrichment Score (ES)
c
p-value = 1.3e-25
R = 0.58
log2(MYC TPM)
log2(DKC1 TPM)
p-value = 3.3e-20
R = 0.52
log2(NAT10 TPM)

Figure 4.3 PRM method validation in reproducibility, consistency, and accuracy.

(a) A Venn diagram illustrating the numbers of RWE proteins quantified in four SILAC experiments. (b) Log_{10}(Ratio) of the average protein expression levels in SW620 over SW480 cells obtained from two forward and two reverse SILAC experiments. One outlier (VIRMA, its quantification results from forward and reverse SILAC experiments display a large discrepancy) was excluded. The top ten up-regulated RWE proteins were labeled in red dots, except DCP2 and TARBP1, which were quantified only in forward SILAC experiments. (c) Western blot analyses of FTO, hnRNPA2B1 and hnRNPC proteins in SW480 and SW620 cells, and the comparison of the quantification data obtained from Western blot (n = 3) and PRM analyses (n = 4). (d) Extracted-ion chromatograms of representative peptides of FTO (LFTVPWPVK), hnRNPA2B1 (IDTIEIITDR), and hnRNPC (MIAGQVLDINLAAEPK), obtained from LC-PRM analysis in one forward (SW_F1) and one reverse SILAC experiment (SW_R1).

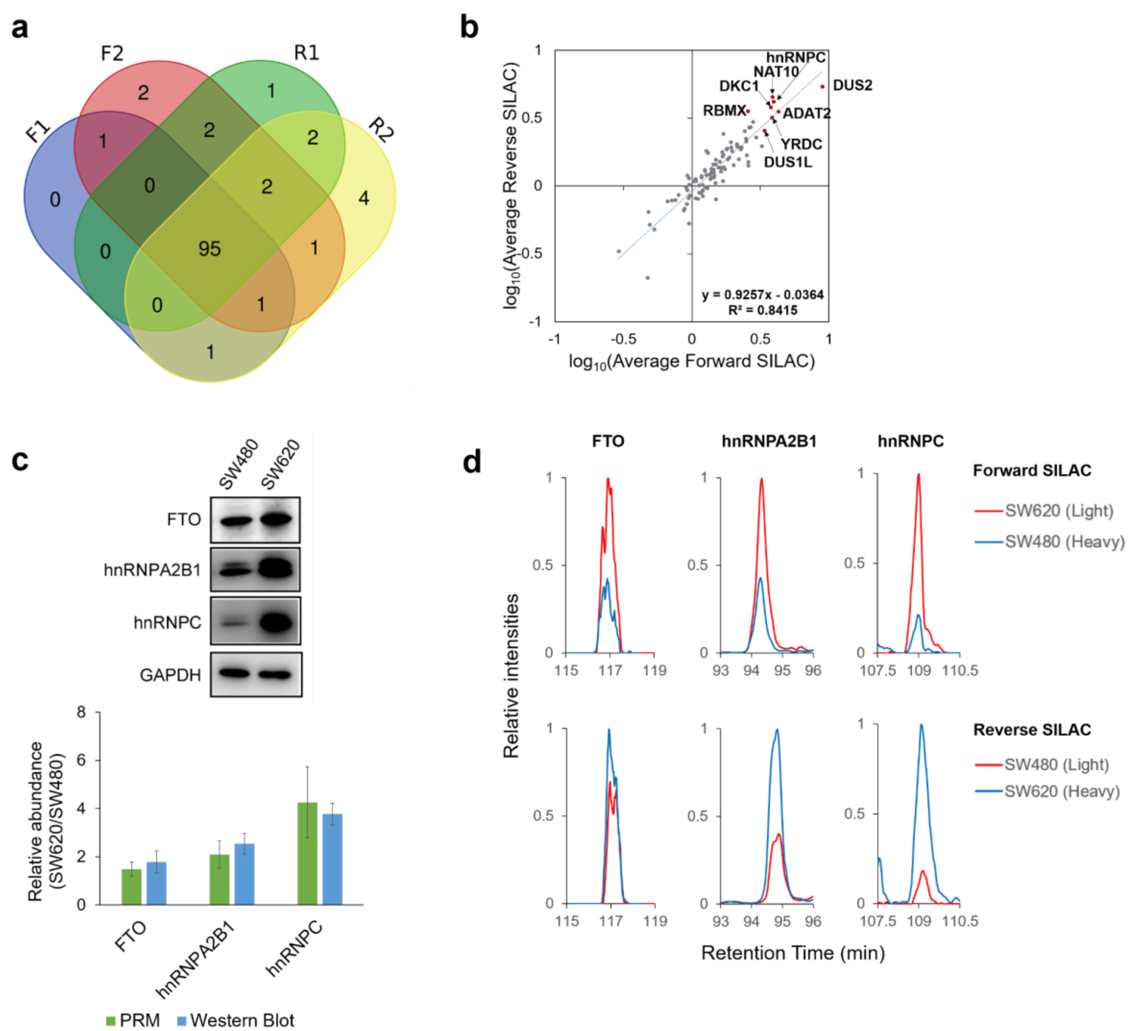


Figure 4.4 Bioinformatics of up-regulated RWE proteins in SW620 over SW480 cells.

(a) mRNA expression levels of the top 10 up-regulated RWE genes identified from LC-PRM analysis. The data were retrieved from the CCLE database. (b) The comparison of relative mRNA levels obtained from the CCLE database and relative protein levels of the top 10 up-regulated RWE genes in SW620 versus SW480 cells, as obtained from LC-PRM analysis. (c) A bar graph illustrating GO analysis on biological process (BP) of all up-regulated RWE genes. The BP results were sorted by Benjamini FDR from smallest to largest using 0.01 as a cutoff. (d) A bar graph showing GSEA results of the sum of particular enriched gene sets of top 10 up-regulated RWE genes identified from LC-PRM analysis. TCGA-COAD dataset was first stratified to high- and low-expression groups using the median value of mRNA expression of a specific RWE protein among all patient tissues. The significance level in GSEA analysis was defined as nominal p-value < 0.01 and FDR q-value < 0.25. The number of permutations was set at 1000.

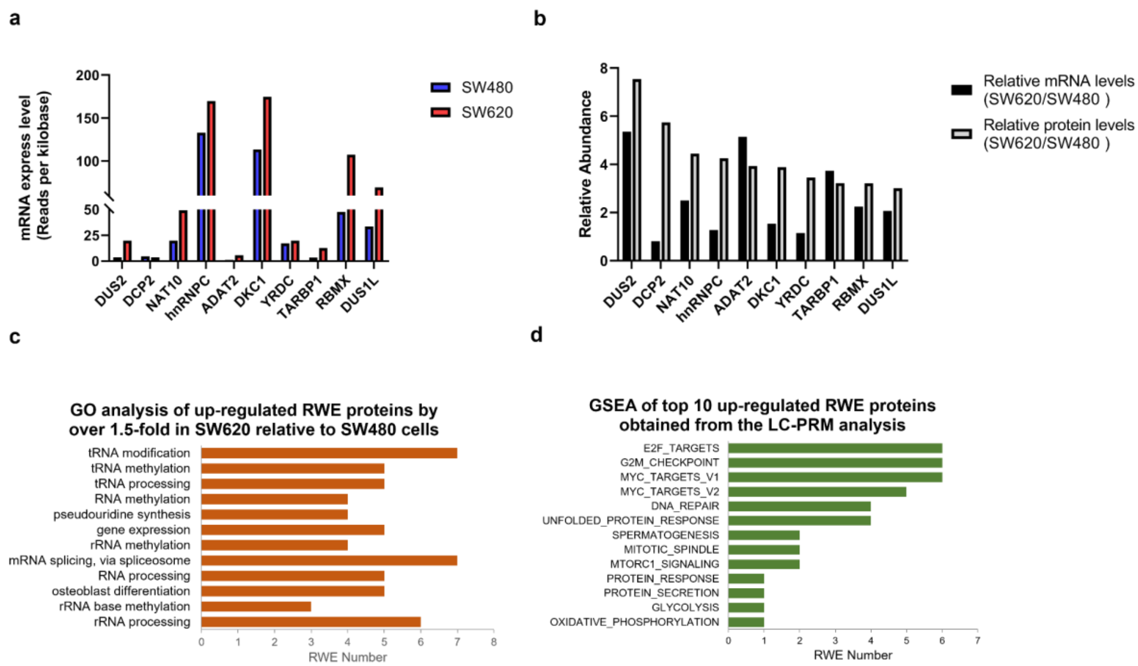


Figure 4.5 Analysis of top 10 RWE proteins in CPTAC and GSE41258 datasets.

(a) A Venn diagram illustrating the number of RWE proteins quantified from the PRM analysis in this study and the previously reported CPTAC analysis, in comparison with the total number of RWE proteins deposited in the PRM library. (b) Relative protein expression levels of NAT10, hnRNPC, DKC1, RBMX, and DUS1L in primary colon tumor tissues and tumor-adjacent normal tissues in the CPTAC samples. Z-values represent S.D. from the median across samples. (c) Relative mRNA expression levels of *DCP2*, *NAT10*, *HNRNPC*, *DKC1*, *YRDC*, *TARBP1*, *RBMX*, and *DUS1L* genes in normal tissues, primary colon tumor tissues, and liver metastasis tissues in GSE41258. The p values were calculated using an unpaired two-tailed Student's t test. For (b-c), the horizontal edges and inner line of the box illustrate the upper/lower quartiles and median, respectively. The top/bottom ends of the whisker denote the maximum/minimum values.

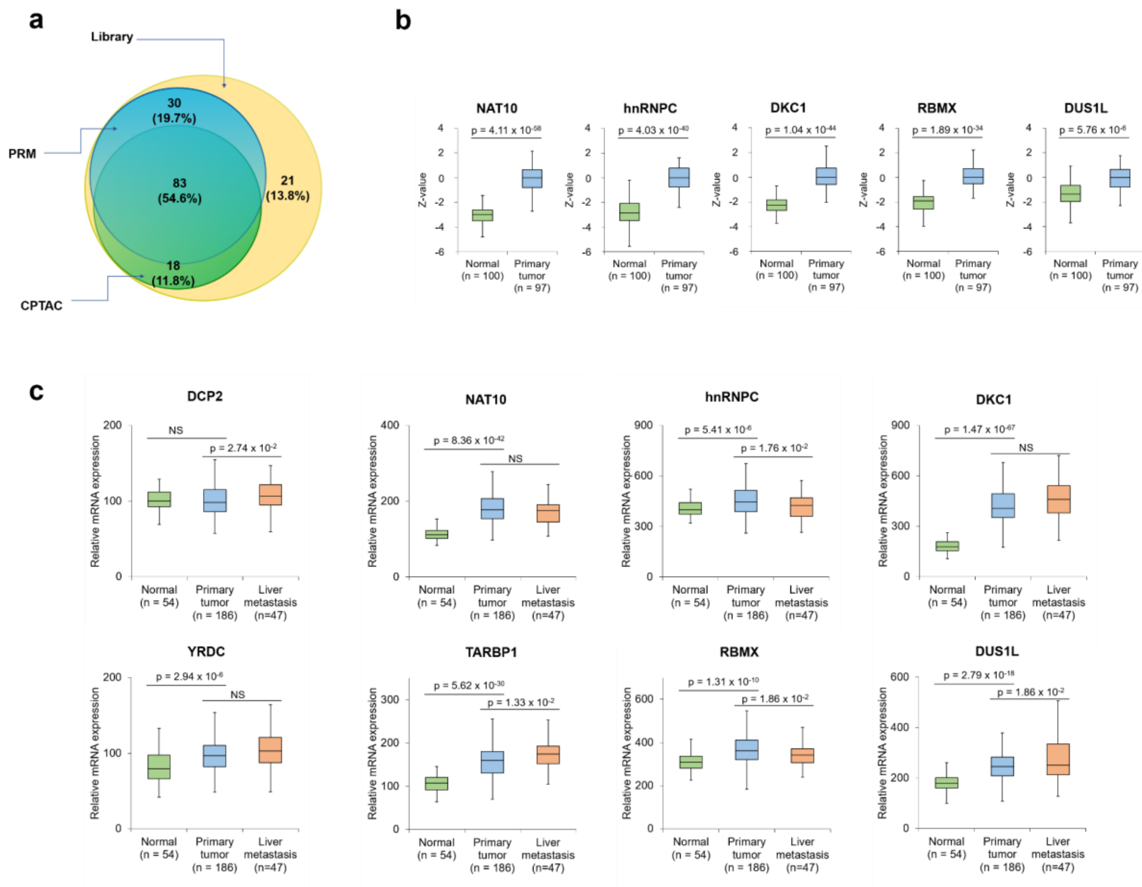


Table 4.1 A summary of GSEA enrichment results of each of the top 10 up-regulated RWE proteins obtained from LC-PRM analysis.

Gene sets with nominal *p*-value (<0.01) and FDR *q*-value (<0.25) were shown.

Gene	Number of gene sets are significant at FDR < 25%	Number of gene sets are significant at nominal pvalue < 1%	Enrichment results	Normalized Enrichment Score (NES)	Nominal *p*-value	FDR *q*-value
DUS2	17	5	E2F_TARGETS	2.14	0.0E+00	1.4E-02
			DNA_REPAIR	2.12	0.0E+00	7.6E-03
			MYC_TARGETS_V2	2.09	0.0E+00	7.5E-03
			MYC_TARGETS_V1	2.07	4.0E-03	7.3E-03
			G2M_CHECKPOINT	2.02	4.1E-03	8.8E-03
DCP2	6	2	MITOTIC_SPINDLE	2.11	2.1E-03	2.4E-02
			SPERMATOGENESIS	1.70	4.1E-03	2.0E-01
NAT10	13	7	UNFOLDED_PROTEIN_RESPONSE	2.36	0.0E+00	0.0E+00
			DNA_REPAIR	2.17	2.0E-03	3.3E-03
			G2M_CHECKPOINT	2.16	0.0E+00	3.3E-03
			MYC_TARGETS_V1	2.08	2.0E-03	8.3E-03
			E2F_TARGETS	2.07	0.0E+00	7.7E-03
			MITOTIC_SPINDLE	2.04	1.9E-03	7.8E-03
			MYC_TARGETS_V2	2.01	0.0E+00	8.8E-03
HNRNPC	15	8	MTORC1_SIGNALING	2.38	0.0E+00	0.0E+00
			G2M_CHECKPOINT	2.38	0.0E+00	0.0E+00
			MYC_TARGETS_V1	2.20	0.0E+00	3.0E-03
			PROTEIN_SECRETION	2.19	3.9E-03	2.7E-03
			UNFOLDED_PROTEIN_RESPONSE	2.19	0.0E+00	2.1E-03
			E2F_TARGETS	2.17	0.0E+00	2.4E-03
			DNA_REPAIR	2.01	9.5E-03	1.5E-02
			SPERMATOGENESIS	1.98	0.0E+00	1.5E-02
DKC1	13	5	UNFOLDED_PROTEIN_RESPONSE	2.25	0.0E+00	9.6E-04
			MYC_TARGETS_V1	2.16	2.0E-03	4.2E-03
			E2F_TARGETS	2.12	2.0E-03	6.5E-03
			G2M_CHECKPOINT	2.07	0.0E+00	8.5E-03
			MYC_TARGETS_V2	1.98	0.0E+00	1.3E-02
YRDC	18	9	UNFOLDED_PROTEIN_RESPONSE	2.46	0.0E+00	5.1E-04
			MTORC1_SIGNALING	2.43	0.0E+00	2.6E-04
			MYC_TARGETS_V1	2.30	0.0E+00	6.3E-04
			DNA_REPAIR	2.26	0.0E+00	8.6E-04
			OXIDATIVE_PHOSPHORYLATION	2.21	2.0E-03	1.5E-03
			E2F_TARGETS	2.20	0.0E+00	1.3E-03
			G2M_CHECKPOINT	2.16	0.0E+00	2.1E-03
			GLYCOLYSIS	2.03	8.0E-03	7.3E-03
			MYC_TARGETS_V2	2.02	0.0E+00	6.9E-03
RBMX	11	3	G2M_CHECKPOINT	2.05	2.0E-03	2.6E-02
			E2F_TARGETS	2.02	0.0E+00	1.8E-02
			MYC_TARGETS_V1	1.85	2.0E-03	4.4E-02
DUS1L	5	1	MYC_TARGETS_V2	1.97	1.9E-03	6.9E-02

5. Chapter 5. Targeted Quantitative Profiling of Epitranscriptomic Reader, Writer and Eraser Proteins Revealed Potential Crosstalk between N^6-Methyladenosine and Other RNA Modifications

5.1 Introduction

There has been a surging interest in the field of epitranscriptomics in recent years. *N*^{6}-methyladenosine (m^6A) in mRNA was first identified in mouse L cells in 1974 (1). Recent pioneering work about m^6A included transcriptome-wide mapping of m^6A (2, 3), as well as the discoveries of MTA70 (METTL3) in bacteria and demonstration of its function as an m^6A methyltransferase (i.e., writer) (4), FTO and ALKBH5 as mammalian m^6A demethylases (i.e., erasers) (5, 6), and m^6A-binding proteins (i.e., readers, e.g., YTHDF2 and YTHDF1) (7, 8). These reader, writer, eraser (RWE) proteins of m^6A assume important roles in modulating the splicing (9), stability (8, 10), and translation efficiencies of mRNA (7, 11, 12).

Several recent studies revealed interplays between m^6A and other RNA modifications (e.g., m^5C and m^1A), though the underlying mechanisms remain poorly investigated. For example, METTL3/METTL14, the core subunits of the m^6A writer complex, and m^5C writer NSUN2 regulate 3′ untranslated region (3′UTR) of *p21* mRNA and synergistically enhance its expression (13). YTHDF2, an m^6A reader protein, is also capable of binding directly with m^5C in RNA, albeit at a lower affinity than that toward m^6A (14). Moreover, YTHDF1-3 and YTHDC1, which are well-established m^6A readers, can also directly recognize m^1A in RNA using the same hydrophobic binding pocket for m^6A binding (15). Insights into the potential interplays between m^6A and m^1A involve the discovery of FTO, a well-recognized eraser for m^6A and m^6A_m, as a demethylase of m^1A in tRNA *in vitro* and *in vivo* (16). Despite the above-described studies, to our knowledge, there has been no systematic investigation about potential crosstalk between m^6A and other RNA modifications.

5.2 Materials and Methods

5.2.1 Cell culture

HEK293T cells were cultured in Dulbecco's modified Eagle's medium (DMEM) medium complemented with 10% fetal bovine serum (FBS) and 1% penicillin−streptomycin. HEK293T cells with *ALKBH5*, *FTO*, and *METTL3*, being individually ablated with CRISPR/Cas 9 genome editing (17). The cells were maintained at 37°C in a humidified chamber supplemented with 5% CO2.

5.2.2 Crude SIL peptides

A total of 48 crude SIL peptides ([$^{13}C_6$, $^{15}N_2$]-Lys and [$^{13}C_6$, $^{15}N_4$]-Arg) representing 45 RWE proteins were synthesized and purified by Vivitide (Gardner, MA). The peptide purity was around 75% and isotopic purity was around 99%. The full list of the SIL peptides can be found in Table 5.1. Each SIL peptide was reconstituted with 15% acetonitrile in 0.1% formic acid. All SIL peptides were mixed as a stock solution for spiking into proteomic samples.

5.2.3 Proteomic sample preparation

ALKBH5$^{-/-}$, *FTO*$^{-/-}$, and *METTL3*$^{-/-}$ and the isogenic parental HEK293T cells were harvested, and proteomic samples were prepared using a filter-aided sample preparation (FASP) method (18) with minor modifications as described elsewhere (19). After desalting tryptic peptides using Pierce C18 Tips (Thermo Fisher), proteomic samples were spiked in a mixture of SIL peptides at a final concentration at 2 fmol/μL. Peptides (500 ng – 1 μg) and SIL peptide mixture (4 fmol) were subjected to LC-PRM for analysis. Two LC-PRM runs were carried out for profiling RWE proteome in the PRM library. Three replicates of each sample were initially prepared; however, one replicate of *METTL3*$^{-/-}$ was removed from the analysis due to a contamination concern observed from an abnormal total ion chromatogram. Therefore, only two replicates of *METTL3*$^{-/-}$ cell samples were analyzed.

5.2.4 LC-PRM data acquisition

The setting of the isotope modifications of the PRM library provided in ProteomeXchange Consortium with the dataset identifier PXD030387 was adjusted in Skyline (20) to reflect [$^{13}C_6$, $^{15}N_2$]-Lys and [$^{13}C_6$, $^{15}N_4$]-Arg, which is different from the isotope modifications in stable isotope labelling by amino acids in cell culture (SILAC). Procedures of LC-PRM data acquisition were the same as described elsewhere (19).

5.2.5 LC-PRM data analysis

After raw data were imported to Skyline, PRM traces were manually examined to remove potential interfering fragment ions, which were not overlayed with other fragment ions and had poor mass accuracy (>20 ppm). A dotp value (21) larger than 0.7, and four-to-six fragment ions eluting at the same retention time, was defined as positive identification. Quantification results, including protein name, peptide name, replicate name, isotope, total area, retention time, and library dotp, were exported from Skyline to Excel. In summary, the ratio of each peptide representing a specific RWE protein was calculated based on a two-step normalization: (1) the peak area of an endogenous peptide is normalized to that of its corresponding SIL peptide or a surrogate standard; (2) further normalized to the sum of peak areas for all light peptides over the sum of peak areas for all heavy peptides in each LC-PRM run. The peptide ratio in each sample, averaged from the quantification results of two or three biological replicates, was represented by mean ± S.D. The relative peptide ratio of the peptide in knockout cells vs. HEK 293T cells was further represented by ratio ± propagation of error. The ratio of a specific RWE protein in knockout cells relative to HEK293T cells was represented by the mean ratio from relative peptide ratios ± the new propagation of error. It is worth noting that if multiple peptides were detected from one RWE protein, only the relative peptide ratio with the propagation of error were used to calculate the mean protein ratio and the new propagation of error.

5.2.6 Western blots

ALKBH5$^{-/-}$, *FTO*$^{-/-}$ and *METTL3*$^{-/-}$ cells and the isogenic parental HEK293T cells were harvested, and lysed with CelLytic M reagent supplemented with 1-2 % protease inhibitor cocktail. After centrifugation at 16,100 g for 25 min, Bradford assay was conducted for total protein quantification. Total proteins from each sample were normalized to the same amount prior to denaturation with Laemmli loading buffer for 10 min at 95 °C. The same amount of proteins were separated using an SDS-PAGE gel. A nitrocellulose membrane was used to transfer proteins from samples at 90 V for 1 h at 4 °C. After blocking the membrane with 5% non-fat dry milk in PBS-T (PBS with 0.1% Tween 20) for 40 min, the membrane was cut into pieces based on the apparent molecular weight of each protein of interest according to the product information provided on https://www.ptglab.com/. Each membrane was incubated at 4 °C overnight with the following antibodies: NOP2

(Proteintech, 10448-1-AP, 1:2000), PUS3 (Proteintech, 17248-1-AP, 1:1000), PUS1 (Proteintech, 11512-1-AP, 1:1000), and GAPDH (Santa Cruz, sc-32233, 1:10,000). The membranes were thoroughly washed with PBS-T five times followed by secondary antibody incubation with donkey anti-rabbit secondary antibody (Sigma, A0545, 1:5,000) for NOP2, PUS3, and PUS1, and anti-mouse secondary antibody (Santa Cruz, m-IgGκ BP-HRP, 1:5,000) for GAPDH. After thorough washing for five times, the membranes with protein of interest were visualized using Amersham ECL™ Western Blot Detecting Reagent. Quantification of Western blot was carried out using Image Studio Lite Ver 5.2.

5.2.7 Bioinformatic analysis of m^6A mapping in HEK293T

Four custom tracks from GSE63753 were imported to UCSC genome browser, including GSE63753_hek293.abcam.CIMS.C2T.bedgraph.gz (CIMS C2T Profile); GSE63753_hek293.abcam.CIMS.tag.uniq.bedgraph.gz (CIMS Unique Tag Profile); GSE63753_hek293.sysy.CITS.m6A.12051.bed.gz (CITS m6A); GSE63753_hek293.sysy.CITS.tag.uniq.bedgraph.gz (CITS Unique Tag Profile). From the location of CIMS C2T and CITS m^6A, we were able to identify the m^6A at single-nucleotide resolution. From CIMS and CITS unique tag profiles, we were able to identify m^6A enriched regions.

5.2.8 RNA-bisulfite sequencing

RNA-bisulfite sequencing was conducted according to a published protocol (22) with some modifications. Small RNAs from *ALKBH5*$^{-/-}$, *METTL3*$^{-/-}$, and HEK293T cells were isolated using mirVana miRNA Isolation Kit (Thermo Fisher). Small RNAs were treated with sodium bisulfite strictly following the published protocol (22). Instead of heating the bisulfite converted RNA in the presence of 20 mM $MgCl_2$ and 50 mM Tris-HCl (pH 7) at 75 °C for 15 min as the protocol indicated, first-strand cDNA was synthesized using SuperScript III Reverse Transcriptase (Thermo Fisher) according to the protocol provided by the vender, with one modification that $MgCl_2$ were added at the final concentration of 3 mM (23). PCR was then conducted using ZymoTaq polymerase (Zymo Research) to amplify the bisulfite converted sequences of $tRNA^{Met}$, $tRNA^{Phe}$, and $tRNA^{Gln}$ (Table 5.2). PCR products were separated on a 3% agarose gel followed by gel extraction using GeneJET Gel Extraction Kit (Thermo Fisher).

5.3 Results and Discussion

5.3.1 LC-PRM Analysis Coupled with the Use of SIL Peptides as Internal or Surrogate Standards for Profiling Epitranscriptomic RWE proteins in HEK293T and the Isogenic *ALKBH5*$^{-/-}$, *FTO*$^{-/-}$ and *METTL3*$^{-/-}$ Cells

To explore the potential crosstalk between m^6A and other RNA modifications, we first modified our recently developed LC-PRM method (19) by employing a mixture of 48 stable isotope-labeled (SIL) peptides representing 45 RWE proteins in the PRM library as internal standards or surrogate standards (Table 5.1). We also used this modified method for high-throughput profiling of a total of 152 epitranscriptomic RWE proteins in HEK293T cells, and isogenic cells with the catalytic subunit of the m^6A writer complex (i.e., METTL3) and m^6A eraser proteins (i.e., ALKBH5 and FTO) being genetically ablated (Figure 5.1a and 5.2). Those peptides with the SIL internal standards were quantified based on their peak areas relative to those of their corresponding SIL peptides, whereas those peptides without SIL internal standards were quantified on the basis of their peak areas relative to those of surrogate standards, which were selected based on similar elution times as those of the target peptides.

By using this LC-PRM method coupled with the use of SIL peptides, we were able to quantify the relative expression levels of 117, 119, and 118 RWE proteins in *ALKBH5*, *FTO*, and *METTL3* versus the parental HEK293T cells, which account for approximately 78% of the proteins in the PRM library (Figure 5.1b). A positive peptide identification requires that dot product (dotp) value for its fragment ions observed in MS/MS is larger than 0.7, and that 4-6 transitions share the same retention time. In addition, for those peptides with SIL internal standards, the analytes and heavy isotope-labeled counterparts have to exhibit the same elution time. Figure 5.3 displays the results from hierarchical clustering analysis of the $\log_2$-transformed LC-PRM quantification results for these RWE proteins in *ALKBH5*$^{-/-}$, *FTO*$^{-/-}$ and *METTL3*$^{-/-}$ cells relative to parental HEK293T cells.

5.3.2 LC-PRM analysis Coupled with the Use of SIL Peptides Being Efficient, Robust, Reproducible, and Accurate

The modified LC-PRM method, coupled with the use of SIL peptides, is efficient, robust, reproducible, and accurate. Compared with SILAC, the utilization of SIL peptides obviates the need of metabolic labeling. In addition, the LC-PRM quantification results of each peptide from two or three biological replicates of HEK293T and the isogenic *ALKBH5*, *FTO*, and *METTL3* knockout cells displayed a mean relative standard deviation (RSD) of 12.7% for peptides quantified based on their corresponding SIL internal standards, and 16.1% for peptides quantified based on surrogate standards. These results demonstrate an excellent reproducibility of the method. We also verified the quantification accuracy of this approach by conducting Western blot analyses. We found that the quantification results obtained from LC-PRM and Western blot analyses are consistent for NOP2 and PUS1 proteins (Figure 5.4). On the other hand, our PRM results showed that PUS3 is down-regulated in *ALKBH5*$^{-/-}$ cells over HEK293T cells, whereas Western blot revealed the up-regulation of the protein in the knockout background. This difference may emanate from difference(s) in post-translational modifications of the protein in the two genetic backgrounds, which may affect peptide detection by LC-PRM or antigen recognition by the antibody employed in Western blot analysis, and/or from the lack of adequate specificity of the primary antibody used in Western blot analysis.

5.3.3 Eight Proteins Altered by More Than 1.5-Fold in the Opposite Directions in *ALKBH5*$^{-/-}$ and *METTL3*$^{-/-}$ Cells Relative to Isogenic HEK293T Cells

We next sought to identify potential targets that may be regulated through an m^{6}A-based epitranscriptomic mechanism. RWE proteins with altered expression by over 1.5-fold in individual knockout cell lines (i.e., *ALKBH5*, *FTO*, or *METTL3*) relative to parental HEK293T cells are illustrated in Figure 5.5a. Notably, when compared to parental HEK293T cells, many more proteins exhibit differential expression in *METTL3*$^{-/-}$ than in *ALKBH5*$^{-/-}$ and *FTO*$^{-/-}$ cells. Among these differentially expressed RWE proteins, four (MRM1, PUS3, NOP2, and TGS1) were down-regulated in *ALKBH5*$^{-/-}$ cells with ratios in the knockout (KO) over parental (WT) cells being < 0.67. These results are accompanied with their upregulations (by >1.5-fold) in *METTL3*$^{-/-}$ cells relative to parental HEK293T cells (Figure 5.5b, top panel), suggesting that m^{6}A in the mRNAs of these genes may modulate their decay. Another four RWE proteins (DUS2, TARBP1, NSUN6, and RBMX) were down-regulated in *METTL3*$^{-/-}$ cells with the ratios detected for KO/WT cells being less than 0.67, which is associated with their marked up-regulation (by at least 1.5-fold) in *ALKBH5*$^{-/-}$ cells relative to parental HEK293T cells (Figure 5.5b, top panel). This result indicates that m^{6}A in the mRNAs of these genes may increase the stability and/or translation efficiency of these mRNAs. Together, eight RWE proteins, namely MRM1, PUS3, NOP2, TGS1, DUS2, TARBP1, NSUN6, and RBMX, displayed opposite trends in expression levels in *ALKBH5*$^{-/-}$ and *METTL3*$^{-/-}$ cells relative to the isogenic parental HEK293T cells, suggesting that their corresponding mRNAs may be subjected to regulation via an m^{6}A-mediated epitranscriptomic mechanism.

It is worth noting that several epitranscriptomic RWE proteins exhibited markedly altered (by at least 1.5-fold) expressions in the same direction in *ALKBH5*$^{-/-}$ and *METTL3*$^{-/-}$ cells relative to parental HEK293T cells. This could be attributed to m^{6}A at different sites in mRNA, and/or the involvement of different m^{6}A reader proteins assuming distinct roles in mRNA decay or translation efficiency in *ALKBH5*$^{-/-}$ and *METTL3*$^{-/-}$ cells.

5.3.4 The Presence of m^{6}A in the mRNAs of *NOP2, PUS3, TGS1* and *RBMX*

We next asked if these eight proteins could be regulated through an m^{6}A-based epitranscriptomic mechanism. We began with examining the presence of m^{6}A in the mRNAs of these eight genes using a publicly available dataset (GSE63753) on single-nucleotide resolution mapping of m^{6}A in HEK293 cells (24). The mapping method relied on UV crosslinking between anti-m^{6}A antibody and m^{6}A-modified mRNA and the resulting C$\rightarrow$T mutation at the +1 position of the cross-linked m^{6}A site, or a truncation at the m^{6}A site, induced by reverse transcription (24). As shown in the integrative genomics viewer (IGV) plots, we observed the presence of m^{6}A sites in the mRNAs of *NOP2*, *PUS3*, *TGS1*, and *RBMX* genes in HEK293 cells. Notably, we found m^{6}A sites in the 3′UTR and the last exon of *NOP2* mRNA, and in the last exon, internal exon, and the 3′UTR of *PUS3*, *TGS1*, and *RBMX* mRNAs, respectively (Figure 5.6). The same dataset, nevertheless, did not reveal the presence of m^{6}A in the mRNAs of the other four genes in HEK293 cells; the exact reason is unclear, though we speculate that this could be attributed to the lack of adequate sensitivity of the m^{6}A mapping method.

Since m^{6}A-based epitranscriptomic mechanism could regulate the stabilities and/or translation efficiencies of mRNA (7, 8, 10), we next asked if elevated expression of NOP2, PUS3, and TGS1 proteins in cells depleted of *METTL3* emanates from the enhanced stabilities of these transcripts. To test this, we are in the process of conducting a real-time quantitative PCR (RT-qPCR) analysis to evaluate the half-lives of *NOP2*, *PUS3*, and *TGS1* transcripts in *ALKBH5*$^{-/-}$, *METTL3*$^{-/-}$ cells relative to HEK293T cells, after treatment of cells with actinomycin D to block transcription prior to mRNA extraction.

We also asked if NOP2, PUS3, and TGS1 are regulated by an m^{6}A reader protein YTHDF2 since it is known to stimulate the degradation of mRNA through binding with m^{6}A (8, 10). NOP2 and TGS1, but not PUS3, are targets of YTHDF2, identified using a photoactivatable ribonucleoside crosslinking and immunoprecipitation (PAR-CLIP) (25) approach to locate YTHDF2 binding sites. Different from a well-recognized model that YTHDF1 and YTHDF3 enhance mRNA translation (7, 26, 27) , Zaccara *et al.* (10) introduced a new model that YTHDF1 and YTHDF3 function together with YTHDF2 in mRNA degradation. Hence, apart from YTHDF2, YTHDF1 and YTHDF3 may also regulate the stabilities of *NOP2*, *PUS3*, and *TGS1* transcripts.

5.3.5 The Up-Regulation of the PUS Family Enzymes in *METTL3*$^{-/-}$ Cells

Aside from the aforementioned PUS3, we found that other members in the PUS family, namely PUS1, TRUB1, PUS7, and PUS7L, were up-regulated in *METTL3*$^{-/-}$ cells by 2.34-, 1.36-, 1.23-, and 2.46-fold from the LC-PRM analysis (Figure 5.7a). Interestingly, by analyzing GSE63753, we identified m^{6}A sites from the C-to-T mutational signatures and/or deletion signatures, in the next-to-last exon of *PUS1*, in the last exon of *TRUB1*, in the 3′UTR of *PUS7*, and near the stop codon of *PUS7L* (Figure 5.7b). PUS1 catalyzes the formation of Ψ from uridine at positions 27/28 in the anticodon stem-loop of some tRNAs, and at positions 34/36 in intron-containing tRNAs (28, 29). PUS3, TRUB1, and PUS7 catalyze Ψ formation in some tRNAs at positions 38/39, 55, and 13, respectively (30-33). Not much is known about the function of PUS7L, where a previous study indicated that it may target position 13 and/or 35 in tRNAs (34). These results suggest a broad role of the m^{6}A writer protein (i.e., METTL3) in modulating ψ biosynthesis in human cells.

We next asked if the up-regulation of the PUS family in *METTL3*$^{-/-}$ cells (Figure 5.7a) indeed modulates Ψ synthesis at specific locations of tRNAs. To this end, we assessed the Ψ levels at specific sites in three representative tRNA substrates (tRNA^{Met} for *PUS1* and *TRUB1*; tRNA^{Phe} for *PUS3*, and tRNA^{Gln} for *PUS3*, *TRUB1*, and *PUS7*) (35, 36) in *ALKBH5*$^{-/-}$ and *METTL3*$^{-/-}$ cells compared with HEK293T parental lines. We employed a single-base resolution RNA-bisulfite sequencing method (22), which involves a deletion signature developed during RNA reverse transcription arising from a bisulfite adduct formation to Ψ (23), coupled with Illumina Sequencing for high-throughput profiling of Ψ in three tRNA substrates. We expect to detect a higher percentage of deletion signatures at positions 27/28 and 38/39, representing Ψ sites, in *METTL3*$^{-/-}$ cells compared with HEK293T cells, and a slightly lower percentage in *ALKBH5*$^{-/-}$ cells compared with the parental HEK293T cells.

We also examined those proteins that are regulated in opposite directions by m^6A writer and FTO, another eraser protein of m^6A. The results also illustrated that MRM1 was down-regulated in *FTO*$^{-/-}$ cells by more than 1.5-fold, and up-regulated in *METTL3*$^{-/-}$ cells by over 1.5-fold (Figure 5.5b, middle panel). The expression fold changes of TYW3, ALKBH8, TRUB2, and MRM1 were both up- or down-regulated by over 1.5-fold between *FTO*$^{-/-}$ cells and *ALKBH5*$^{-/-}$ cells, relative to HEK293T cells (Figure 5.5b, bottom panel). These findings suggest that reversible methylation at the N^6 position of adenosine in the mRNAs of these genes, mediated by METTL3 and FTO, may modulate the stabilities and translation efficiencies of these mRNAs.

5.4 Conclusion and Future Work

In summary, we modified our recently developed LC-PRM method by incorporating SIL peptides as internal or surrogate standards. By using this modified targeted proteomic method, we were able to commonly quantify 114 RWE proteins, representing 75% of the RWE proteome in the PRM library, in *ALKBH5*$^{-/-}$, *FTO*$^{-/-}$, *METTL3*$^{-/-}$ cells, and their isogenic parental HEK293T cells. NOP2, PUS3, and TGS1 were up-regulated in *METTL3*$^{-/-}$ cells by over 1.5-fold, and down-regulated in *ALKBH5*$^{-/-}$ cells by at least 0.67-fold, compared with the isogenic parental HEK293T cells. In addition, bioinformatic analysis of an m^6A mapping study revealed the presence of m^6A in mRNA of *NOP2*, *PUS3*, *TGS1*, and *RBMX* genes. It will be important to examine if the up-regulation of NOP2, PUS3, and TGS1, and down-regulation of RBMX in *METTL3*$^{-/-}$ cells arise from altered mRNA stabilities, and/or the binding of m^6A reader proteins, YTHDF1, YTHDF2, and YTHDF3. We are also interested in knowing if a higher percentage of Ψ sites at positions 27/28 and 38/39 in *METTL3*$^{-/-}$ cells, compared with HEK293T cells, will be observed using Illumina Sequencing on three bisulfite-converted tRNA species. The current model of m^6A and Ψ crosstalk is illustrated in Figure 5.8. Together, we revised our LC-PRM method by employing SIL peptides, as internal or surrogate standards. The new method should be amenable for assessing quantitatively the expression levels of epitranscriptomic RWE proteins in tissue samples.

References

1. Perry RP, Kelley DEJC. Existence of Methylated Messenger RNA in Mouse L Cells. 1974;1:37-42.

2. 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;485:201-6.

3. Meyer KD, Saletore Y, Zumbo P, Elemento O, Mason CE, Jaffrey SR. Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. Cell. 2012;149:1635-46.

4. 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 (New York, NY). 1997;3:1233-47.

5. 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;7:885-7.

6. Zheng G, Dahl JA, Niu Y, Fedorcsak P, Huang CM, Li CJ, et al. ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. Mol Cell. 2013;49:18-29.

7. Wang X, Zhao BS, Roundtree IA, Lu Z, Han D, Ma H, et al. N(6)-methyladenosine Modulates Messenger RNA Translation Efficiency. Cell. 2015;161:1388-99.

8. Wang X, Lu Z, Gomez A, Hon GC, Yue Y, Han D, et al. N6-methyladenosine-dependent regulation of messenger RNA stability. Nature. 2014;505:117-20.

9. Xiao W, Adhikari S, Dahal U, Chen YS, Hao YJ, Sun BF, et al. Nuclear m(6)A Reader YTHDC1 Regulates mRNA Splicing. Mol Cell. 2016;61:507-19.

10. Zaccara S, Jaffrey SR. A Unified Model for the Function of YTHDF Proteins in Regulating m(6)A-Modified mRNA. Cell. 2020;181:1582-95.e18.

11. Meyer KD, Patil DP, Zhou J, Zinoviev A, Skabkin MA, Elemento O, et al. 5' UTR m^{6}A promotes cap-independent translation. Cell. 2015;163:999-1010.

12. Choe J, Lin S, Zhang W, Liu Q, Wang L, Ramirez-Moya J, et al. mRNA circularization by METTL3–eIF3h enhances translation and promotes oncogenesis. Nature. 2018;561:556-60.

13. Li Q, Li X, Tang H, Jiang B, Dou Y, Gorospe M, et al. NSUN2-Mediated m5C Methylation and METTL3/METTL14-Mediated m6A Methylation Cooperatively Enhance p21 Translation. J Cell Biochem. 2017;118:2587-98.

14. Dai X, Gonzalez G, Li L, Li J, You C, Miao W, et al. YTHDF2 Binds to 5-Methylcytosine in RNA and Modulates the Maturation of Ribosomal RNA. Anal Chem. 2020;92:1346-54.

15. Dai X, Wang T, Gonzalez G, Wang Y. Identification of YTH Domain-Containing Proteins as the Readers for N1-Methyladenosine in RNA. Anal Chem. 2018;90:6380-4.

16. Wei J, Liu F, Lu Z, Fei Q, Ai Y, He PC, et al. Differential m(6)A, m(6)A(m), and m(1)A Demethylation Mediated by FTO in the Cell Nucleus and Cytoplasm. Mol Cell. 2018;71:973-85.e5.

17. Yang Y-Y, Yu K, Li L, Huang M, Wang Y. Proteome-wide Interrogation of Small GTPases Regulated by N6-Methyladenosine Modulators. Anal Chem. 2020;92:10145-52.

18. Wisniewski JR, Zougman A, Nagaraj N, Mann M. Universal sample preparation method for proteome analysis. Nat Meth. 2009;6:359-62.

19. Qi TF, Miao W, Wang Y. Targeted Profiling of Epitranscriptomic Reader, Writer, and Eraser Proteins Accompanied with Radioresistance in Breast Cancer Cells. Anal Chem. 2022;94:1525-30.

20. 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 (Oxford, England). 2010;26:966-8.

21. Kawahara R, Bollinger JG, Rivera C, Ribeiro ACP, Brandão TB, Paes Leme AF, et al. A targeted proteomic strategy for the measurement of oral cancer candidate biomarkers in human saliva. Proteomics. 2016;16:159-73.

22. Khoddami V, Yerra A, Mosbruger TL, Fleming AM, Burrows CJ, Cairns BR. Transcriptome-wide profiling of multiple RNA modifications simultaneously at single-base resolution. Proc Natl Acad Sci U S A. 2019;116:6784-9.

23. Fleming AM, Alenko A, Kitt JP, Orendt AM, Flynn PF, Harris JM, et al. Structural Elucidation of Bisulfite Adducts to Pseudouridine That Result in Deletion Signatures during Reverse Transcription of RNA. J Am Chem Soc. 2019;141:16450-60.

24. Linder B, Grozhik AV, Olarerin-George AO, Meydan C, Mason CE, Jaffrey SR. Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. Nat Methods. 2015;12:767-72.

25. Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, et al. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. Cell. 2010;141:129-41.

26. Li A, Chen YS, Ping XL, Yang X, Xiao W, Yang Y, et al. Cytoplasmic m^6A reader YTHDF3 promotes mRNA translation. Cell Res. 2017;27:444-7.

27. Shi H, Wang X, Lu Z, Zhao BS, Ma H, Hsu PJ, et al. YTHDF3 facilitates translation and decay of N(6)-methyladenosine-modified RNA. Cell Res. 2017;27:315-28.

28. Chen J, Patton JR. Cloning and characterization of a mammalian pseudouridine synthase. RNA (New York, NY). 1999;5:409-19.

29. Motorin Y, Keith G, Simon C, Foiret D, Simos G, Hurt E, et al. The yeast tRNA:pseudouridine synthase Pus1p displays a multisite substrate specificity. RNA (New York, NY). 1998;4:856-69.

30. Lecointe F, Simos G, Sauer A, Hurt EC, Motorin Y, Grosjean H. Characterization of yeast protein Deg1 as pseudouridine synthase (Pus3) catalyzing the formation of psi 38 and psi 39 in tRNA anticodon loop. J Biol Chem. 1998;273:1316-23.

31. Chen J, Patton JR. Pseudouridine synthase 3 from mouse modifies the anticodon loop of tRNA. Biochemistry. 2000;39:12723-30.

32. Zhang W, Pan T. A dual function PUS enzyme. Nat Chem Biol. 2020;16:107-8.

33. Kaya Y, Ofengand J. A novel unanticipated type of pseudouridine synthase with homologs in bacteria, archaea, and eukarya. RNA (New York, NY). 2003;9:711-21.

34. de Crécy-Lagard V, Boccaletto P, Mangleburg CG, Sharma P, Lowe TM, Leidel SA, et al. Matching tRNA modifications in humans to their known and predicted enzymes. Nucleic Acids Res. 2019;47:2143-59.

35. Suzuki T, Suzuki T. A complete landscape of post-transcriptional modifications in mammalian mitochondrial tRNAs. Nucleic Acids Res. 2014;42:7346-57.

36. de Brouwer APM, Abou Jamra R, Körtel N, Soyris C, Polla DL, Safra M, et al. Variants in PUS7 Cause Intellectual Disability with Speech Delay, Microcephaly, Short Stature, and Aggressive Behavior. Am J Hum Genet. 2018;103:1045-52.

Figure 5.1 (a) Workflow of LC-PRM analysis coupled with the use of SIL peptides as internal or surrogate standards for profiling epitranscriptomic RWE proteins in HEK293T and the isogenic *ALKBH5*$^{-/-}$, *FTO*$^{-/-}$ and *METTL3*$^{-/-}$ cells. (b) A Venn diagram depicting the numbers of quantified RWE proteins in HEK293T, *ALKBH5*$^{-/-}$, *FTO*$^{-/-}$ and *METTL3*$^{-/-}$ cells, compared with those deposited in the PRM library.

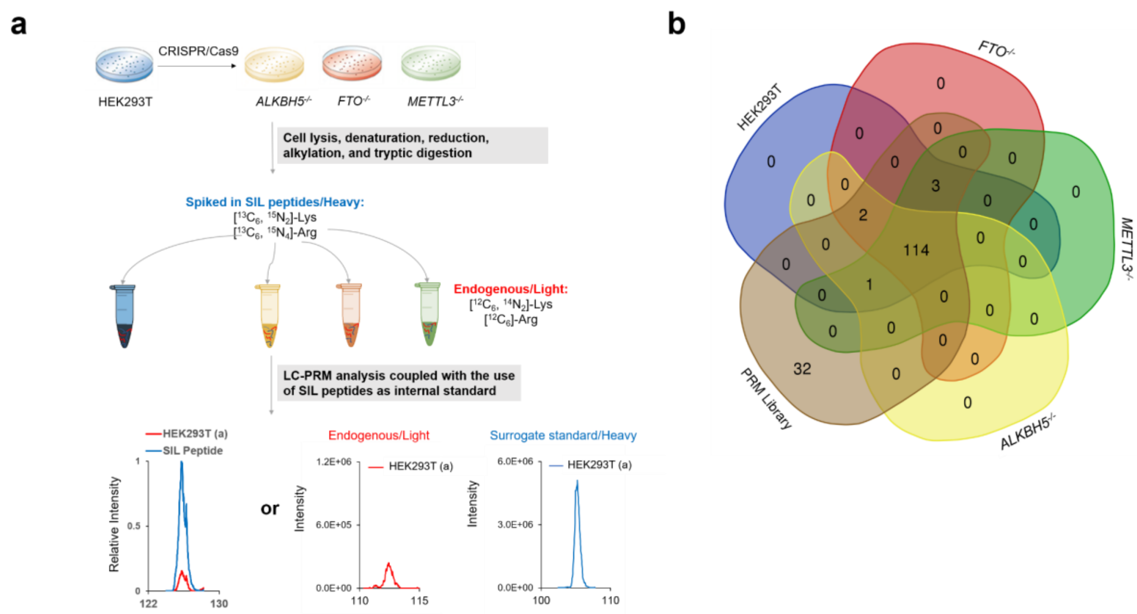


Figure 5.2 PRM traces showing the evidence of successful knockout of *ALKBH5*, *METTL3*, and *FTO* in HEK293T.

Dotp values are 0.3, 0.33, and 0.35 of VSEPVLSLPVR from ALKBH5 in *ALKBH5*$^{-/-}$, NPEAALSPTFR from METTL3 in *METTL3*$^{-/-}$, and ILIGNPGCTYK from FTO in *FTO*$^{-/-}$ respectively, compared with 0.96, 0.99, 0.9 in their isogenic HEK293T cells. The SIL peptide of ILIGNPGCTYK was not detected.

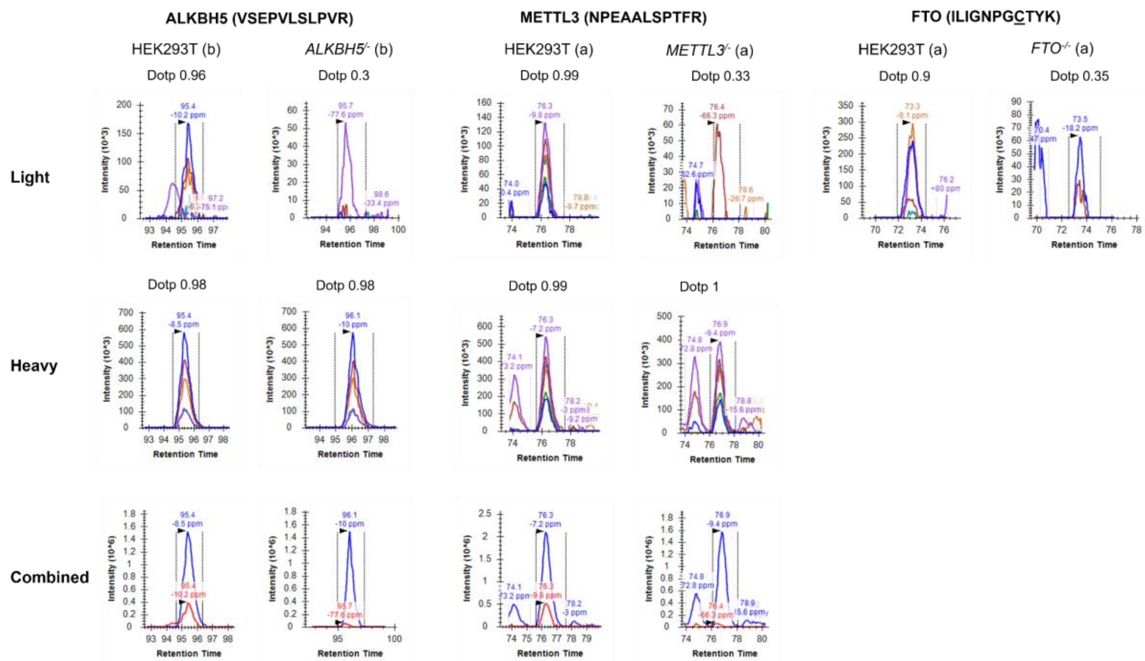


Figure 5.3 Hierarchical clustering illustrating the $\log_2$-transformed expression ratios of RWE proteins in *ALKBH5*$^{-/-}$ (n = 3), *FTO*$^{-/-}$ (n = 3) and *METTL3*$^{-/-}$ (n = 2) cells relative to parental HEK293T cells (n = 3).

Hierarchical clustering was generated using Perseus, where red and blue boxes designate up- and down-regulation of the RWE protein in the knockout cells compared to HEK293T cells. White boxes illustrate no changes in expression between knockout and HEK293T cells. Gray boxes indicate missing data.

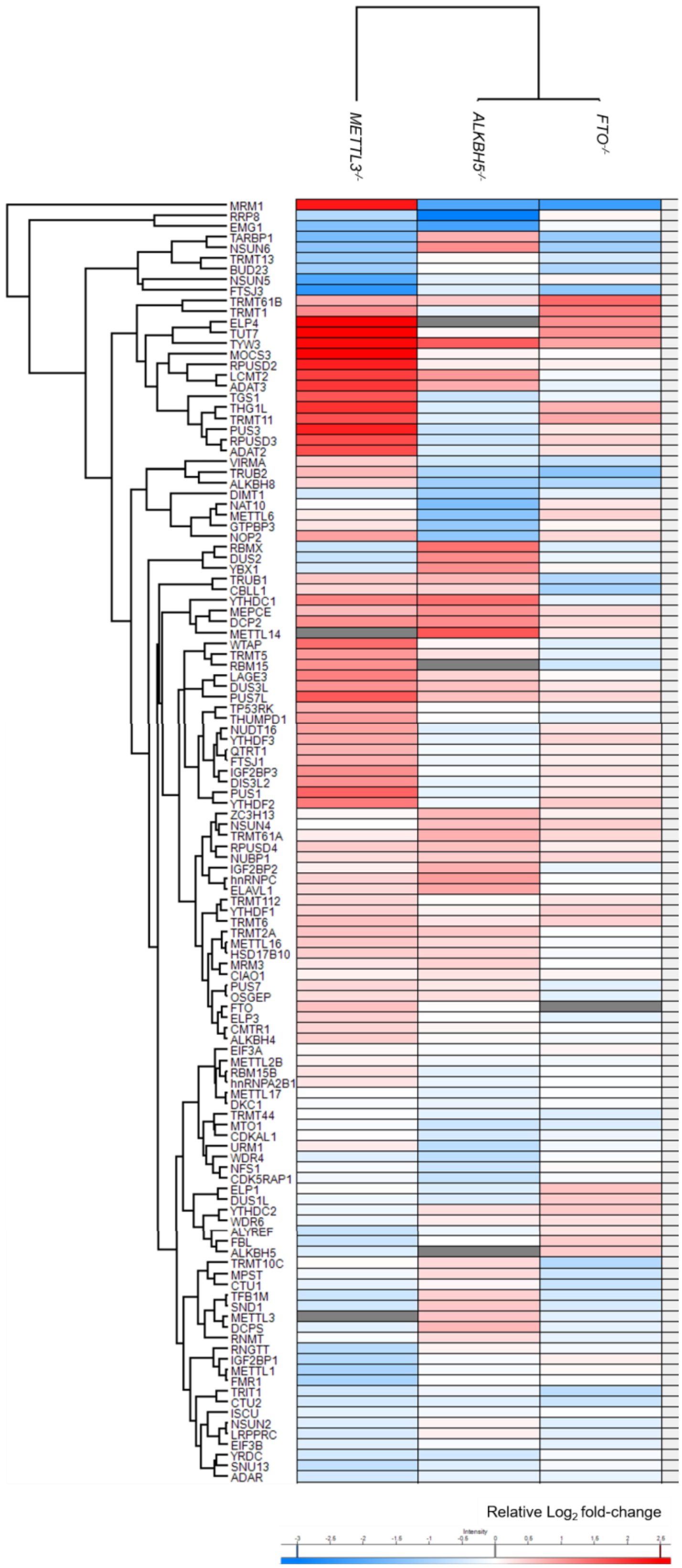

METTL3-/-
ALKBH5-/-
FTO-/-
MRM1
RRP8
EMG1
TARBP1
NSUN6
TRMT13
BUD23
NSUN5
FTSJ3
TRMT61B
TRMT1
ELP4
TUT7
TYW3
MOCS3
RPUSD2
LCMT2
ADAT3
TGS1
THG1L
TRMT11
PUS3
RPUSD3
ADAT2
VIRMA
TRUB2
ALKBH8
DIMT1
NAT10
METTL6
GTPBP3
NOP2
RBMX
DUS2
YBX1
TRUB1
CBLL1
YTHDC1
MEPCE
DCP2
METTL14
WTAP
TRMT5
RBM15
LAGE3
DUS3L
PUS7L
TP53RK
THUMPD1
NUDT16
YTHDF3
QTRT1
FTSJ1
IGF2BP3
DIS3L2
PUS1
YTHDF2
ZC3H13
NSUN4
TRMT61A
RPUSD4
NUBP1
IGF2BP2
hnRNPC
ELAVL1
TRMT112
YTHDF1
TRMT6
TRMT2A
METTL16
HSD17B10
MRM3
CIAO1
PUS7
OSGEP
FTO
ELP3
CMTR1
ALKBH4
EIF3A
METTL2B
RBM15B
hnRNPA2B1
METTL17
DKC1
TRMT44
MTO1
CDKAL1
URM1
WDR4
NFS1
CDK5RAP1
ELP1
DUS1L
YTHDC2
WDR6
ALYREF
FBL
ALKBH5
TRMT10C
MPST
CTU1
TFB1M
SND1
METTL3
DCPS
RNMT
RNGTT
IGF2BP1
METTL1
FMR1
TRIT1
CTU2
ISCU
NSUN2
LRPPRC
EIF3B
YRDC
SNU13
ADAR
Relative Log2 fold-change
Intensity
-3
-2.5
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5

Figure 5.4 Western blot analyses for validating the protein expression levels of PUS1, NOP2, and PUS3.

Extracted-ion chromatograms of a representative peptide TIEDDLVSALVR from PUS1 and its corresponding SIL peptide (a) LGVTNTIISHYDGR from NOP2 and its surrogate standard AATACFGFPK (b), a representative peptide ILAWAPVEPSFSAR from PUS3 and its surrogate standard GFAFVQYVNER (c). Shown on the right are the Western blot results (n =3) of PUS1 (a), NOP2 (b) and PUS3 (c) proteins in $\mathit{ALKBH5}^{-/-}$, $\mathit{FTO}^{-/-}$ and $\mathit{METTL3}^{-/-}$ cells relative to parental HEK293T cells.

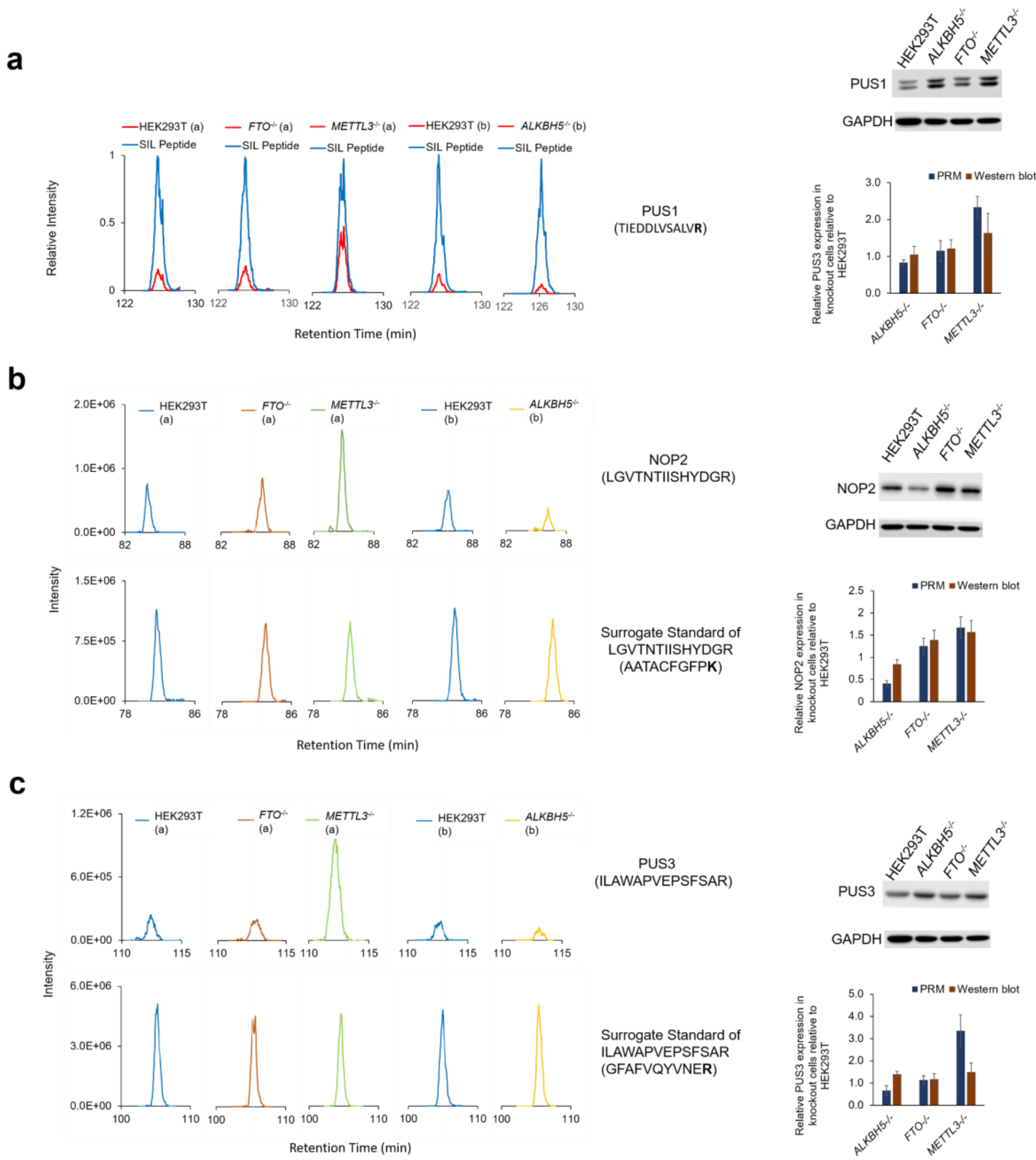


Figure 5.5 (a) LC-PRM quantification results of RWE proteins in *ALKBH5*$^{-/-}$ (n = 3), *FTO*$^{-/-}$ (n = 3) and *METTL3*$^{-/-}$ (n = 2) cells relative to parental HEK293T cells (n = 3).

Only proteins with ratios in knockout/parental cells being > 1.5 or < 0.67 are displayed. The ratio of each peptide representing a specific RWE protein was determined following the procedures described in Materials and Methods in the Supporting Information. (b) Scatter plots depicting the LC-PRM quantification result of RWE proteins in one knockout over HEK293T cells vs. another knockout over HEK293T cells. The expression fold change of those RWE proteins altered over 1.5-fold in both the knockout cells and HEK293T cells are labeled in red.

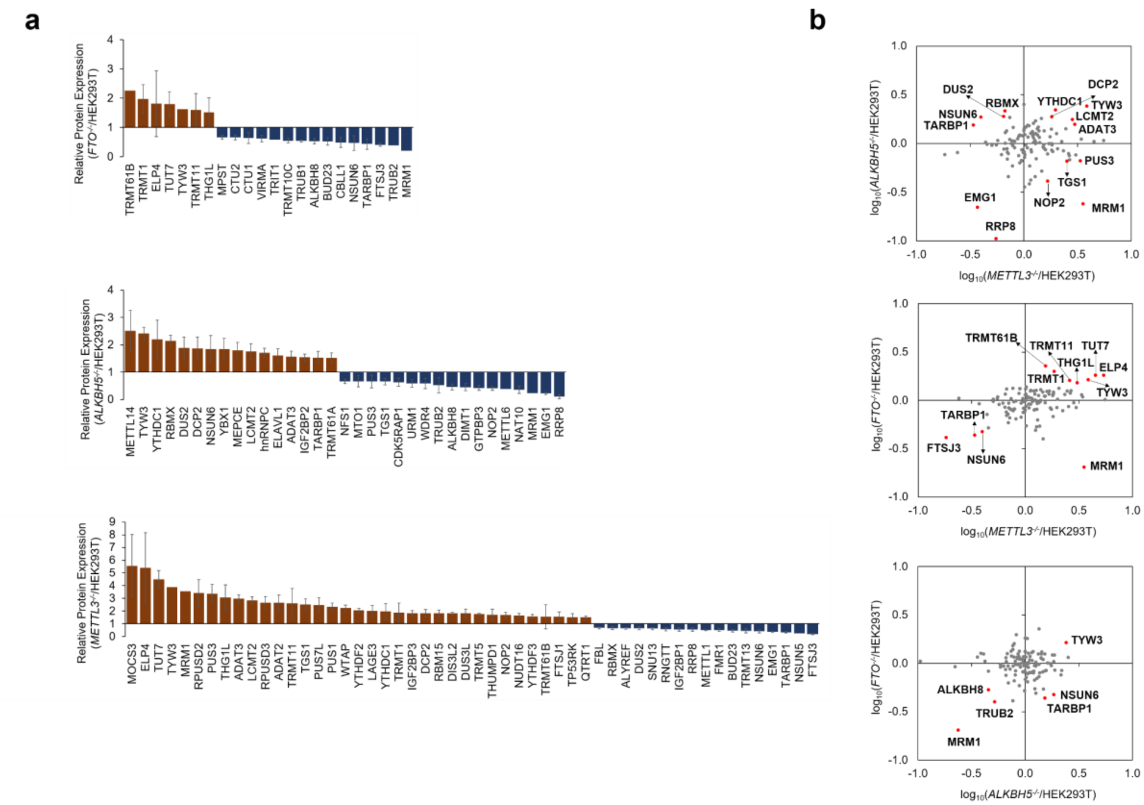


Figure 5.6 Bioinformatic analysis depicting m^{6}A mapping results for in NOP2 (a), PUS3 (b), TGS1 (c), RBMX (d) mRNAs in GSE63753.

The lower panel is the full view of the gene. The upper panel shows the zoom-in view of the marked signal to check genomic sequences. The potential m^{6}A sequence motif is highlighted in red boxes. For those genes located on the reverse strand, the converted complementary sequence was revealed on top, together with the identification of m^{6}A site labeled as a red circle beneath determined based on the location of CIMS C→T mutational signature and/or CITS deletion signature.

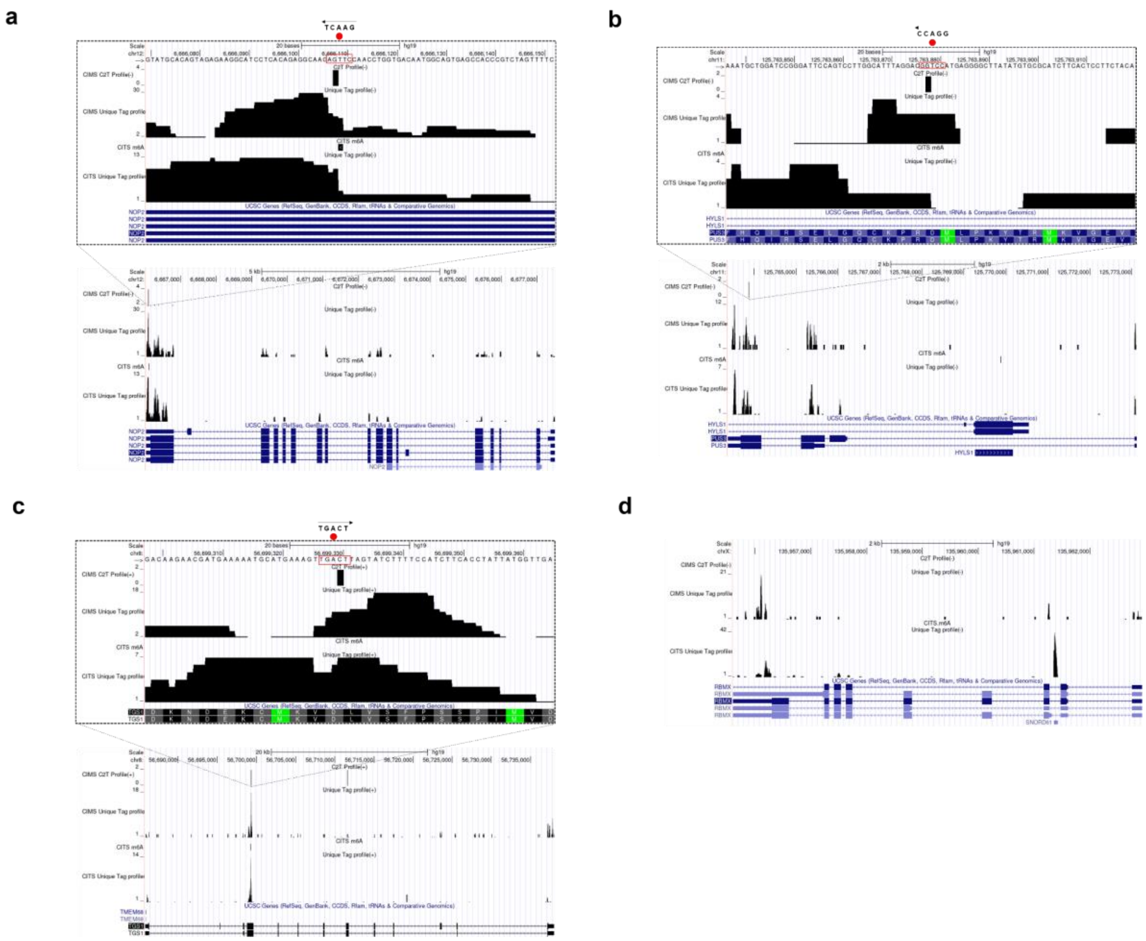


Figure 5.7 LC-PRM quantification results and Bioinformatics of m^6A mapping of the PUS family enzymes.

(a) LC-PRM quantification results of PUS1, PUS3, TRUB1, PUS7, and PUS7L in *ALKBH5*$^{-/-}$, *FTO*$^{-/-}$ and *METTL3*$^{-/-}$ cells, compared with the isogenic parental HEK293T cells. (b) Bioinformatics of m^6A mapping in *PUS1*, *TRUB1*, *PUS7*, and *PUS7L* genes in GSE63753. Additional description is shown in the legend of Figure S2.

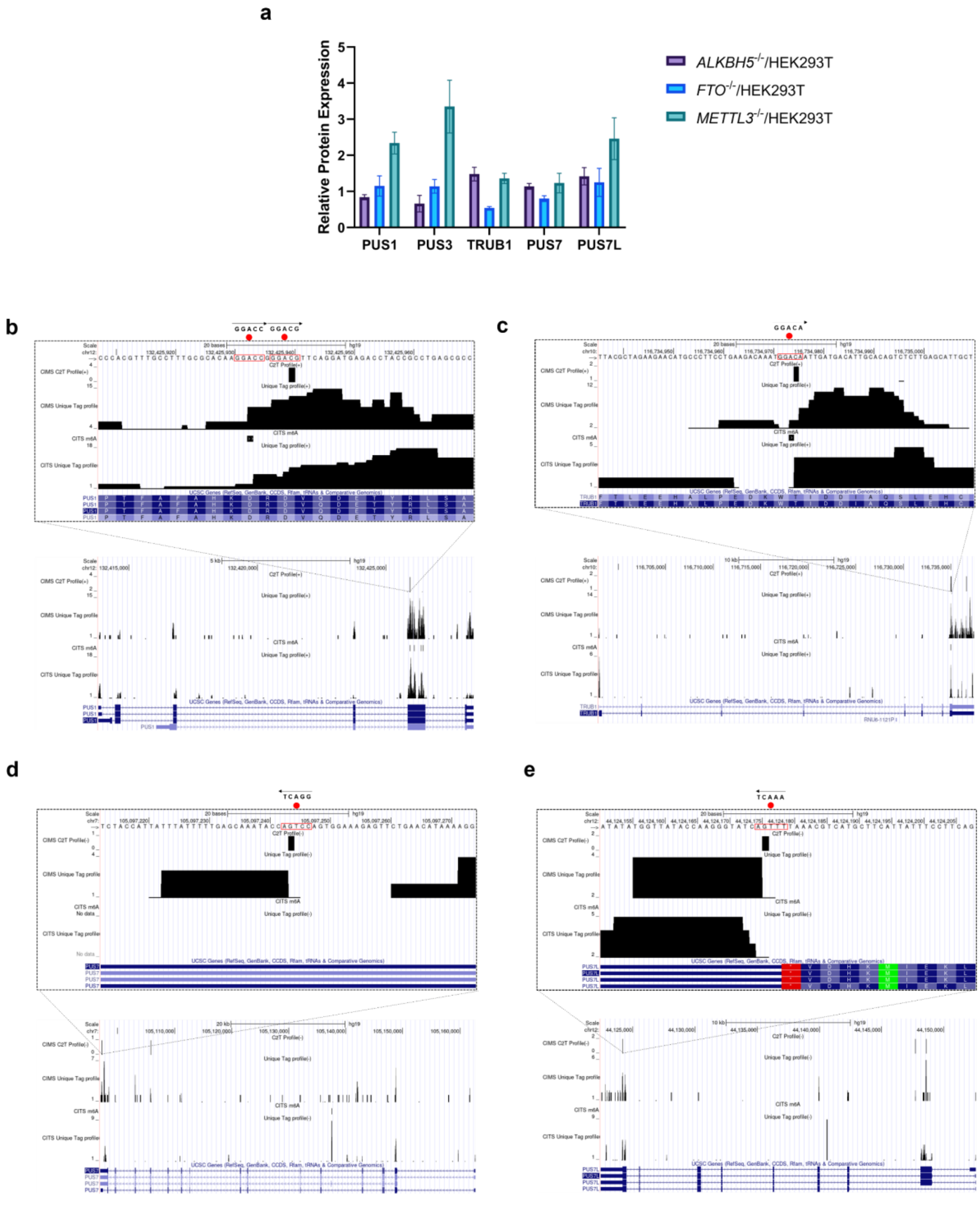

a
Relative Protein Expression
ALKBH5-/-/HEK293T
FTO-/-/HEK293T
METTL3-/-/HEK293T
PUS1
PUS3
TRUB1
PUS7
PUS7L
b
GGACC GGACG
c
GGACA
d
TCAGG
e
TCAAA

Figure 5.8 Current model of m^6A and Ψ crosstalk.

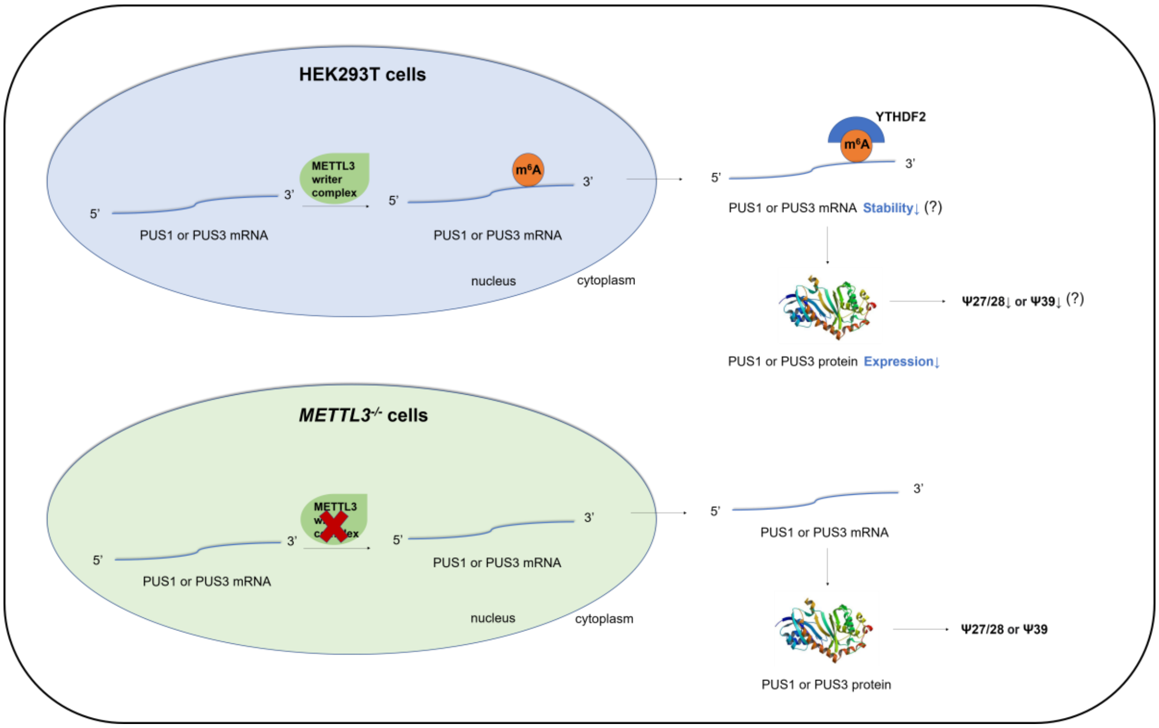


Table 5.1 A list of SIL peptides used in this study. Heavy isotope-labeled amino acid is marked in bold.

Protein	Peptide	Protein	Peptide
ALKBH1	AEAGILNYY**R**	IGF2BP1	ITISSLQDLTLYNPE**R**
ALKBH4	ELSAEFGPGG**R**	IGF2BP2	LYIGNLSPAVTADDL**R**
ALKBH5	VSEPVLSLPV**R**	IGF2BP3	ITISPLQELTLYNPE**R**
ALKBH8	SGIITSDVGDLTLS**K**	WTAP	QVQQPSVAQL**R**
FTO	ILIGNPGCTY**K**	VIRMA	FP**<u>C</u>**VVYINEV**R**
YTHDF1	SPVDYGTSAGVWSQD**K**	RBM15	LQQLALG**R**
YTHDF2	DGLNDDDFEPYLSPQA**R**	EIF3A	VLLATLSIPITPE**R**
YTHDF3	AITDGQAGFGNDTLS**K**	EIF3B	GTQGVVTNFEIF**R**
YTHDF3	GNVGIGGSAVPPPPI**K**	hnRNPA2B1	IDTIEIITD**R**
YTHDF3	VPGISSIEQGMTGL**K**	hnRNPC	GFAFVQYVNE**R**
YTHDC1	GVWSTLPVNE**K**	PUS1	TIEDDLVSALV**R**
YTHDC2	VVLIVGETGSG**K**	TRUB1	AAAAVVAAAA**R**
METTL3	NPEAALSPTF**R**	DKC1	EVVAEVV**K**
METTL14	LEIDEIAAP**R**	ALYREF	QQLSAEELDAQLDAYNA**R**
METTL1	AAPAGGFQNIA**<u>C</u>**L**R**	NSUN4	VLVDVP**<u>C</u>**TTD**R**
METTL16	EDFGLSIDIPLE**R**	ALKBH2	EVEYFTGALA**R**
METTL2B	TQTPPVEENVTQ**K**	ALKBH3	EDITYQQP**R**
NSUN2	FYALDPSFP**R**	ALKBH6	VPALEPF**R**
NSUN6	EVASYQPLQ**R**	ALKBH7	DEESFFGE**R**
DNMT1	FFLLENV**R**	METTL6	QNPLYDTE**R**
TRMT61A	T**<u>C</u>**QALAA**R**	NSUN5	YSAVLDAVIASAGLL**R**
TRMT61B	DISGATEDI**K**	TRDMT	YAMDVEN**K**
TRMT6	AATA**<u>C</u>**FGFP**K**	RBM15B	NLDADLV**R**
PUS7	FGTTAVPTYQVG**R**	ZCH3H13	LISDSVE**R**

Table 5.2 PCR primers for RNA bisulfite sequencing.

Description	Sequence
Bisulfite converted $tRNA^{Met}$ PCR Forward Primer	5'-GTT GAA ATA GTT TAG TTG GGA-3'
Bisulfite converted $tRNA^{Met}$ PCR Reverse Primer	5'-TAC CAA AAC CCA AAA TCA-3'
Bisulfite converted $tRNA^{Gln}$ PCR Forward Primer	5'-GGT TTT ATG GTG TAA TGG-3'
Bisulfite converted $tRNA^{Gln}$ PCR Reverse Primer	5'-AAA TCC CAC CAA AAT TTA-3'
Bisulfite converted $tRNA^{Met}$ PCR Forward Primer	5'-GTT TTG TTA GTG TAG TAG G-3'
Bisulfite converted $tRNA^{Met}$ PCR Reverse Primer	5'-TAA TAC CCC ATA TAA AAA TCA-3'

6. Chapter 6. Concluding Remarks

In this dissertation, I described two research projects, namely untargeted quantitative profiling of paired primary/metastatic melanoma cells, with the focus being placed on TBC domain-containing proteins, and targeted quantitative profiling of epitranscriptomic RWE proteins.

In Chapter 2, an off-line SCX fractionation prior to LC-MS/MS enables the quantification of a total of 7387 proteins in the matched pair of primary/metastatic melanoma cells (i.e., WM-115/WM-266-4). Among them, 1551 (21%) proteins display at least 1.5-fold differences between these two cell lines. I placed my emphasis on studying one subfamily of small GTPase regulatory proteins, i.e., TBC domain-containing proteins, since some members of this family, including TBC1D3, TBC1D8 and TBC1D16, are known to assume important functions in tumorigenesis or tumor progression. Our proteomic data enabled the quantification of 24 TBC proteins, accounting for 55% of predicted members of the TBC protein subfamily. More importantly, a role of TBC1D7 in melanoma cell invasion was unveiled from the proteomic data, together with follow-up cell-based assays and bioinformatic analysis of publicly available data of melanoma patients and cell lines. Moreover, the quantitative proteomic results provided an important basis for investigating the roles of other proteins, especially the ones with significant fold changes and clinical relevance, in modulating melanoma metastasis in the future.

The formation of m^6A, the most abundant internal RNA modification in mRNA, is reversible. m^6A can be deposited by m^6A writer complex, METTL3-METTL14 heterodimer and other subunits, and removed by eraser proteins, ALKBH5 and FTO. m^6A

readers translate m^6A into function, namely mRNA splicing, export, stability, and translation. To our knowledge, there is no systematic study about how epitranscriptomic RWE proteins modulate cancer radioresistance and metastasis. Toward this end, I developed an LC-PRM method for high-throughput profiling of a total of 152 epitranscriptomic RWE proteins, modulators of m^6A and other ribonucleoside modifications.

In Chapter 3, I employed this LC-PRM method coupled with SILAC to assess the differential expression of epitranscriptomic RWE proteins in two matched pairs of radioresistant/parental breast cancer cells (i.e., MDA-MB-231/C5 and MCF-7/C6). 65% and 70% of the epitranscriptomic RWE proteome was quantified in the matched pair. TRMT1 (an $m^{2,2}G$ writer) may assume a role in promoting breast cancer radioresistance because of its clinical relevance and its correlation with DNA repair gene sets.

In Chapter 4, I further applied this LC-PRM method coupled with SILAC to explore the differences in expression levels of epitranscriptomic RWE proteins in paired metastatic/primary CRC cells (i.e., SW620/SW480). I was able to quantify 74% of the epitranscriptomic RWE proteome in the PRM library. Among these quantified proteins, 48 and 5 were up- and down-regulated by larger than 1.5-fold in SW620 over SW480 cells, respectively. The roles of some of those proteins with marked up-regulation in metastatic CRC cells, including NAT10, hnRNPC, and DKC1, in the metastasis of CRC and other cancer were demonstrated, which validated our proteomic findings.

In Chapter 5, I modified this LC-PRM method by combining the use of a mixture of 48 SIL peptides representing RWE proteins as internal or surrogate standards. I utilized this method to investigate potential crosstalk between m^6A and other modified nucleotides in RNA by exploring the differential expression levels of RWE proteins in *ALKBH5*$^{-/-}$, *FTO*$^{-/-}$, *METTL3*$^{-/-}$ cells, and their isogenic parental HEK293T. Approximately 78% of the epitranscriptomic RWE proteome in the PRM library was quantified. I also found that the expression levels of the Ψ synthases, i.e., PUS1 and PUS3, were up-regulated in *METTL3*$^{-/-}$ cells and down-regulated in *ALKBH5*$^{-/-}$ relative to parental HEK293T cells. Bioinformatic analysis of published m^6A mapping results revealed the presence of m^6A in the mRNAs of PUS1 and PUS3. It will be important to investigate if the up- and down-regulations of PUS1 and PUS3 in *METTL3*$^{-/-}$ cells and *ALKBH5*$^{-/-}$ cells emanate from an m^6A-based epitranscriptomic mechanism. In addition, it will be necessary to interrogate the levels of Ψ at sites 27/28 and 39, regulated by PUS1 and PUS3, respectively, in *ALKBH5*$^{-/-}$ and *METTL3*$^{-/-}$ cells relative to HEK293T cells. This study may ultimately lead to the discovery of a crosstalk between m^6A and Ψ.

Together, in this dissertation, I presented a novel targeted quantitative proteomic method for profiling epitranscriptomic RWE proteins. This proteomic method was coupled with two labeling methods individually, namely, SILAC and SIL peptides, both of which were found to provide highly efficient, sensitive, accurate, and reproducible quantification of epitranscriptomic RWE proteins. This LC-PRM method allowed for the investigation of the roles of epitranscriptomic RWE proteins in breast cancer radioresistance (Chapter 3), CRC metastasis (Chapter 4), and understanding new regulatory mechanisms of the epitranscriptome, namely, a potential m^6A-Ψ crosstalk (Chapter 5). We envision that the

LC-PRM method coupled with SIL peptides will be amenable for high-throughput profiling of epitranscriptomic RWE proteins in biofluid and tissues.