UNIVERSITY OF CALIFORNIA
RIVERSIDE

Interrogation of Small GTPase Proteome and METTL3 Interactome

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

Doctor of Philosophy

in

Chemistry

by

Yen-Yu Yang

March 2023

Dissertation Committee:
Dr. Yinsheng Wang, Chairperson
Dr. Joseph Genereux
Dr. Haofei Zhang

The Dissertation of Yen-Yu Yang is approved:

Committee Chairperson

University of California, Riverside

ACKNOWLEDGEMENTS

Ph.D. journeys are full of trial and hardship, but ultimately rewarding. My personal journey would not have been possible without the research laboratory that gave me the opportunity to pursue my Ph.D. studies and the amazing support system from my family and friends. I would like to express my gratitude to both for getting me here today.

I would also like to express my deepest appreciation and gratitude to my doctoral advisor Dr. Yinsheng Wang for opening the doors of mass spectrometry, analytical chemistry and cellular biology for me. His passion for science in general and expertise in mass spectrometry-based proteomic analysis was contagious, and his guidance and mentorship throughout my Ph.D. studies have deeply instilled in me a love for proteomics. His understanding and encouragement were the source of my continued resilience, even during the most difficult times. He encouraged me to embrace unexpected findings with critical interpretations and be confident in front of my own results. He also showed me the importance of reading literature, asking thought-provoking questions, and being thoughtful about experimental design. Moreover, he always provided timely and insightful feedback and constructive criticism on the experimental results, which were instrumental in improving the quality of my work, as demonstrated in this dissertation. In addition, he also revised my manuscripts, abstracts for conferences, and applications for scholarships and awards thoroughly and thoughtfully, which taught me effective time management skills and had a great impact on my productivity. Lastly, I would also like to extend my sincere thanks to him for encouraging and supporting me to apply for several awards and to attend

many excellent conferences, which have provided me with new perspectives, deeper understanding of mass spectrometry, and valuable connections to colleagues. From him, I developed so much, not the least of which being a strong work ethic and a commitment to excellence. I will always cherish the impact he has had on my scientific and personal growth.

I would also like to express my gratitude to the "mass spec folks" in the lab. Special thanks to Dr. Yuxiang Cui, Dr. Gwendolyn Gonzales, and Dr. Ming Huang, all of whom supported and trained me on the instruments with patience and understanding. Their expertise and experience in mass spectrometry and liquid chromatography have been invaluable in helping me get familiar with the instruments crucial for proteomic analysis. I would also like to thank Jiekai Yin, Dr. Tianyu Qi, Dr. Xuejiao Dong, Dr. Quanqiang Zhang and my successor Zhongwen Cao for their unwavering support in the day-to-day maintenance of instruments and troubleshooting. The mass spec folks in Wang lab have been such a great team to work with. I hope that we can continue our journey of supporting each other to success together as we move forward in our careers.

I would like to extend my thanks to all other current and previous members in the Wang lab. Special thanks to Dr. Chengjie Ma, Jun Yuan, Dr. Drew Kellum, Dr. David Bade, Dr. Yuxiang Sun, Dr. Xiaomei He, Dr. Xiaochuan Liu, Dr. Rong Cai and everyone in the above-mentioned mass spec team for bringing positive energy and a sense of humor to the lab. Your friendly nature and support created a real difference in my Ph.D. journey and have made those challenging days enjoyable. Research-wise, your generosity and willingness to share your knowledge have helped me to complete my dissertation studies

and to develop a deeper understanding in biology. I am thankful for your knowledge, your guidance, and your inspiration, which has helped my Ph.D. journey in so many ways.

I would like to express my gratitude to my committees and many professors in the department. I am grateful for Dr. Joseph Genereux who taught me chemical biology and provided constructive feedback and critical questions on my dissertation research throughout my Ph.D. studies. I would also like to thank Dr. Wenwan Zhong and Dr. Ryan Julian who taught me separation sciences and mass spectrometry, respectively. Coming from a non-chemistry background, I was intimidated by analytical chemistry knowledge and techniques. But they have been supportive and knowledgeable professors who taught me from scratch and helped me accumulate knowledge in these subject matters in such a short period of time. Your patience and expertise helped me build a strong foundation during the first year of my Ph.D. studies and subsequently a strong passion for the field.

I would also like to acknowledge my collaborators Ruthia Soh, Madeline Vera-Colón, and Dr. Nicole zur Nieden. Thank you for your collaboration on osteogenesis research and providing me with the materials to make our project a success. Many thanks to Dr. Nicole zur Nieden for her feedback and guidance on the osteogenesis project, which laid important foundations on the work. I am so grateful to Ruthia Soh and Dr. Nicole zur Nieden for training me on culturing and differentiating stem cells. It was a wonderful learning experience.

Next, I would like to thank my family and my friends for being extremely supportive throughout my Ph.D. studies outside of the research laboratory. Special thanks to Brendan for endless kindness and understanding, which helped me to navigate emotional

challenges I have faced over years in my PhD studies. I appreciate his willingness to listen to my difficulties and offer guidance with practical solutions. I also appreciate the time we learned Python together and developed new skills during the pandemic. In line with that, I would also like to thank my two cats, Mr. Kitty and Noodles who have been a source of comfort and joy after long days in the laboratory, especially those days when the experiment did not work out as expected. Your furry companionship helped me to manage stress and anxiety throughout my Ph.D. journey.

My appreciation should also go to my friends Nina, Frank, and Roger, of whom we shared experiences, both good and bad together, and helped one another to feel not as isolated. We encouraged each other, chatted about scientific research, and spent quality time together. Your friendship has brought so much light and laughter into my Ph.D. journey. I would also like to thank the Yang and the Brown family. Seeing you on the holidays has been very meaningful to me and I am grateful for the memories we created together.

I would like to mention several institutes and foundations that have helped me during my Ph.D. years. Special thanks to the Department of Chemistry at UC Riverside for providing such a great learning environment to grow my professional credentials in so many different ways. I would like to thank Christina Youhas and Jeffery Matute for all their assistance with class enrollment and reimbursements. Their attention to details and timely processing has made the journey much smoother and less stressful. I would also like to thank the YinChin foundation for their financial support when I first moved to the US to start my Ph.D. studies. I also appreciate the Department of Chemistry at UC Riverside

along with the Chinese American Chromatography Association for recognizing my work and providing me awards to help further my research. I am also very grateful to the University of California for providing me an opportunity to attend the Lindau Nobel Laureate Conference, which profoundly inspired me towards my scientific career development.

Lastly, I would like to acknowledge my appreciation for my grandparents and their unconditional love, raising me, and teaching me important values of life. I am extremely grateful for your presence in my life and the memories that we share together forever.

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 1: Yang YY., Huang M., and Wang Y. Targeted Proteomic Analysis of Small GTPases in Murine Adipogenesis. *Analytical Chemistry*, 2020, 92, 9, 6756-676

Chapter 3: Yang YY., Yu K., Li L., Huang M., and Wang Y. Proteome-Wide Interrogation of Small GTPases Regulated by N^6 -Methyladenosine Modulators. *Analytical Chemistry*, 2020, 92, 14, 10145-10152.

DEDICATION

To my grandparents, Tsu-Fang Yang and Su-Hsia Kung,
who support me with unconditional love.

ABSTRACT OF THE DISSERTATION

Interrogation of Small GTPase Proteome and METTL3 Interactome

by

Yen-Yu Yang

Doctor of Philosophy, Graduate Program in Chemistry
University of California, Riverside, March 2023
Dr. Yinsheng Wang, Chairperson

Recent advances in proteomics, including instrumentation, data acquisition approaches, and sample preparation methods, allow us to investigate the proteome of interest with high sensitivity, selectivity, and throughput. The focus of this dissertation is placed on the applications of modern proteomics to systematically interrogate an important subset of the proteome. The research reported in this dissertation is divided into two parts, i.e., applying multiple-reaction monitoring (MRM)-based targeted proteomics to systematically quantify the expression of the entire small GTPase superfamily; and employing ascorbate peroxidase (APEX)-based proximity labeling to characterize the interacting proteins of METTL3/METTL14, the catalytic heterodimer of the major N^6-methyladenosine (m^6A) methyltransferase complex.

I utilized MRM-based targeted proteomics coupled with scheduled-liquid chromatography (LC) and synthetic stable isotope-labeled (SIL) peptides to systematically quantify the differential expression of small GTPase upon the adipogenesis of murine

mesenchymal stem cells (MSCs), including 3T3-L1 and C3H10T1/2 cells. The results revealed Rab32 as a novel adipogenesis regulator, where its downregulation is required for successful adipogenic differentiation in MSCs. In addition, I further employed the LC-MRM method to interrogate systematically the temporal responses of the entire small GTPases proteome during the course of osteogenic differentiation of H9 human embryonic stem cells. The quantification results revealed altered expression of a large number of small GTPases accompanied with osteogenic differentiation, especially those involved in autophagy. The results also documented a previously unrecognized role of KRAS in osteogenesis, where it regulates the accumulation of extracellular matrix for mineralization through attenuating the activity of secreted matrix metalloproteinase-9 (MMP9).

To understand how m^6A in mRNA regulates the expression of small GTPases, I also adopted the LC-MRM platform to assess the differential expression of small GTPases in cells upon genetic depletion of m^6A reader, writer, and eraser proteins. The results demonstrated that depletion of METTL3 and ALKBH5 (m^6A demethylase) resulted in substantially altered expression of a subset of small GTPase proteins, including RhoB and RhoC. The results further demonstrated that the mRNA stability of RhoB is significantly augmented in *METTL3*$^{-/-}$ cells, suggesting a m^6A-mediated decay of RhoB transcript.

To investigate the protein interactomes of METTL3/METTL14, I employed APEX-based proximity labeling followed by LC-MS/MS analysis to examine systematically the interacting proteins of the METTL3 and METTL14 in HEK293T cells. The results revealed a subset of histone modification writer and eraser enzymes to be enriched in the proximity proteome of METTL3, including histone acetyltransferase 1 (HAT1). Consistently, the

acetylation level of H3K9 diminished in *METTL3*$^{-/-}$ cells, which could be restored upon complementation with wild-type, but not catalytically inactive METTL3. We also discovered that the METTL3-catalyzed m^{6}A in chromatin-associated RNAs (caRNAs) promotes the chromatin occupancy of the long α isoform of HAT1 (HAT1 α), thereby stimulating H3K9Ac, especially at loci of LTR and LINE retrotransposons. The HAT1 occupancy in chromatin was diminished upon genetic depletion of METTL3 or METTL14 while its global expression was unaltered, suggesting that METTL3-containing methyltransferase complex regulates chromatin localization of HAT1 without altering its expression level. Results from our *in vitro* acetylation assay demonstrated the ability of HAT1 α to catalyze H3K9Ac in nucleosomes. Moreover, HAT1's chromatin occupancy and its function in acetylating H3K9 in chromatin requires its nuclear localization and is promoted by YTHDC1, an m^{6}A reader protein that also interacts with METTL3.

Together, the research presented in this dissertation displayed the robustness of modern proteomics to systematically investigating the proteome of interest with excellent sensitivity, selectivity, and throughput, and its power in offering new biological insights.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 Chapter 1. Introduction

1.1 General Overview

Mass spectrometry-based proteomics constitutes a powerful tool for investigating systematically the proteome with high throughput and identification/quantification accuracy. Advancements of mass spectrometry in the last two decades, including the development of high-resolution mass spectrometers, acquisition methods, and data analysis tools, foster the discoveries of numerous important biological regulations. Moreover, many proteomics-driven biomedical findings were translated into therapeutic developments. In modern mass spectrometry-based proteomics, top-down approach, which analyzes intact protein, provides comprehensive and specific molecular information. On the other hand, bottom-up proteomics, which interrogates proteolytically cleaved peptides, is the most widely employed technique due to its high throughput. In this chapter, I will discuss modern bottom-up proteomics analysis methods, including untargeted discovery proteomics operated in data-dependent acquisition (DDA) or data-independent acquisition (DIA) mode; and targeted proteomics operated in multiple-reaction monitoring (MRM) or parallel-reaction monitoring (PRM) mode.

Next, I will discuss the principles and the method innovations for modern bottom-up proteomic analysis, including recent developments in instrumentation and sample preparation techniques. Subsequently, I will describe quantitative proteomics approaches, which encompass the MS^1- and MS^2-based quantification methods, followed by the principles of data processing and interpretation. Common quantitative proteomics approaches will be elaborated, including stable isotope labeling by amino acids in cell

culture (SILAC), label-free quantification (LFQ), non-isobaric and isobaric labeling, and stable isotope-labeled (SIL) standard peptides. In the end, I will discuss state-of-the-art proteomic methods for investigating systematically protein-protein interactions, including co-immunoprecipitation (co-IP), proximity labeling, and cross-linking (XL). I will describe the basic principles and explain the advantages and limitations of each method.

Next, I will move on to discuss the biological regulations studied in this dissertation. I will first introduce the Ras superfamily of small GTPases, including their cellular functions and important members in each of the small GTPase subfamilies. I will also introduce the N^6-methyladenosine (m^6A), the most abundant internal modification installed post-transcriptionally on mRNA. Moreover, I will elaborate the reader, writer, and eraser proteins of m^6A for rapid and dynamic RNA regulations. In the end, I will introduce the scope of the dissertation.

1.2 Bottom-Up Proteomic Techniques

From receiving environmental stimuli through receptors located on the plasma membrane to gene expression regulations in the nucleus, proteins are the direct functional molecules in living organisms. Dynamic regulations of proteins, which include translation, post-translational modifications (PTMs), subcellular localization, protein-protein interactions, and turnover, contribute to most if not all biological processes. The size of human proteome was originally anticipated as much as 100,000 but was later revealed to be ~ 20,000 by human genome project if a single representative protein is used as the translation product of each gene[1]. The divergences are in part due to the different functional protein products that could be derived from a single gene, of which the variations arise

from in allelic alteration (i.e., single-nucleotide polymorphism, SNP), alternative splicing/translation, and PTMs. Approximately a decade ago, Kelleher et al.[2] proposed the term 'proteoform' to describe all those different protein products generated from a single gene. The different proteoforms, although derived from the same gene, have differential cellular functions.[3] Thus, investigating the proteome composition is indispensable in studying dynamic cellular regulation. At the systematic regulation level, the spatiotemporal protein-protein interactions further lay the foundation for complex cellular functions. Thus, given that proteins are the building blocks of life, investigating protein-protein interaction networks will be a direct and precise surrogate in revealing biological regulations.

Mass spectrometry emerged as a powerful tool for interrogating the proteome. In the last two decades, marked improvements have been made in the resolution, dynamic range, sensitivity, and scanning speed of mass spectrometers. These instrumental advances are further compounded by whole genome sequencing that enables sequence-based approach to map proteomes, of which the framework expedites downstream data analysis[4]. Thus, proteomics has been widely employed in basic and clinical research, and pronouncedly advanced our knowledge in biomedical studies such as disease progression[5,6]. Currently, two approaches, referred to as 'top-down' and 'bottom-up', are commonly employed in proteomics analysis, while 'middle-down' approach has been employed for better PTM mapping (Figure 1.1). In top-down proteomics, intact proteins are introduced into mass spectrometers for fragmentation. When available, top-down proteomics provides the richest information about protein proteoforms such as PTMs. Recently, Ge and colleagues[7] demonstrated, in top-down proteomics coupled with two-

dimensional chromatography separation, the successful measurement of intact proteins up to 223 kDa as well as the identification of more than 4000 unique protoeforms. Robinson[8] and colleagues[8] also employed top-down approach to characterize membrane proteins and their associations with ligands, such as lipids and metabolites. Thus, the recent advancements highlight the potential of top-down proteomics in analyzing most proteins. However, this approach remains challenging to execute in general due to the complexity of fragmentation pattern and other technical limitations such as protein solubility.

Bottom-up proteomics, which was pioneered by Yates and colleagues[9] nearly thirty years ago, relies on proteolysis by using proteases (e.g., trypsin) and the resulting peptides are subsequently subjected for tandem mass spectrometry (MS/MS) analysis (Figure 1.1). The peptides are mapped by their mass-to-charge ratio (*m/z*) of the precursor ions and a cross-correlation of the observed fragment ions to those predicted from the peptide sequence *in silico*. Thereafter, bottom-up approach was further advanced by peptide separation techniques such as liquid chromatography in the nano flow range[10]. In addition, various quantitative proteomics methods were developed to improve the quantification accuracy and throughput, including metabolic and isobaric labeling.

With the recent developments in instruments and quantitative approaches, bottom-up proteomics allows for high-throughput characterizations of the proteome. In this section of the Introduction, I will focus on the recent developments in bottom-up proteomics. I will first introduce discovery proteomics operated in DDA and DIA modes and targeted proteomics operated in MRM- or PRM-based approaches. Next, I will discuss the recent

developments of mass spectrometers for optimizing peptide sensitivity. Finally, I will introduce the sample preparation workflow for bottom-up proteomics.

1.2.1 Discovery Proteomics

In DDA mode analysis, peptides co-eluted from liquid chromatography are subjected to a survey scan at MS^1 level (a.k.a. full scan) to analyze the *m/z* of precursor ions. Afterwards, the mass spectrometer, depending on the intensities of the precursor ions measured in survey scan, isolates the top *N* ion species in real-time for fragmentation, and the fragment ions are scanned at MS^2 level (Figure 1.2). In a typical cycle of survey and MS^2 scan, the instrument selects top 10 and sometimes up to top 25 precursor ions while 12-15 MS^2 scans were shown to give the best protein and peptide identification[11]. Notably, a minimum intensity threshold should be employed to avoid isolating low-abundance precursor ions, which often take up scanning time but fail to produce sufficient quality of MS^2 spectra for peptide identification. Yates and colleagues[12] demonstrated that the optimal threshold of minimal intensity is around the noise level of the mass analyzer.

The proteome coverage is further improved by introducing dynamic exclusion that precludes the mass spectrometer from isolating the same precursor ions for fragmentation within a user-defined time window. In the top *N* selection strategy used in DDA mode analysis, the original setup of precursor ion selection will lead to redundant acquisition of MS^2 spectra of the same high-abundance peptides. By incorporating dynamic exclusion into decision making, the same precursor ions of which the MS^2 spectra have just been acquired in the previous MS^2 scan will be disregarded. Instead, the mass spectrometer will isolate peptides ions of lower abundances, allowing for a better peptide coverage. Florens

and colleagues[13] demonstrated that by applying dynamic exclusion in DDA analysis, the proteome coverage could be improved by 2-fold. Recently, supervised learning algorithm was also adapted for on-the-fly dynamic exclusion decision to further allow mass spectrometer to acquire MS^2 for more low-abundance precursor ions[14].

After acquisition, data generated in DDA mode are subjected to search engines for peptide identification. Commonly used tools include MaxQuant[15], Fragpipe[16], and commercialized software such as Proteome Discoverer[17]. Typically, the data are searched against a confined proteome database (e.g., human proteome) rather than *de novo* to improve the speed and to reduce the computational cost. Pioneered by Yates and colleagues[9], the data search engines extract the MS^2 spectra and match the fragment ions to those generated from *in silico* digestion under a user-defined mass tolerance. The best assignment is considered as the peptide spectrum match (PSM), which is subsequently scored according to the matching quality. To gain confidence of peptide identification, a false discovery rate (FDR) control should be applied. Among the FDR control methods, target-decoy approach proposed by Gygi and colleagues[18] is currently the most widely used. Briefly, decoy peptides comprised of amino acid sequences that do not exist in nature are incorporated in data search. The scores of the target and decoy PSMs are used for posterior FDR control with search engine-defined algorism. Taking MaxQuant as an example, the PSMs are subsequently ranked based on the above-mentioned scoring system and integrated into the identified peptide list until the percentage of decoy PSMs exceeds the pre-defined threshold (e.g., 1%). The data computation for DDA mode analysis is discussed in great detail in a recent review[19].

Despite being commonly used in discovery proteomics, DDA mode analysis suffers from the number of peptides that could be analyzed in MS^2 scans, which is intrinsically limited by the sampling speed of mass spectrometers. Dozens or hundreds of precursor ions are detected in each of the survey scan yet only a small portion of them could be subjected to MS/MS analysis prior to next survey scan. The limitation of DDA leads to the development of DIA-based discovery proteomics analysis, which allows all concurrent peptides within a predefined *m/z* range to be isolated, co-fragmented, and scanned at MS^2 level simultaneously (Figure 1.2). Since the precursor ions are all subjected to MS/MS analysis, DIA mode allows for re-mining of unidentified peptides from the once-and-forever acquired data sets. Moreover, since all precursor ions are unbiasedly co-fragmented without intensity-based selection, DIA approach often improves the reproducibility in protein identification across multiple experiments as compared to top *N*-selecting DDA analysis. A study showed that over 98% of proteins are commonly identified across 24 samples in DIA analysis, while over 50% of protein identity is missing when the samples were analyzed in DDA mode[20].

Different DIA methods have been introduced in the past two decades to optimize the acquisition conditions for discovery proteomics while only a few are still being used in proteomics studies. Pioneered by Aebersold and colleagues[21] in early 2000s, sequential window acquisition of all theoretical mass spectra (SWATH) is one of the most popular approaches in DIA analysis. In SWATH acquisition, the mass spectrometer is set to cycle through 32 consecutive precursor ion *m/z* range increments in the range of *m/z* 400-1200 in a user-defined retention time window (Figure 1.3). Note that in DIA analysis, the

isolation *m/z* window needs to be optimized. If the window is set too wide, the MS^2 spectra will be complicated by too many concurrent precursor ions and undesirable interferences will be introduced. Consequently, data analysis and interpretation become challenging and/or error-prone. Further information on DIA-based bottom-up proteomics can be found in recent reviews[22-24].

Although promising in principle, DIA is challenging in data analysis due to the complexity of the acquired MS^2 spectra. In the last decade, several search engines, including DirectDIA[25] and PECAN[26], were developed to optimize the data analysis for results obtained in the DIA mode. Shortly after introducing SWATH approach, Aebersold and colleagues[27] also developed openSWATH, a DIA data analysis tool which applies a workflow conceptually similar to that of targeted proteomics (section 1.2.2). After acquisition, the results are matched to a comprehensive or project-specific spectral library. The retention time, relative intensities of fragment ions, and the isotopic distribution of precursor and fragment ions are referenced for peptide identification[27,28]. It should be noted that openSWATH, unlike tools that automate all processes, requires manual inspection of extracted-ion chromatograms (XICs). Although time-consuming, openSWATH ensures rigorous data analysis and the reported data are typically of better confidence. Recently, tools like Skyline[29] seem to bridge the gaps of the automated tools and manual inspection. Skyline provides a user-friendly environment for the visualization of XICs when inspecting DIA data.

Together, mass spectrometry methods and data analysis approach are continuously improving for DIA-mode discovery proteomics. The high reproducibility, proteome

coverage, and the capability of being re-mined prompt DIA mode analysis as an attractive method for large-scale bottom-up proteomics.

1.2.2 Targeted Proteomics - MRM and PRM

While discovery proteomics allows for unbiased investigation of protein expression, it could be unsatisfactory when a project is seeking to systematically analyze a subset of proteins of interest. For systematic screening of a subset of low-abundance proteome (e.g., biomarkers), targeted proteomics, in which the mass spectrometer is programmed to specifically monitor the selected precursor ions, has become the method of choice. Unlike discovery proteomics, for which the goal is to identify as many proteins as possible, targeted proteomics aims to monitor selected proteins with high sensitivity, reproducibility, and quantification accuracy. Currently, MRM (a.k.a. selected-reaction monitoring, SRM) and PRM approaches are commonly used depending on the types of mass spectrometers and the nature of samples (Figure 1.2).

MRM analysis is often operated on a triple-quadrupole mass spectrometer (QqQ). The first quadrupole mass analyzer is programmed to isolate a specific precursor ion for fragmentation. In the third quadrupole, pre-defined selected fragment ions are filtered and guided to the detector for quantification (Figure 1.2). The pair of targeted precursor ion and the ensuing product ion is termed ‘transition’. Thus, in MRM, the mass spectrometer is monitoring the trace of the transition over its retention time. For confident identification and unbiased quantification, the transition list for the peptide must be carefully chosen.

The selection of precursor ions requires up-front analysis in discovery proteomics. Once the discovery proteomics data are acquired, transitions that will be used to quantify

protein expression could be picked from the MS^1 and MS^2 spectra of the peptide. To facilitate the development of MRM platform, transition should be selected with the following criteria: (1) The peptide should give optimal signal intensity in both MS^1 and MS^2 spectra. (2) The peptide should be unique to the protein of interest. (3) Fragment ions of the highest intensities in the MS^2 spectra acquired in discovery proteomics should be selected. Typically, 3 to 5 transitions per peptide are selected for its quantification. Notably, all transitions from the same precursor should be concurrent (Figure 1.4), which is used to ensure identification accuracy. (4) Peptides with modifications or missed cleavage sites should be avoided since their quantification accuracy could be compromised by proteoforms. Further information about the experimental setup of MRM analysis could be found elsewhere[30,31].

It should be noted that quadrupole is a low-resolution mass analyzer, and the peptide identification/quantification accuracy for complex samples is significantly hampered. In this regard, crude enrichment is essential for successful quantification of MRM-based targeted proteomics. In a recent study published by our group[32], we employed SDS-PAGE to separate small GTPase proteins from the rest of the proteome based on their molecular weights. In this vein, over 100 small GTPases could be consistently quantified, representing approximately 60% of the human small GTPase proteome. More importantly, selected small GTPases were validated by orthogonal methods such as Western blot analysis. The results demonstrate the high sensitivity and quantification accuracy of the MRM-based targeted proteomics method.

PRM experiments are typically conducted on high-resolution mass spectrometers (HRMS) such as quadrupole-Orbitrap hybrid instruments[33]. Due to the use of HRMS, PRM is empowered to analyze complex samples such as whole-cell lysate. Like MRM, the mass spectrometer is programmed to isolate precursor ions of interest for fragmentation, but all fragments are subjected to spectrum acquisition at MS^2 level in the PRM workflow (Figure 1.2). For quantification, PRM approach monitors the intensities of selected fragment ions like MRM. Typically, 5-8 most abundant concurrent fragment ions are quantified to represent the intensity of the peptide[34]. To strengthen the identification confidence, survey scan could be implemented to acquire the spectra of precursor ions. The peptide selection criteria for PRM analysis are similar to that of MRM analysis. Unique peptides with good intensities in both MS^1 and MS^2 are typically good candidates.

Mass spectrometer should be fine-tuned to ensure optimal sensitivity and quantification accuracy. To avoid interferences in MS^2 spectra such as concurrent peptides that could perturb quantification accuracy, the quadrupole isolation window is often set at approximately 1 *m/z*. While further narrowing down the isolation window could improve selectivity, the sensitivity is also compromised due to the decreased ion transmission efficiency in the quadrupole. Thus, finding a balance between the selectivity and sensitivity is crucial for PRM experiments. To improve sensitivity for increased proteome coverage, maximum accumulation time could be increased, which, however, comes at a cost of instrument hygiene[34]. In recently introduced instruments, the acquisition parameters such as maximum injection time and fragmentation energy could be specified for each of the

precursor ion monitored, allowing for a more efficient PRM acquisition method. The implementation and application of PRM are further discussed in recent reviews[34,35].

Despite showing high sensitivity and quantification accuracy, targeted proteomics is limited to the number of peptides it could analyze within a single injection. Depending on the types of instruments, the maximum scan rates of mass spectrometers are typically about a dozen scans per second. In addition, increasing the number of precursor ions will fundamentally reduce the ion accumulation time, which compromises sensitivity. To resolve the problem, different approaches have been developed for targeted proteomics. Scheduled liquid chromatography (scheduled LC) has been used to improve the throughput of target proteomics analysis[32,36,37].

First introduced by Rinner and colleagues[38], normalized retention time (iRT) approach is currently routinely employed in target proteomics experiments. The principle of iRT-based schedule is that, although peptides' absolute retention times fluctuate depending on instrumental conditions, their relative retention times to pre-selected standard peptides remain steady. Hence, through obtaining the retention times of standard peptides in calibration runs, the retention times of the monitored peptides could be predicted. In this vein, the mass spectrometer is programmed to analyze the peptide only within a narrow retention time window, thereby decreasing the number of concurrently monitored peptides (Figure 1.5).

In early 2000s, Domon and colleagues introduced on-the-fly scheduling acquisition in targeted proteomics analysis, which fine-tune the predicted retention time according to the spike-in standard peptides in real-time[39,40]. While the approach is conceptually similar

to iRT, the dynamic correction of retention time allows for the acquisition window to be further reduced to as narrow as one minute. Thus, although scheduled LC requires extensive up-front analysis, it significantly improves the throughput in targeted proteomics.

Together, careful selection of peptides from the up-front discovery mode analysis, is essential for successful targeted proteomics analysis. Although setting up targeted proteomics together with scheduled LC is a tedious process, a robust platform can be used for routine protein quantification once established. Due to its high sensitivity, quantification accuracy and throughput, targeted proteomics is proven to be the method of choice for systematic and reproducible analysis of a subset of proteins of interest. In this regard, MRM-based approach is recently gaining attention in biomarker screening[41].

1.2.3 Mass Spectrometer Instrumentation in Bottom-Up Proteomics

Recent developments in mass spectrometers and the related instruments have greatly advanced the depth and throughput in bottom-up proteomics. Given that mass spectrometer analyzes ion species, the ionization efficiency of peptides is critical for successful bottom-up proteomics. A major driving force for bottom-up proteomics is the innovation of nanoelectrospray ionization (nESI), which is compatible with liquid chromatography (Figure 1.6). Developed by Wilm and Mann[42] in 1990s, nESI is now commonly used in tandem with online peptide separation. The sensitivity of electrospray ionization depends on the concentration of analytes and prefers low flow rate[43]. In nESI, ultralow flow rate (i.e., few hundred nanoliter per minute) is applied on capillary column without sheath gas, allowing for the minimization of the sample consumption and simultaneous analysis of peptide mixtures with good sensitivity[42]. Moreover, owing to the

small size of droplets it generates, desolvation is also improved in nESI, which further enhances the ionization efficiency of peptides[43]. Currently, the development of ionization techniques is still active in mass spectrometry field, and more innovations are being introduced to advance proteomics analysis[44,45].

Another soft ionization technique commonly employed in bottom-up proteomics is the matrix-assisted laser desorption/ionization (MALDI). In MALDI, the sample is mixed with matrix material and spotted onto a metal plate. A laser is then applied to vaporize the samples at ambient temperature and the ionized peptide ions are guided into a mass spectrometer[46]. Although the approach is not compatible with liquid chromatography, samples separated by off-line HPLC could be spotted onto MALDI plates. To achieve online separation, capillary electrophoresis has been implemented to be operated in tandem with MALDI followed by mass spectrometry analysis[47].

In tandem mass spectrometry, the precursor ions are isolated and fragmented to acquire MS^2 spectra for mapping peptide sequence. In this regard, optimal fragmentation is crucial for bottom-up proteomics. While sufficient bond cleavage is required, excessive fragmentation yielding too many low *m/z* ions should also be avoided as they provide modest values in peptide identification. In the past decades, collision-induced/activated dissociation (CID/CAD) is one of the most commonly used fragmentation techniques in triple-quadrupole and linear ion trap mass spectrometers[48]. Recently, beam-type fragmentation that uses higher amount of energy is introduced on the quadrupole-time-of-flight (QTOF) mass spectrometer, which later becomes available on Orbitrap instruments as higher-energy collision dissociation (HCD). HCD typically results in more extensive

fragmentation and generates informative fragment ions in low *m/z* region. When implemented on high-resolution mass spectrometers such as Orbitrap, HCD benefits the isobaric labeling-based quantification approaches such as tandem mass tag (TMT, see section 1.3.2.1), which quantify the relative abundances of low-mass reporter ions[49].

Both CID and HCD start with collisions of neutral gas such as nitrogen, during which the kinetic energy is transferred into internal vibrational energy, which is redistributed in the precursor ions. Once the vibrational energy exceeds a certain threshold, it breaks the weakest bond. CID and HCD efficiently fragment the peptide bond and generate robust coverage of b- and y-ions of a given peptide (Figure 1.7). Other ions such as immonium ions or internal fragments could also be generated from alternative fragmentation pathways. Mann and colleagues[49] showed that in both HCD and CID, the MS^2 spectra of most peptides could be interpreted in DDA analysis, suggesting the robustness of the fragmentation methods in bottom-up proteomics.

Alternatively, electron transfer dissociation (ETD) could be used for peptide fragmentation. An advantage of ETD over CID and HCD is that labile modification moieties such as phosphate group are better preserved, allowing for effective PTM mapping. ETD induces cleavage predominantly at the N-C_α bond and thus generates c- and z-ions. Moreover, ETD is shown to outperform CID and HCD for highly charged precursor ions[50-52]. Thus, the fragmentation methods and conditions should be carefully evaluated and chosen for different experimental objectives. A detailed study of fragmentation methods in bottom-up proteomics was conducted by Mohammed and colleagues[52].

Mass analyzers such as quadrupole, linear ion trap (LIT), Orbitrap, and time-of-flight (TOF) are commonly used in bottom-up proteomics. Currently, Thermo's quadrupole-Orbitrap hybrid system dominates the DDA mode analysis. Orbitrap mass analyzer, since its innovation in the late 1990s, has demonstrated excellent mass accuracy and resolution, which is highly desirable for DDA mode bottom-up proteomics of complex peptide mixtures[53]. The development of parallelized Quadrupole/Linear Ion Trap/Orbitrap Tribrid mass spectrometers further improves the acquisition rate by allowing survey and MS^2 scan to be conducted in different mass analyzers simultaneously. Recently, Bruker's timsTOF Pro, a hybrid time-of-flight mass spectrometer coupled with trapped ion mobility spectrometry (TIMS) and parallel accumulation-serial fragmentation (PASEF), shows excellent sensitivity and proteome coverage, allowing for the analysis of low-quantity protein samples[54]. Remarkably, Andrew and colleagues[55] were able to identify 3,666 proteins in a single 5-min LC gradient on the timsTOF Pro in the DDA mode. In addition, Mann and colleagues[56] showed that nearly 4,000 proteins could be identified with as low as 10 ng of proteins when operated in DIA mode on the timsTOF Pro, highlighting the superb sensitivity of the instrument. Together, the continuing improvement in sensitivity, resolution, and mass accuracy of mass spectrometers will further enhance the efficiency and robustness in analyzing complex biological samples[57].

1.2.4 Protein Extraction and Digestion

1.2.4.1 Cell Lysis Methods

To prepare samples for bottom-up proteomic analysis, proteins are extracted from biological samples such as cultured cells or tissues. Commonly used approaches for

efficient protein extraction include ionic detergents such as sodium dodecyl sulfate (SDS) or chaotropic agents such as urea. Notably, the extraction condition should be compatible with downstream steps. For example, co-immunoprecipitation requires preservation of protein structure, thus mild lysis conditions (e.g., nonionic detergent) should be used. On the other hand, to fully extract insoluble proteins (e.g., stress granules), harsh conditions such as the combination of urea and sonication could be employed to facilitate protein extraction[58]. Nevertheless, the protein extraction is the first crucial step and may need to be optimized according to the nature of samples and experimental needs.

1.2.4.2 Protease Digestion

In the peptide-centric bottom-up proteomic analysis, efficient protease digestion is essential for successful mass spectrometry analysis. Moreover, an unbiased digestion procedure is required for accurate relative or absolute protein quantification. For complete proteolysis, proteins should first be denatured by using surfactants or chaotropic agents to disrupt protein tertiary structures, which consequently allows for the exposure of protease digestion sites buried within the proteins. SDS and urea are commonly used due to their denaturation efficiency. However, strong denaturation reagents are incompatible with most if not all protease digestion; thus, these reagents need to be sufficiently removed prior to proceeding with enzymatic digestion. In this vein, alternative denaturation approaches have been developed to assist sample preparation[59]. For instance, Yu et al.[60] demonstrated that acid-labile anionic surfactant, which is sufficient in denaturing proteins, does not inhibit the activity of commonly used proteases such as trypsin.

Cysteine residues in proteins could form disulfide bonds to assist protein folding. Those disulfide bonds, however, will lead to cross-linked peptides that pose extra challenges in ionization efficiency due to their higher molecular weight as well as data search due to the complexity of disulfide bond combinations. To break the disulfide bonds, proteins are incubated with reduction reagents such as dithiothreitol (DTT). To prevent the disulfide bonds from reformation during the ensuing sample preparation steps, cysteines should be further alkylated by reagents such as iodoacetamide (IAA). Recently, Müller and Winter[61] conducted a systematical analysis of protein reduction/alkylation conditions, and they found effective combinations that could be adapted into sample preparations workflow in proteomic analysis. Subsequently, the protease is added to digest the proteins into peptides.

Notably, the protease must be carefully selected to yield peptides length that is compatible with mass spectrometry analysis. Ideally, the resulting peptides should be approximately 6-25 amino acids for the commonly used electrospray ionization. Currently, trypsin that cuts the amide bonds on the C-terminal sides of lysine and arginine is commonly used due to its proteolysis efficiency and the frequencies of the two amino acid in proteins. However, alternative proteases and/or multi-protease digestion could be used for different experimental purposes or to improve proteome coverage. Further information about protease selection is discussed elsewhere[62,63]. Finally, the sample may need to be cleaned up to remove MS-incompatible components (e.g., salts) prior to LC-MS/MS (provide full name for the first time) analysis.

Based on the aforementioned general guidelines, several methods have been developed to streamline sample preparation procedures, each with its strengths and limitations. The most straightforward strategy is the single-pot in-solution digestion. In this approach, all steps (i.e., from protein denaturation, reduction/alkylation to protease digestion) are sequentially conducted in the same tube without transferring samples[64]. If protease-inhibiting denaturing reagent is used, the samples should be diluted sufficiently to reach protease-tolerable concentrations of denaturing reagents. The protease is subsequently added, and the resulting peptide mixture could be acidified for C_{18}-based peptide clean-up and enrichment. Although being straightforward and easy to perform, in-solution digestion could not remove impurities and detergents. Consequently, protease digestion efficiency may be compromised, and mass spectrometer could be contaminated.

Alternatively, in-gel digestion method, pioneered by Mann and colleagues[65,66] about three decades ago, has been extensively used in proteomics due to its robustness in sample clean-up prior to protease digestion. After electrophoresis, the gel is fixed and subsequently cut into small pieces. The gel cubes are incubated with organic solvent (e.g., acetonitrile) to shrink and then expanded in different aqueous buffers for reduction/alkylation and protease digestion. Notably, the protein bands could be visualized through gel staining (e.g., with Coomassie brilliant blue) and excised according to their molecular weights for fractionation, which reduces the complexity of the ensuing samples for better sensitivity and proteome coverage[32].

An advantage of in-gel digestion is its compatibility with a variety of detergents for effective protein solubilization and denaturation. The surfactants along with other

impurities are efficiently removed during gel electrophoresis or wash steps, and thus do not interfere with LC-MS/MS analysis. Notably, in-gel digestion requires small and robust proteases such as trypsin that can efficiently diffuse into the gel pieces, which limits the choices of proteases. A major drawback of in-gel digestion is the relatively poor peptide recovery, making it unsuitable for processing low quantities of samples. In addition, the method could not be easily automated, rendering it incompatible with high-throughput sample preparation. Nevertheless, in-gel digestion with the unique features of gel electrophoresis is well established and proved to be a simple and efficient method for preparing samples for bottom-up proteomics[67-69].

To employ solution-based protease digestion without the aforementioned impurity issues, Mann and colleagues[70] developed filter-aided sample preparation (FASP) approach. In FASP, commercially available microcentrifugation device with user-selected molecular weight cutoff is used to conduct all steps in the filter cup, and the solution could be removed by centrifugation prior to the next step. Given that the denaturation reagents along with low-molecular weight impurities could be efficiently removed through centrifugation, strong surfactants or chaotropic agents could be incorporated into the protein extraction and denaturation steps. In addition, excess reagents such as iodoacetamide, which is used for cysteine alkylation, could also be depleted prior to protease digestion. Furthermore, during peptide elution, the filter retains high molecular weight impurities or unwanted products such as partially digested proteins, which may interfere with downstream LC-MS/MS analysis. Thus, due to its clean-up capability and better peptide recovery, FASP

has emerged as an attractive sample preparation method for bottom-up proteomics and many filter-based approaches have since been developed to meet users' needs[71,72].

1.3 Quantitative Proteomics

While RNA sequencing has been employed as a surrogate to quantify gene expression, the results do not always align with those of protein expression. Translational efficiency and protein turnover both contribute to the dynamic regulation of protein expression on top of its mRNA expression. Moreover, proteins can be chemically modified to execute different functions. Thus, quantitative proteomics could provide important insights into biological regulations. To date, various proteomic approaches have been developed to quantify protein expression using MS^1 or MS^2 scan. In this section of introduction, I would discuss the recent developments and applications of quantitative proteomics.

1.3.1 MS^1-Based Quantification

1.3.1.1 SILAC

Recent developments in mass spectrometry-based proteomics enable automatic analysis of thousands of proteins in a single LC-MS/MS run. To quantify the expression of those proteins, different peptide labeling approaches have been developed. Among them, SILAC, which was introduced by Ong et al.[73] in the early 2000s, has been extensively employed in quantitative proteomics due to its quantification accuracy. In SILAC experiments, the cells are grown in media containing unlabeled (light) or stable isotope-labeled (heavy) essential amino acids, which could not be biosynthesized in mammalian

cells. In this vein, the only source of these amino acids will be from the culture medium, thereby forcing their incorporation into proteomes of the cells.

Currently, the combination of trypsin digestion, a protease that cleaves amide bonds on the C-terminal side of lysine and arginine, and SILAC labeling with ^{13}C- and ^{15}N-labeled lysine and arginine is the most commonly used. In this combination, all tryptic peptides except the C-terminal peptides of some proteins contain at least one heavy isotope-labeled amino acids, allowing for comprehensive and unbiased peptide quantification. Notably, arginine is actually not characterized as an essential amino acid; however, its inclusion in culture media is required for normal proliferation in many cell lines[74]. Arginine was reported to be converted between proline in some cell lines (e.g., HeLa cells)[75]. The conversion may complicate data analysis and, if the conversion is not incorporated into data search properly, distorts the quantification results. To resolve this issue, one can supplement the SILAC medium with excess unlabeled proline and/or labeled arginine manipulate the conversion equilibrium[76].

For successful SILAC-based quantification, complete incorporation of the heavy isotope-labeled amino acids into the proteome is essential. To achieve this, the cells must be cultured in SILAC medium for sufficient time depending on the cell doubling time, and the incorporation rate must be examined up-front through LC-MS/MS analysis[77]. To prevent biased quantification results arising from heavy isotope-labeled amino acids, the assignments of SILAC media to the testing samples is often swapped in different biological replicates.

Immediately after harvesting, cells cultured in different SILAC media are combined by either equal cell number or protein quantity, and the ensuing mixture is subjected to protease digestion. Given that all samples are combined at the very beginning, variations in subsequent steps of sample preparation do not affect the quantification accuracy. Thus, SILAC is a straightforward and highly accurate quantitative proteomics approach. In addition, the light and heavy forms of each peptide, due to their identical amino acid sequence, have almost identical elution profile in LC. Consequently, the interferences emanating from variations in instrument conditions, e.g., ionization efficiency, has negligible effects on relative peptide quantification. During LC-MS/MS analysis, each peptide appears as a pair in both the MS^1 and MS^2 spectra, with the quantification being predominantly achieved through the chromatographic peak area of the former.

To bring the advantage of SILAC-based quantification into animal studies, Wu et al.[78] developed, approximately two decade ago, the stable isotope labeling of mammals (SILAM) in rats through ^{15}N-labeled spirulina diet. Shortly after, Kruger et al.[79] introduced SILAM of ^{13}C-substituted lysine in C57B1/6 mice. These pioneering developments allow SILAC to be incorporated into mammalian models for quantitative proteomics analysis. Currently, SILAC approach is also available for various other model organisms, including fruit fly[80] and zebrafish[81].

Despite showing success in animal studies, the SILAC approach is not compatible with clinical proteomics. Geiger et al.[82] partially resolve this incompatibility by using heavy SILAC-labeled peptides derived from five cell lines, termed super-SILAC mix, as

isotopically labeled internal standards for relative quantification of different tissue samples. Through spiking-in equal amount of super-SILAC mix, the quantification of protein expression in clinical samples could be accurately normalized by its isotopically labeled counterparts. In line with this, Jaffe et al.[83] employed super-SILAC approach to profile histone modifications and successfully identified that the demethylation of histone H3 at Lys-36 is a common feature in the 115 cancer cell lines investigated in their study. Thus, although SILAC could not be directly employed in clinical proteomics, the principle has been successfully transferred for accurate quantification.

Since its first introduction about two decades ago, SILAC approach has revealed numerous important biological regulations such as protein expression[84], PTMs[83], and protein-protein interactions[85]. Moreover, because of that the heavy isotope-labeled amino acids are incorporated into proteins during translation process, pulsed SILAC (pSILAC) strategy has been used for direct investigation of protein translational efficiency[86]. Recently, the multiplexity of SILAC is also expanded up to 5-plex, although the 2- and 3-plex remain the most commonly used[87]. Together, SILAC provides straightforward and accurate protein quantification for biological samples in general and cultured cells in particular. The experimental set up and data analysis of SILAC approach is discussed in great detail elsewhere[77,88,89].

1.3.1.2 Non-Isobaric Chemical Labeling

SILAC labeling procedures, despite being successfully adapted to higher model organisms (e.g., mice), is inherently very involved. Consequently, the approach is not cost-effective for most projects that include animal studies and clinical samples. The

development of the above-mentioned super-SILAC mix approach partially resolves the issue, yet proper selection and evaluation of cell lines could be time-consuming and may not contain all proteins that are present in the samples to be analyzed. To enable straightforward MS^1-based quantification similar to that of SILAC for samples that metabolic incorporation could not be easily achieved, stable isotope labeling approaches through click chemistry have been developed.

The non-isobaric labeling for peptide quantification at MS^1 is pioneered by the development of isotope-coded affinity tag (ICAT) in the late 1990s[90,91]. The approach involves reaction of free cysteine residues in proteins with thiol-specific ^{13}C-labeled iodoacetamide probe, and the labeled peptides are subsequently enriched through affinity purification and analyzed by LC-MS/MS. However, due to the relatively low frequency of occurrence of cysteines in proteins, it is estimated that only approximately 10-20% of the whole cell proteome could be quantified through ICAT, which left a large number of proteins uncharacterized. One way to resolve this problem is to label the peptide through their free amines (N-terminus and the side chain of lysine), which are carried by most tryptic peptides.

Reductive dimethyl labeling, which was pioneered by Hsu et. al[92] and further optimized by other groups[93,94], is an efficient and inexpressive approach to convert primary amines to dimethylamines with different combinations of ^{13}C and deuterium-labeled reagents (Figure 1.8A). Recently, mass difference tandem mass tagging (mTMT)[95] and tag for relative and absolute quantitation (mTRAQ)[96] that carry amine-reactive NHS-ester moiety were developed as an alternative approach. Unlike dimethyl labeling that may

suffer from retention time drift due to the use of deuterium, mTMT and mTRAQ that employ ^{13}C and ^{15}N stable isotopes allow for the coelution of the labeled peptides. Thus, the quantification will not be affected by fluctuation of instrument conditions.

Non-isobaric stable isotope labeling is an attractive approach for MS^1-based quantitative proteomics. In addition, the approach could improve the quantification precision and accuracy, particularly in targeted proteomics[97]. Furthermore, a study that compared SILAC- and mTRAQ-based approach demonstrated comparable quantification results, highlighting non-isobaric stable isotope labeling as a useful alternative to metabolic isotope labeling[98]. Together, non-isobaric stable isotope labeling is a robust alternative approach for SILAC.

1.3.1.3 Label-Free Quantification

Over the years, mass spectrometry-based quantitative proteomics becomes indispensable in investigating biological regulations. While the aforementioned labeling approaches have been developed for relative quantification, they are not always applicable due to the nature of the samples. For instance, although SILAC method provides high accuracy, metabolic labeling procedures cannot be transferred to clinical studies in general. The non-isobaric labeling approach, despite resolving sample incompatibility issue, introduces extra steps that will result in additional sample loss, thus prevents it from being used for low-quantity samples. In this regard, label-free quantification, which requires no additional chemical labeling or sample preparation, is gaining popularity in fields like clinical proteomics.

Unlike the labeling-based MS^1 quantification approaches that combine multiple samples into a single LC-MS/MS analysis, label-free quantification simply involves analyzing these samples separately by LC-MS/MS. Label-free quantification is subsequently achieved through comparing peak areas in the extracted-ion chromatograms (XICs) of precursor ions. Chelius and Bondarenko[99] demonstrated that the peak area correlates linearly with peptide concentrations in the sample mixtures, laying the foundation of chromatogram-based label free quantification. However, caution should be taken about the dynamic range of the mass analyzer to prevent over- or underestimating the relative quantification results.

Alternatively, MS^2 spectra counting, pioneered by Liu et al.[100] has been used as a surrogate for relative peptide abundance. In this approach, the number of MS^2 spectra of a given peptide are summed as the spectral count. To quantify a protein, the spectral count of all the peptides associated with the protein are compared across samples. Data processing approaches have been developed to unbiasedly translate the differential spectral counts into their relative protein expression[101]. However, it has been shown that the method still performs poorly in terms of quantification accuracy. Moreover, spectral counting approach requires high spectrum scanning speed, which is not always achievable. In addition, the required recurrent MS^2 acquisition conflicts with the aforementioned dynamic exclusion that prevents repetitive precursor ion selection. In addition, in contrast to chromatographic intensities over retention time, the spectral counting approach does not measure any physical property of the peptide. Besides, the approach has limited linear

range, further compromising its application in quantifying protein expression. Thus, spectral counting approach currently remains controversial[102].

Aside from its compatibility to most samples, the quantification accuracy in label-free approach has been significantly improved owing to the recent developments in computational proteomics[103-105]. A recent extensive assessment of label-free quantification software by Shweiki et al.[106] showed that tools like MaxQuant or Progenesis QIP can accurately quantify those proteins with abundances varying by two orders of magnitude. Another advantage of label-free quantification is the deep proteome coverage. In the labeling-based approaches, the same peptide needs to be scanned at least twice depending on multiplexity, which takes up more instrument time to obtain the quantification data of the same peptide.

A common drawback of label-free approach is the quantification reproducibility, which arises from the differential instrumental conditions such as liquid chromatography or ionization efficiency. In lieu of this, missing data, where the protein is only detected in some samples, is also commonly observed. The absence of protein intensity in some samples complicates data interpretation. It is noteworthy that while label-free quantification commonly uses MS^1-based quantification, the method could also be implemented in MS^2-based design. Indeed, label-free quantification of the product ion intensities is routinely used in targeted proteomics[107]. Together, due to its simplicity and sample compatibility, label-free quantification is a promising technique in quantitative proteomics analysis. Detailed discussion of label-free quantification can be found elsewhere[108,109].

1.3.2 MS^2-Based Quantification

1.3.2.1 Isobaric Labeling

Large-scale quantitative proteomics is powerful in systematically investigating biological regulations directly at the protein level. To improve the data acquisition rate and allow for simultaneous analysis of multiple samples on mass spectrometers, isobaric mass tags, which label peptides after protease digestion, have been developed. In a recent study, Gygi and colleagues[110] interrogated the proteome of 375 cell lines listed on the Cancer Cell Line Encyclopedia (CCLE), of which the samples were consolidated into only 42 runs by using tandem mass tag (TMT)-based isobaric labeling.

The isobaric mass tag is comprised of a reactive group to react with peptides, a reporter group linked by an isotopic mass balancer group, of which the bond is prone to cleavage during fragmentation. The nominal mass of each of the isobaric mass tags is identical while the stable heavy isotopes (^{13}C, ^{15}N, or ^{18}O) are distributed differentially between the reporter group and the balancer group (Figure 1.8B). In this regard, isobaric labeled peptides share almost identical liquid chromatography elution profile and is undistinguishable in the MS^1 scan. During fragmentation, reporter ions are released, and the balancer group is lost as a neutral molecule. The intensities of the ensuing reporter ions are used to achieve relative quantifications of samples. It is noteworthy that the reporter ions are in the low *m/z* range; hence, they do not interfere with the sequence-informative b- and y- fragment ions distributed in higher *m/z* range.

Both TMT and isobaric tags for relative and absolute quantitation (iTRAQ) are commonly used in isobaric labeling. In the early 2000s, Thompson and the colleagues[111]

introduced, for the first time, TMT strategy for achieving relative quantification. While only 2-plex was available at the time, the multiplexing capacity has been improved to 16-plex to date, with 6- and 10-plex being the most widely used[112]. In TMT 6-plex, the reporter ions generated are of *m/z* 126, 127, 128, 129, 130, and 131[113]. The 6-plex was later expanded into 10-plex without altering the chemical structure of the tag[114]. Instead, the isotopic mass shift of reporter ions introduced by ^{13}C is resolved from that of ^{15}N by 6.32 mDa. Although the mass difference is extremely small, the lately developed high-resolution mass spectrometers such as Orbitrap are capable of resolving these reporter ions to achieve relative quantification. Recently, Wang et al.[115] demonstrated that combining TMT reagents of different chemical structures, which result in mass shifts of precursor ions, allows for the MS^1-based quantification. Using this approach, the multiplexing capacity could be further increased up to 27-plex, showing that the throughput could be further enhanced by combining the MS^1- and MS^2-based approaches.

The iTRAQ was introduced by Ross et al.[116] as 4-plex shortly after the development of TMT. In 4-plex iTRAQ, the reporter ions have *m/z* of 114, 115, 116, 117, which was later expanded to 8-plex with reporter ions at *m/z* from 113 to 119 and 121[117]. Notably, the reporter ion 120, of which the mass is nearly identical with the immonium ion of phenylalanine, was omitted to prevent distorting the quantification results. Other isobaric labeling techniques include *N*,*N*-dimethyl leucine (DiLeu)[118,119] and deuterium isobaric amine reactive tag (DiART)[120]. In-depth discussion about different isobaric labeling approaches is available elsewhere[121].

An advantage of using isobaric labeling as compared to metabolic labeling such as SILAC is that the samples are combined into a single injection and the labeled peptides co-elute as a single peak in MS^1, allowing for enhanced signal intensity. Because the quantification is achieved by the relative intensities of reporter ions at MS^2 level, low-abundance peptides that would not trigger MS/MS fragmentation can still pass the intensity threshold by the contribution of peptides from other samples. In a recent paper, Lian et al.[122] spiked in 'boosting' samples, which contain peptides of identical amino acid sequences but in much higher quantity, into the sample mixtures to increase the proteome coverage in their TMT-based quantitative phosphoproteomics. The results demonstrated a 4-fold increase in the phosphorylation site coverage, suggesting the robustness of isobaric labeling in assisting the quantification of low-abundance peptides.

While metabolic labeling has been applied to animal studies, the technique is predominantly suitable only for cultured cells. Isobaric labeling, similar with the aforementioned non-isobaric labeling, is immune to the restriction and has been extensively used in clinical protoemics[113,115]. Another advantage of using isobaric labeling is that the incidence of missing values is typically very rare. Given that missing quantification information complicates downstream data processing and often requires imputations for statistical analysis, isobaric labeling allows for more straightforward data interpretation. A recent study demonstrated that more than 99% of proteins are commonly quantified within the same TMT batch with the coefficient of variation (CV) being 3-fold lower than that of label-free quantification[123]. Another study further showed that the approximately 90% of proteins are commonly quantified even in the 24 batches experiment

of 10-plex TMT-labeled samples, highlighting the robustness of isobaric labeling approach in reducing missing values[124].

A well-known limitation of isobaric labeling is ratio compression, where the co-eluting and nearly isobaric peptides contaminate the targeted one, leading to distorted quantification results. Given that precursor ions of similar quantity can generate reporter ions of which the intensities differ by more than 2 orders of magnitude, any high-abundance co-fragmentated peptides could pronouncedly skew the quantification results. To resolve the inherent challenges, McAlister et al.[125] developed isolation waveforms with multiple frequency notches (SPS MultiNotch) MS^3 method, which isolates multiple MS^2 fragment ions for further fragmentation and scan the reporter ions in MS^3 level. Because it is unlikely for the contaminating precursor ion to generate identical fragment ions, the approach allows for the reporter ions to be derived exclusively from the targeted peptide. Alternative approach includes narrowing down the *m/z* window for more effective precursor ion isolation[126]. Another drawback is that isobaric labeling is not suitable for targeted proteomics in general. This is due to the fact that the labeled peptides could not be distinguished in MS^1, which hampers the two-step selection advantage in either MRM or PRM to achieve superior selectivity.

Large-scale quantitative proteomics is becoming indispensable for gaining insights into biological regulations directly at the protein level rather than using mRNA expression as surrogates. Those experiments compare dozens or hundreds of samples. Thus, isobaric labeling enabling multiplexing in LC-MS/MS analysis will allow for acquiring results in a timely fashion. Since their initial development in the early 2000s, isobaric labeling

approach is now extensively used in both basic research and clinical studies. In summary, isobaric labeling-based quantification approach is a powerful technique to multiplex samples and will continue to be widely employed in large-scale quantitative proteomics studies.

1.3.2.2 Stable Isotope-Labeled (SIL) Peptide for SRM

In targeted proteomics, stable isotope-labeled (SIL) peptides of identical amino acid sequence, which is spiked-in prior to sample injection, is increasingly used for relative quantification of an important subset of proteome[127]. The mass spectrometer is programmed to monitor the transition of both the unlabeled sample peptides and the SIL standards. Relative quantification is achieved through normalizing the intensities of fragment ions from the sample peptide to the SIL standards, of which equal amount was spiked-in to different samples. The SIL-based approach effectively overcomes the limitation of label-free quantification of differential instrumental conditions and matrix effects. In addition, the SIL peptides are spiked-in to the digested peptides and consequently do not introduce extra steps into the sample preparation workflow. Similar to the isobaric and non-isobaric labeling approaches, SIL-based quantification could also be adapted to any biological samples.

When choosing the isotopic labeling strategies, ^{13}C, ^{15}N, and ^{18}O are superior to deuterium because they typically do not introduce obvious retention time shifts for peptides, thereby minimizing the quantification bias resulting from differential ionization efficiency. In addition, sufficient mass difference (> 6 Da) should be incorporated into the heavy SIL standards to allow for appropriate mass shifts in doubly or triply charged peptide

ions. This is especially important in MRM analysis operated on a triple-quadrupole instrument, which has low resolving power. In a recent study from our lab, we spiked-in a crude pool of synthetic peptides with a C-terminal [$^{15}N_2$,$^{13}C_6$]-labeled lysine (+8.0 Da) or [$^{15}N_4$,$^{13}C_6$]-labeled arginine (+10 Da) to achieve successful high-throughput MRM-based quantification of small GTPases in brain tissues from Alzheimer disease patients[36], showing the flexibility of SIL-based approach as well as its robustness in quantification accuracy.

Ideally, all peptides should be normalized by its own SIL peptide. However, synthetic peptides are expensive, thus making it cost-ineffective when minor changes are made in the targeted proteomics method. Alternatively, one can rely on surrogate SIL standard peptides that share similar elution profile to the targeted peptide for normalization. Quantification results obtained from surrogate peptides exhibit high accuracy in PRM analysis[128]. In addition, absolute quantification of targeted peptides could also be achieved when the quantity of SIL peptide is available. In the AQUA workflow developed by Gigy and colleagues[127], the absolute protein quantification is determined through their relative intensities to SIL standards in MRM-based analysis.

1.3.3 Data Analysis for Quantitative Proteomics

After acquisition in quantitative discovery proteomics, the raw data are often subjected to search engine for peptide/protein identification and quantification. The quantification methods used in the experiment is specified during search setting. Frequently used methods such as label-free, SILAC, and isobaric labeling are available in most modern data search engines. Currently, MaxQuant[15] is an open resource and has been

routinely used for data search in discovery quantitative proteomics since its first introduction. The recently developed FragPipe[16], which is also made available to the public, has demonstrated tremendous improvement in search speed. A drawback for FragPipe, however, is that the output information is not as comprehensive as compared to MaxQuant. Most mass spectrometer vendors also developed commercial software products, such as Proteome Discoverer from Thermo or PaSER from Bruker. The commercial tools often provide an all-in-one platform for peptide/protein search, downstream analysis, and data visualization.

The raw data generated from different mass spectrometers are in proprietary formats, which support new features of their instruments. Those instrument-specific formats, however, may not be compatible with the data search engine. An alternative approach is to convert the raw data into open spectrum formats (Figure 1.10) through a wide variety of tools currently available such as MSConverter[129]. Currently, Human Proteome Organization (HUPO) suggests Extensible Markup Language (XML)-based open spectrum formats such as mzXML and the newer version mzML for discovery proteomics, which are highly compatible with most modern data search software. Notably, converting raw data into open formats also grants the results to be read by most software without licensing restriction, making sharing data more openly in the community. Different file formats for proteomics are discussed in detail elsewhere[130].

The searched results for quantitative proteomics often require additional analysis to translate the quantification values into unbiased and meaningful biological findings. Such downstream data processing, however, is currently not standardized in the proteomics

field. Data analysis for quantitative proteomics typically requires normalization and statistical correction to remove inherent technical errors introduced during sample preparation or instrument condition. For quantitative proteomics, there are a number of reasons why normalization must be applied to the datasets, including unequal quantities of starting proteins, differential labeling efficiencies, and variations in instrumental conditions (e.g., alterations in ionization efficiency). Just like other omics studies, the ratio between two samples is typically first subjected to $\log_2$ transformation[131]. An obvious advantage of such transformation is that the ratios of peptides or proteins become a continuously distributed spectrum with the up- and down-regulated peptides/proteins being represented in a similar fashion. Without proper transformation, the down-regulated peptides/proteins, which are concentrated between 0 to 1, are treated differently from the up-regulated ones in statistical analysis.

After transformation, normalization is applied to correct for systematic errors. All normalization methods used in quantitative proteomics always make different assumptions. Thus, one should carefully select the methods depending on the nature of the samples as well as the experimental design. Common normalization methods used in MS^1-based quantification include quantile normalization[132,133] which enforces similar distributions of samples. Median/mean normalization[134,135] scales the protein/peptide expression to make all samples have the same median or mean. Valikangas et al.[136] recently evaluated systematically different normalization methods in label-free quantification and found that variance stabilization normalization (Vsn) performed the best in correcting bias introduced by technical errors. Other more complex normalization procedures include MaxLFQ,

which was developed by Cox et al.[103] The MaxLFQ was integrated into the recent versions of MaxQuant and adapted into other data search engines such as FragPipe.

Similar to label-free quantification, isobaric labeling often requires normalization to achieve accurate and reproducible quantification results[124]. Plubell et al.[137] introduced an internal reference scaling (IRS) method that normalizes the proteins/peptides in each sample to achieve same intensity of the selected house-keeping genes, which shows improved normalization as compared to methods relying on global reporter ion intensities. However, the approach may not be sufficient when compositional bias is present in the samples. That is, given that equal amount of starting protein is typically used in most sample preparation procedures, the changes in a subset of abundant proteins can cause systematical quantification bias for low-abundance proteins. Recently, trimmed mean of M values (TMM) approach has been employed in TMT experiment for removing such compositional bias[138-140]. Furthermore, while quantification results are often accurate within the same TMT experiment, the variation becomes obvious in large-scale proteomics experiment that compares multiple TMT batches. In this vein, Brenes et al.[124] introduced a 'reference channel' approach, which includes a relevant internal reference sample across all batches. Their results showed that the quantification accuracy in multiple TMT batches could be significantly corrected.

In some cases, normalization is suggested to be avoided. For instance, in proximity labeling experiments (see section 1.4.2) that investigate protein interaction networks, the assumption that the majority of the proteome in the sample is unaltered, which is adapted in most normalization methods, is not valid[85]. In summary, proper transformation and

normalization are essential for data processing in quantitative proteomics to prevent distorted data and misinterpretation.

1.4 Protein-Protein Interactions

Proteins typically do not work alone to regulate numerous cellular functions. It was estimated that more than 80% of proteins form complex(es) to attain spatiotemporal functions, of which the composition is dynamically regulated by ligand binding or post-translational modifications[141]. For example, G protein-coupled receptors (GPCRs), receptor proteins localized on the plasma membrane, alter their associations with G_{α} subunit along with other signal transduction proteins upon binding with extracellular ligands[142].

Given the importance of protein complexes, perturbation in protein-protein interaction (PPI) is responsible for the progression of numerous diseases. For example, in the brain tissues of amyotrophic lateral sclerosis (ALS) patients, TAR DNA-binding protein 43 (TDP43), an RNA splicing protein, is found to mis-localize to cytosol and aggregate, which subsequently sequesters important cellular proteins such as nuclear pore complex. The distorted interaction of TDP43 with other nuclear pore proteins disrupts the nucleocytoplasmic transport machinery[58,143]. On the other hand, aggregation of TDP43 impairs its ordinary interactions with other RNA-binding proteins, and the subsequent dysregulated RNA splicing was found to be associated with neuronal degeneration[144]. In this regard, it is of great interest to characterize the dynamic protein interaction networks in cells. In this section of introduction, I will introduce the state-of-the-art proteomics

approaches to systematically investigate transient and/or stable protein-protein interactions.

1.4.1 Co-Immunoprecipitation (Co-IP)

Co-IP followed by mass spectrometry analysis is extensively used in investigating protein interaction networks[145]. In co-IP experiments, an antibody is immobilized or captured by protein A/G on a bead to recognize and subsequently purify a member of the protein complex from lysates. During the process, the antibody pulls down the target protein, with which the associated proteins that binds directly or indirectly are co-enriched. Following adequate washing, proteins could be either digested on-beads or eluted for standard sample preparation workflow. Through proteomics analysis, the interaction network of the targeted protein could be identified. Successful enrichment with minimum background is of the essence of co-IP experiment. In this regard, antibody with high specificity and affinity must be carefully chosen and tested up-front. Alternatively, tags such as Flag could be genetically incorporated to be expressed in-frame with the protein of interest. Given that commercial antibodies for the frequently used tags are often in good quality, tagging approach typically allows for clean and effective enrichment.

Proper controls should be incorporated into the experimental design to avoid false discovery. Given that only a small portion of the bead's surface is conjugated with antibody, proteins could be adsorbed non-specifically. Thus, the beads surface should be sufficiently blocked, which is typically achieved with bovine serum albumin (BSA) but could be further strengthened by combining with glycogen and/or DNA. In addition, proper

controls such as cells without tagged protein must be incorporated into the experiment to inform nonspecific background proteome.

Ideally, proteins are enriched at endogenous level to better reveal its native protein-protein interactions. However, this could sometimes be challenging when the expression level of the targeted protein is low, which hinders immunoprecipitation efficiency, thereby decreasing the number of associating proteins that could be confidently identified. To resolve the problem, the protein of interest could be overexpressed to improve the pull-down efficiency. Moreover, the overexpression design could be achieved in tandem with the above-mentioned tagging approach, which allows for sufficient co-immunoprecipitation with high quality of antibody-conjugated beads.

An advantage of employing co-IP to study protein complexes is that the approach could reveal physical protein-protein interactions. In lieu of this, to prevent the false identification of irrelevant proteins that happen to bind to different motif on the same DNA/RNA, the lysate should be treated with nuclease to digest the nucleic acids and allow only proteins with physical interaction to be preserved. One drawback for co-IP approach, however, is the obligatory use of mild binding/washing condition. The condition, although prevents antibody denaturation, often leads to high backgrounds in the proteomics results. In addition, proteins in the water-insoluble fractions such as membrane or stress granules are not suitable for immunoprecipitation as they could not be preserved under mild lysis conditions. Nevertheless, co-IP is an easy to establish workflow for systematical investigation of physical protein-protein interactions and is further discussed elsewhere[146].

1.4.2 Proximity Labeling

The aforementioned antibody-based affinity purification in combination with mass spectrometry is a widely used approach in investigating protein-protein interactions. Although the co-IP method reveals the proteins that physically and stably bind to the protein of interest, the resulting MS data are fundamentally limited by the nonspecific binders arising from mild wash conditions. Moreover, transient or weak protein-protein interactions are often missed since their bindings could not survive the cell lysis or wash steps. To overcome these limitations, scientists have developed complementary approaches based on proximity labeling for better interactome profiling.

Among the emerging proximity labeling methods, ascorbate peroxidase (APEX)-based approach pioneered by Alice Ting and colleagues[85] is one of the most employed (Figure 1.11). The engineered APEX enzyme is genetically encoded to express in frame with the protein of interest. Upon transient H_2O_2 exposure, the enzyme immediately (< 1 min) converts the biotin phenol into a phenoxy radical, which attacks nearly electron-rich amino acids such as tyrosine. In this regard, the proteins within a radius of 10-20 nm to the bait protein is covalently labeled with biotin tags. Those labeled proteins can be purified by using streptavidin-conjugated beads. Given that the biotin-streptavidin interactions are so strong, nonspecific binding proteins can be efficiently removed by using harsh washing conditions (e.g., with a urea-containing buffer).

It should also be noted that the short time frame of APEX labeling enables us to snapshot transient protein-protein interactions. Paek et al.[147] used the APEX-based approach to successfully identify the transient interactome of G-protein-coupled receptors

within as short as 30 seconds upon stimulation. The authors demonstrated the time-resolved interactome of angiotensin II type 1 receptor upon agonist activation and dynamic endocytosis regulations. In addition, the *in situ*-generated short-lived phenoxy radical is membrane-impermeable, allowing for fingerprinting the dynamic subcellular localization of the protein of interest in living cells[148,149]. Recently, Marmor-Kollet et al.[150] employed APEX to investigate the stress granule proteomes with multiple bait proteins. They discovered over a hundred novel stress granule proteins, showing the ability of APEX in profiling spatiotemporal proteomic interaction networks.

Different approaches were proposed to further optimize the APEX-based proximity labeling. Tian and colleagues[151] designed biotin phenol derivatives for better biotinylation in living cells. The analog probe BP5 was shown to generate more active free radicals and label proteins more efficiently. The analog BN2, on the other hand, has weaker labeling activity as compared to the original probe. The authors used BN2 to successfully identify proteins that bind directly with the bait protein due to its smaller labeling radius.

Due to the limited frequency of electron-rich amino acids, the majority of peptides derived from protein-level enrichment do not contain biotin tags, rendering the identification of interactome ambiguous. To improve the identification confidence and remove nonspecific peptides from the samples, biotinylated peptides could be directly enriched for proteomic analysis[152,153]. Through the evidence of biotinylation modification on the peptide, the proximity proteome could be revealed with even higher confidence. In addition, through comparing the biotinylating peptides, alterations of the interaction

surfaces of the associating proteins could also be revealed. Thus, peptide-level enrichment has become an attractive alternative.

Aside from APEX, biotin protein ligase (BirA) from *E. coli* could be fused with the protein of interest to convert the biotin and ATP to biotinyl-AMP ester that reacts with lysine residues in nearby proteins. The engineered BirA was pioneered by Choi-Rhee et al.[154] in the early 2000s through introducing an R118G mutation, which renders its enhanced affinity with biotinyl-AMP, thereby providing improved selectivity for labeling only proteins in close proximity. In 2011, Roux et al.[155] streamlined the $BirA^{R118G}$-based approach for efficient proximity labeling in living cells, termed BioID, coupled with mass spectrometry to systematically investigate protein interaction networks. In their proof-of-concept experiment, they successfully identified most proteins that were previously reported to be in close proximity of lamin-A, a filament protein constituent of the nuclear lamina, suggesting the selectivity and efficiency of the BioID platform in investigating protein-protein interactions. To date, BioID has been applied to identify the biological regulations of important cellular proteins such as E3 ligase and p38 MAPK.

A limitation of BioID approach is the long labeling time (18 h), which limits temporal resolution. To resolve the problem, Ting and colleagues[156] evolved BioID to TurboID through introducing 15 additional mutations. The resulting new enzyme has a 23-fold higher catalytic activity compared to BioID and the labeling time is significantly reduced to 10 min, allowing for capturing transient protein-protein interactions. To reduce potential interference from the relatively large BirA-derived enzymes fused to the targeted proteins, the same group also developed a N-terminal truncated enzyme with 13 mutations,

termed miniTurbo[156]. The smaller enzyme was shown to have just 20% less catalytic activity as compared to TurboID and still achieve sufficient proximity labeling in as short as 10 min. Thus, the later developed TurboID and miniTurbo greatly improve the temporal resolution of BirA-based proximity labeling. Chua et al.[157] employed TurboID to investigate the T cell receptor (TCR)-stimulated protein interaction networks of Lck, a protein that plays an important role during the initiation of TCR signaling pathway. The improved catalytic activity of TurboID allowed for capturing 27 proteins that were in close proximity to Lck during the initial 30 min of TCR signaling.

BirA-based approaches (e.g., BioID and TurboID) do not require H_2O_2 induction and thus are compatible with animal studies. This is especially true when the tissues are too thick for efficient H_2O_2 diffusion. In this vein, Uezu et al.[158] successfully identified synapse protein complex upon neuronal activity inhibition through injecting AAV-packed BirA into the cortex of mouse pups at day 0 postnatal followed by 7 days of daily biotin injection. Their results revealed the first application of BioID in an animal model and successfully identified nearly all the previous reported proteins at the inhibitory postsynaptic density. Recently, Murata et al.[159] developed brain and muscle ARNT-like 1 (BMAL1)-BioID knock-in mice and successfully induced biotinylation through high-biotin diet in multiple tissues including brain, heart, testis, and liver without adverse effects. Those studies demonstrated the applications of BioID in investigating protein-protein interactions in higher model organisms.

Alternatively, proximity labeling could be conducted on dissected tissue samples without genetic engineering of the model animals. Horseradish peroxidase (HRP) is a

peroxidase that also converts biotin phenol to phenoxy radicals as that of APEX enzyme. Notably, although HRP has higher catalytic activity, it is not active in a reducing environment such as cytosol, thus limits its application in live cell labeling[160]. However, through conjugating HRP with an antibody that recognizes the targeted protein, proximity labeling could be achieved for *ex vivo* samples by using a workflow similar to that of APEX-based approach. Bar et al.[161] employed HRP-conjugated antibody to study the proximity proteome of lamin A protein in primary human muscle tissues and successfully identified nuclear envelope proteins, the known binding partners of lamin A. Given the commercial availability of HRP-conjugated primary or secondary antibody, the approach allows for investigating protein-protein interactions of animal models or clinical samples without tedious up-front genetic engineering.

To further advance the capability of proximity labeling and obtain higher spatiotemporal resolutions, the enzymes have been evolved into two inactive fragments that are fused with different proteins. When the two proteins of interest are in close proximity, the two subunits reconstitute into an active peroxidase or biotin ligase. The approach allows for investigating protein-protein interactions in a specific organelle in cells or the protein complex-specific protein-protein interactions. The splitting approach was first introduced by Munter et al.[162] in 2017 and the authors successfully identified the substrates of protein phosphatase 1 (PP1) upon heterodimerization with PP1-interacting proteins (PIPs).

Cho et al.[163] adapted, a few years after, the concept and employed a split-TurboID approach to investigate proteome composition specifically at the endoplasmic reticulum

(ER)-mitochondria contact sites, which is responsible for important biological regulations such as calcium ion signaling. The two spitted fragments were fused to the selected ER and mitochondrial membrane proteins, respectively, and were reconstituted into the active TurboID exclusively at the sites where mitochondria are associated with ER, allowing for investigating the proteome at a very confined cellular region. Similarly, splitting approach has been successfully adapted for APEX-based proximity labeling to achieve better temporal resolutions[164].

Proximity labeling has also been successfully applied to reveal stable or transient RNA-binding proteins. Han et al.[165] fused the engineered APEX to CRISPR-Cas13, an engineered enzyme that recognizes specific RNA through user-designed guide RNA. Consequently, the CRISPR-Cas13-fused APEX labeled the proteins near the RNA motif of interest, and those RNA-binding proteins could be identified through proteomic analysis. In their proof-of-concept investigation, the peroxidase was subsequently targeted to the telomerase RNAs to identify novel binders such as m^6A demethylase ALKBH5, which removes the m^6A modification on telomerase RNA and subsequently regulate telomerase assembly and downstream functions. Alternatively, RNA of interest could be genetically engineered to fuse with sequence encoding the bacteriophage MS2. Through fusing the APEX or biotin ligase with MS2 coat protein (MCP) that binds specifically to MS2, the enzyme could be guided to the RNA for labeling RNA-binding proteins[166].

Taken together, proximity labeling-based approaches such as APEX and BioID in combination with mass spectrometry allow for high-throughput profiling of spatiotemporal protein-protein interactions that were previously difficult to assess, including the transient

or weak interactions. In addition, the incorporation of biotin into the purification steps enables harsh binding/washing conditions to remove non-specific binders, thereby facilitating the discovery of interacting proteins. Collectively, proximity labeling has emerged as a powerful tool in investigating protein-protein interactions. Further information on proximity labeling can be found in recent review articles[167,168].

1.4.3 Cross-Linking

Chemical cross-linking followed by mass spectrometry analysis (XL-MS) emerges as an alternative approach to study protein-protein interactions. The cross-linker forms covalent bonds with different amino acids that are positioned within a defined distance in a protein or protein complex. After cross-linking, protein lysates are subjected to protease digestion followed by enrichment of the ensuing cross-linked peptides for LC-MS/MS analysis (Figure1.12). Notably, the cross-linking condition must be empirically tested for each experiment to achieve sufficient labeling with high specificity.

An advantage of XL-MS approach is the structural information that the method provides. The defined length of the cross-linker constraints the distance of two amino acids that can be cross-linked, thereby revealing their relative positions within a protein complex. Following cross-linking reaction, the labeled peptides could be dead-end (only one end of the cross-linker is covalently bound to an amino acid), intra- or intermolecular cross-linked, with the last one providing the most biological information. On the other hand, the dead-end and intramolecular cross-linked peptides, although does not reveal protein-protein interactions, is a surrogate of solvent accessibility of the amino acids.

Given that the cross-linked products are often of very low abundance in the resulting peptide mixtures, effective enrichment steps prior to MS analysis are often required to achieve sufficient sensitivity. Because cross-linking systematically shifts the peptide to higher molecular weight, they could be separated from the unmodified or dead-end peptides by size exclusion chromatography (SEC)[169]. Schmidt and Sinz[170] demonstrated a streamlined single-step enrichment of the cross-linked products by using C_{18}-strong cation exchange (SCX) StageTips and commercial Oasis mixed-mode cation-exchange (MCX) cartridge, enabling straightforward enrichments without the need of additional instrument. Recently, cross-linker was further implemented with an enrichment handle such as azide, which is used to incorporate an affinity tag (e.g., biotin tag) onto the cross-linked products through click chemistry for affinity purification. In this vein, Petrotchenko et al.[171] demonstrated successful enrichment of cross-linked peptides for MS analysis by using cyanurbiotindipropionylsuccinimide (CBDPS) cross-linker that allows affinity enrichment with avidin.

To identify protein-protein interactions in their native environment, Huang and colleagues[172] developed membrane-permeable probe for labeling protein complexes in living cells. The study provides the foundation for employing XL-MS to study protein-protein interactions *in vivo*. Recently, the XL-MS was further evolved to study the system biological regulations at tissue level. Chavez et al.[173] employed a cross-linking approach to investigate protein-protein interactions in murine heart tissues and the results revealed over two thousand Lys-Lys linkages. In this study, the dissected mouse heart was incubated together with a membrane-permeable cross-linker, and the ensuing cross-linked peptides

were enriched for MS analysis. The results demonstrate the applicability of XL-MS in animal studies.

One major challenge in XL-MS is the subsequent data processing due to the extremely large combinations of potential cross-linked peptides. Although several computational approaches have been developed to resolve the problem, the complexity in searching cross-linked peptides still exponentially increases the data search space[174]. To resolve the computational challenges, Petrotchenko et al.[175] developed cleavable cross-linkers. In such experimental design, the precursor ions are first subjected to soft fragmentation by using CID or ETD to release the two parts of the cross-linked peptides. The released peptides are then further fragmentated and subjected to MS^3 analysis for peptide identification. Since the cross-linked peptides are released into two individual peptides during fragmentation, data analysis could be achieved through regular data search with variable modification settings. In this vein, by combining MS^2 and MS^3 spectra, cross-linked peptides could be successfully identified. Owing to the greatly improved efficiency in data search, MS-cleavage strategy is currently incorporated into most cross-linker designs[172,176].

XL-MS, with the recent advancements in cross-linker design, including cell permeability, affinity purification handles, and the MS-cleavable moiety, becomes another powerful technique to reveal protein interaction networks. Similar to proximity labeling, the formation of covalent bond allows for capturing transient or weak interactions. In addition, XL-MS reveals structural information about how proteins interact. With the advancement of mass spectrometers, cross-linker designs, and computational strategies,

XL-MS is emerging as an orthogonal method in studying systems biology. Further information about XL-MS could be found in recent articles[176,177].

1.4.4 Outlook of Proteomics-Based Approaches in Studying Protein-Protein Interactions

Protein interaction networks are dynamically altered to attain proper regulations. Recent developments in proteomics techniques allow for systematic investigation of protein-protein interactions. IP-MS enables us to capture the native physical interactions. However, such studies often fail to capture transient or weak interactions because they could not survive the sample preparation steps. Proximity labeling emerges as an alternative approach to overcome the limitation. The robustness of proximity labeling is evidenced by numerous findings that reveal comprehensive and insightful biological regulations, which could not be achieved through IP-MS. It is noteworthy that proximity labeling results only indicate that the identified proteins are in close proximity, and whether they interact physically requires further validation. XL-MS that cross-links amino acids of a specific distance could be employed to reveal the structural information of protein-protein interactions. Notably, the results from XL-MS approach are often less comprehensive in interrogating protein interaction networks. However, the approach provides the highest precision and richest information for studying their interactions. Together, proteomics, together with those state-of-the-art approaches, enables systematical exploration of systems biology with high precision and accuracy.

1.5 Small GTPases and Cellular Regulations

Small GTPases (a.k.a. small G-proteins) are low-molecular-weight (15-30 kDa) guanosine triphosphate (GTP)-binding proteins that regulate nearly every aspect of cellular functions in eukaryotic systems. The small GTPases belong to the Ras superfamily, which is termed by the founding members of rat sarcoma (Ras) proto-oncogenes discovered in 1980s[178]. Upon binding to GTP, the small GTPases are in their active states and could interact with downstream effector proteins to achieve cellular regulations. The intrinsic GTPase activity of the proteins in the superfamily, upon the stimulation of GTPase-activating proteins (GAPs), promotes the hydrolysis of GTP into guanosine diphosphate (GDP) and subsequently converts the small GTPases into a GDP-bound inactive state. Currently, the Ras superfamily is further divided into five different families according to their sequence homologies and cellular functions, including Ras, Rho (Ras homology), Arf (ADP-ribosylation factor), Rab (Ras-like in brain), and Ran (Ras-like nuclear) (Figure 1.13A).

For the Ras and Rho families, the inactive small GTPases could be further sequestered by guanine nucleotide inhibitors (GDIs) to prevent them from interacting with regulatory proteins and only allow for reactivation upon specific stimulation. The GDI-mediated prolonged inhibition affords additional regulations of protein activities[179]. For small GTPase reactivation, the bound GDP will be replaced by GTP. However, due to its high binding affinity to guanine nucleotides, small GTPases require guanine nucleotide exchange factors (GEFs) to facilitate the dissociation of GDP to allow for GTP to bind. Notably, small GTPases bind tightly with both GTP and GDP, and GEFs promote the

exchange in both directions. The fact that GEFs could promote the conversion to GTP-bound active state is due to a 10-fold higher cellular concentration of GTP than GDP[179,180]. Moreover, the binding affinity of small GTPases to guanine nucleotides is at the picomolar range, which is far below the micromolar concentrations of GDP and GTP in the cells, allowing the regulations to be resistant from the fluctuations in guanine nucleotide concentrations.

From a structure standpoint, members of the small GTPase superfamily contain the approximately 170 amino acid long G domain, which is defined by five highly conserved G-boxes shared by all small GTPases to coordinate nucleotide binding and GTP hydrolysis (Figure 1.13B). While G domain is shared among all small GTPases, other components of the proteins modulate protein-protein interaction or subcellular localizations to allow for diverse roles in cellular regulations. In this vein, the switch I and switch II, which undergo conformational changes when the small GTPase is associated with GTP, are involved in binding with effector proteins. The N- and C-termini provide further regulations that govern the membrane localization of small GTPases through lipidations such as prenylation, which is crucial for members such as the Rab subfamily that requires membrane attachment to function. In addition, the C-terminal extension of small GTPases is highly diverse in amino acid sequence even within the same subfamily, but could share similar chemical properties such as charges and hydrophobicity[181].

Through switching between the GTP- and GDP-bound states, small GTPases constitute binary switches of cellular signal transductions. In this regard, abnormal expression or activity of small GTPases are found to drive the progression of many

diseases[182]. In this section of introduction, I will discuss the roles and important members of each small GTPase family.

1.5.1 Ras

Ras proteins are the founding members of small GTPase superfamily and by far the most well-characterized family[178]. The small GTPases in the Ras family are molecular switches that turn important signal transduction pathways on and off in response to the stimulation, which results in regulating cell proliferation and differentiation. In mammalian system, HRas, NRas and KRas are the most important members that, when activated, bind to and activate downstream effector proteins such as Raf kinases and PI3K[183]. An important pathway that could be activated by Ras is the RAF-MEK-ERK signaling, which promotes the phosphorylation of important transcription factors such as c-Fos to drive the expression of proliferation-promoting genes[183]. In addition, HRas was shown to activate transcription factor Egr1 also through MEK-ERK pathway, which results in the activation of genes driving neuronal differentiation[184].

Ras proteins are also involved in cell survival and growth. For instance, Ras proteins could bind to PI3K, which leads to the activation of Akt to initiate mammalian targeting of rapamycin (mTOR) kinase-centered signaling. The activation of mTOR leads to cell growth in response to nutrition or growth factor. In healthy cells, the Ras small GTPases are tightly controlled to afford proper cell differentiation and proliferation. However, in cancer cells, Ras small GTPases are often dysregulated, leading to continuous activation of downstream signaling pathways. For instance, G12D substitution in KRas, a

mutation commonly found in lung or pancreatic ductal adenocarcinomas, causes its constant hyperactivation and thus uncontrolled cell proliferation and survival[185].

Together, Ras small GTPases are crucial for cell proliferation, growth, and differentiation. Abnormal expression or activity of the Ras family is associated with tumor developments. Hence, understanding the mechanism of action of the Ras small GTPases, including their regulatory proteins such as GAPs and GEFs, would provide novel insights into the therapeutic developments. Further information about the Ras family could be found elsewhere[186].

1.5.2 Arf

Arf small GTPases, including Arf1 to Arf6, steroidogenic acute regulatory (Sar), and Arf-like (Arl) proteins, are important regulators of the biosynthesis process and the bidirectional vesicular trafficking (i.e., post-Golgi secretory, endocytosis, and endo-lysosomal system). Members in the Arf family contain an additional 14 amino acids at the N-terminus, of which the extension distinguishes the Arf proteins from the other small GTPases families. The sequences of switch I and switch II, which are used for interacting with effector proteins, are almost identical for all Arf small GTPases[187]. It this vein, it is still unclear how different Arf proteins achieve intracellular membrane specificity. Based on the similarity of amino acid sequence, the Arf subfamily could be further divided into three classes: Arf1 and Arf3 in class I (human lacks Arf2), Arf4 and Arf5 in class II, and Arf6 in class III.

Many of the Arf proteins are localized on the Golgi apparatus. Those Golgi-associated Arf small GTPases recruit coat protein 1 (COPI) and adaptor protein 1

(AP1)/clathrin complex to the membrane. COPI is a member of the coat protein complex family. Upon being recruited to the Golgi membrane, COPI assists the formation of vesicles to transport proteins through the endomembrane system. For instance, Arf1 localizes on the cis-Golgi membrane and could form a stable complex with COPI upon GEF activation[188].

In contrast, Arf6 predominately localizes to plasma membrane to regulate endocytosis[189]. A study conducted on Mardin-Darby canine kidney (MDCK) epithelial cells demonstrated that overexpression of SMAD1, a cadherin-binding GAP of Arf6, inhibited the activation of Arf6 and consequently strongly blocked the E-cadherin-mediated membrane internalization[190]. In addition, Arf6 is also involved in plasma membrane recycling. Prigent et al.[191] demonstrated that GTP-bound Arf6 interacts with Sec10 protein to prompt the redistribution of ruffling areas to reorganize plasma membrane. Thus, Arf small GTPases such as Arf6 are active regulators of plasma membrane homeostasis.

In summary, small GTPases in the Arf family are crucial in recruiting membrane coating proteins for the formation of vesicles and the endomembrane trafficking. Interestingly, recent studies also showed that the Arf proteins may coordinate with small GTPases of other subfamily for distinct regulations such as cytoskeleton rearrangement[192]. Strikingly, the Arf small GTPases share high sequence homology and subcellular distribution. The underlying mechanism of how different Arf small GTPases achieve specificity in regulations and cellular functions remains to be further investigated. Detailed

discussion about the Arf family-centered intracellular trafficking regulations could be found in a recent review[193].

1.5.3 Rab

The Rab family is the largest branch in the small GTPase superfamily with approximately 60 members being identified in mammalian cells so far. Similar to Arf small GTPases, the Rab family also play important roles in regulating endomembrane systems such as endoplasmic reticulum, Golgi apparatus, endosomes and lysosomes. Upon activation by GEF proteins, the Rab small GTPases translocate onto specific organelle or vesicles, and recruit effector proteins that modulate membrane trafficking. Interestingly, unlike Arf family that share high sequence homology, the switch regions and the N- or C-terminus of the proteins may differ pronouncedly, resulting in the Rab small GTPases being further divided into different subfamilies.

Rab5 is an endomembrane-regulating Rab small GTPase for endosome biogenesis. Zeigerer et al.[194] demonstrated that depletion of Rab5 in adult mouse liver elicited significant reductions in the numbers of early and late endosomes, along with diminished lysosomes. The distorted endosome formations disrupted the endocytosis of low-density lipoprotein and eventually led to dysfunctional delivery of apical proteins to bile canaliculi. The findings suggest the importance of Rab-centered endomembrane trafficking regulations for normal physiology[194]. Rab23 that expresses in the nerve system, in contrast, localizes to the plasma membrane to regulate membrane internalization and relocation within the cells[195,196]. Furthermore, Eggenschwiler[197] and Anderson demonstrated that Rab23 could regulate vesicle trafficking between the plasma membrane and the cytoplasm

compartment to modulate Hedgehog signal transduction, which is involved in neuronal differentiation.

In addition to the typical Rab proteins, there are also Rab-like (Rabl) small GTPases, including Rabl5 (a.k.a. IFT22), Rabl4 (a.k.a. IFT27), and Rabl2. The Rabl small GTPases do not contain cysteine residues in their C-termini for prenylation, which guides the membrane association of other Rab small GTPases[198]. Instead, the Rabl small GTPases are incorporated into the intraflagellar transport (IFT) complexes to regulate ciliary protein trafficking. Rabl4, for instance, binds to IFT-B upon GEF activation, thereby regulating retrograde cargo transport. The Rabl4-IFT-B complex dissociates to terminate the trafficking regulation when GTP is hydrolyzed into GDP[199].

Altogether, the Rab family is critical for guiding vesicles to the correct destinations in cells. In this vein, dysregulations of Rab proteins have been identified to induce a wide spectrum of disorders such as neurodegeneration[200], inflammatory diseases[201,202], and cancers[203]. Thus, investigating the detailed molecular regulations of the Rab family will lay the foundations of designing therapeutic approaches for many diseases.

1.5.4 Rho

Proteins in the Rho family are crucial in modulating cytoskeleton dynamics to regulate cell migration, adhesion, and division. When activated, Rho small GTPases associate with the membrane through a lipid moiety at their C-termini and activate kinase cascades to regulate cytoskeleton assemblies. The inactive GDP-bound Rho small GTPases could be further sequestered by GDI proteins and pulled into the cytosol for prolonged inactivation in response to cells' needs. Through shuttling between the membrane-bound

active state and the cytosol-localized inactive state, the Rho small GTPases could coordinate cell movements.

In the canonical mechanism of cell migration in polarized cells, the active GTP-bound RhoA and Rac1 are at separated zones. The active RhoA predominantly localizes at the contractile tail to drive cell retractions. Rac1, in contrast, distributes at the leading edges to induce cell protrusions. Together, the RhoA-Rac1 machinery contributes to cell migrations. However, more recent findings suggest that the RhoA is also activated at the leading edges in certain cell types[204]. The active RhoA is spatiotemporally segregated from Rac1 at the edge front. Rac1, along with Cdc42, another Rho small GTPase, is subsequently activated approximately 40 seconds after the reinforcement and stabilization of RhoA-driven newly expanded protrusions[205,206]. The periodic propagating waves of active RhoA and Rac1 coordinating leading and tailing edges of cells during migration. Thus, RhoA-Rac1, together with other members in the Rho family, are the regulatory hubs for cell migration.

Cytoskeletal dynamics is critical in cellular regulations. Hence, it is not surprising that dysregulated expression or activity of Rho small GTPases, e.g., RhoA and Rac1, is associated with the progression of diseases such as tumor formation and neurodegeneration. Interestingly, while dysregulation of Rho small GTPases are found oncogenic, other studies discovered tumor-suppressive effects of some hyperactive Rho proteins. Therefore, the involvement of Rho small GTPases in oncogenesis remains to be further characterized[207,208]. In the central nerve system, Rho small GTPases are involved in regulating neuronal networking and synaptic plasticity[209,210]. Chan et al.[211] demonstrated

that mutations in leucine-rich repeat kinase 2 (LRRK2), one of the most common genetic defects found in familial Parkinson's disease patients, attenuates the activity of Rac1, thus causes actin filaments disassembly and neurite retraction.

Together, the Rho small GTPases are key players of cytoskeleton rearrangement, which is crucial for various physiological regulations such as cell development and podosome formation in many immune cells. Aside from the aforementioned RhoA-Rac1 machinery, many other Rho small GTPases are also master regulators of cell protrusions and migration[208]. Further information about the Rho small GTPases could be found in recent reviews[212,213].

1.5.5 Ran

The Ran small GTPase, which contains only one member, is the most abundant small GTPase in cells and regulates nucleocytoplasmic molecule transport through the nuclear pore complex (NPC)[214]. Ran is an atypical small GTPase in that it does not contain cysteine residue at the C-terminus for lipidation-dependent membrane localization[215]. Although Ran protein also switches between the GTP- and GDP-bound binary states under the regulations of GAP and GEF proteins, different binding states of Ran do not directly translate into protein activation and inactivation as most other small GTPases. Instead, different guanine nucleotide binding states of Ran interact with different effector proteins to regulates the cargo import and export through NPC[216].

RCC1 (Regulator of chromosome condensation 1) is the GEF protein of Ran that mainly localizes in the nucleus, while RanGAP is cytosolic[217,218]. The cellular distribution of GEFs and GAPs for Ran small GTPases creates a gradient of GTP/GDP-binding states

of Ran across the nuclear envelope with higher concentration of GTP- and GDP-bound Ran localizes in the nucleus and cytosol, respectively. The RCC1 protein binds to cargos and transport them from the cytosol into the nucleus through NPC. Once in the nucleus, the GTP-bound Ran binds to importin, causing its structural alteration and consequently helps the cargos to be released from importin[219].

On the other hand, cargos are exported to the cytosol through binding with exportin proteins. The loading of cargos with exportin in the nucleus requires its association with GTP-bound Ran protein. Once in the cytosol, the RanGAP together with Ran-binding protein 1 (RanBP1) promotes GTP hydrolysis and the resulting GDP-bound Ran dissociates from exportin, leading to cargos being unloaded from exportin (Figure 1.14)[219].

Together, Ran protein is a unique member of the small GTPase superfamily. By employing the GTP/GDP-binding binary switches, Ran establishes the cargo loading and unloading system in the cells for nucleocytoplasmic transport. Moreover, Ran protein also binds to different effector proteins such as the RanBP family through its C-terminus to further fine-tune its association with importin and exportin. Further discussion of Ran small GTPases could be found elsewhere[220,221].

1.6 Epitranscriptomic Regulations

DNA undergoes dynamic chemical modifications to alter gene expression. The DNA (cytosine-5)-methyltransferases (DNMTs) catalyze the formation of 5-methylcytosine (5mC) and consequently silences gene expression[222]. Together with histone post-translational modifications and small non-coding RNA (ncRNA)-mediated regulations, the transcriptional regulations are strictly and dynamically regulated in

response to cellular needs such as proliferation and differentiation. Such gene expression regulation without altering the DNA sequence is termed ‘epigenetics’ and has been extensively investigated[223-225].

RNA modifications such as pseudouridine (Ψ) were identified in around the same time as DNA methylation[226]. However, the field only emerged after the advancement of modification-specific sequencing techniques together with the discovery of modification removal proteins such as RNA demethylase fat mass and obesity-associated (FTO) protein that demethylates m^6A[227,228]. Given that the post-transcriptional modifications are also involved in mediating gene expression without altering the RNA sequence, the regulation is consequently coined ‘epitranscriptomics’[229].

In mammalian cells, transfer RNAs (tRNAs), which is the adaptor molecule that brings amino acids to the mRNA and pairs the corresponding codon through its anticodon during protein synthesis[230], are the most extensively modified RNA in the cells[231,232]. The modifications in the tRNA anticodon regions are crucial for its base pairing with messenger RNA (mRNA), thus controls reading frame maintenance and translational efficiency[233]. In addition, those modifications on the tRNA body were found to regulate its metabolism such as stability[234]. In ribosomal RNA (rRNA), the component of the protein translational machinery ribosome complex, post-transcriptional modifications are also found to regulate the translational efficiency of mRNA[235].

In mRNA, the *N*7-methylated guanosine (m^7G) is installed to the first nucleotide of the RNA[236]. The resulting ‘mRNA cap’ is crucial for post-transcriptional processing of mRNAs, including polyadenylation, splicing, and nuclear export[237]. In the cytosol, the m^7G

cap is involved in recruiting translational initiation factors for protein synthesis, thereby regulating translational efficiency[238]. In addition, the m^7G prevents the mRNA from exonuclease-driven RNA cleavage, suggesting its crucial role in RNA stability[239]. Aside from the constitutively installed m^7G, N^6-methyladenosine (m^6A) is the most abundant internal post-transcriptional modification in mRNAs[240]. Recently, m^6A is identified as a reversible mRNA modification that regulates a wide spectrum of RNA metabolisms. Furthermore, the methylation allows for rapid and dynamic regulation of gene expression in response to external stimuli or cellar needs, which is achieved through its reader, writer and eraser proteins[240]. In this section of Introduction, I will briefly discuss the roles of m^6A modification in RNA biology and the ensuing cellular regulations.

1.6.1 m^6A Installation and Removal

In mammalian cells, the m^6A is installed on a consensus motif GAC or AAC in mRNAs and non-coding RNAs with the central adenosine being methylated. In the current paradigm, the methylation is installed by a methyltransferase complex centered by methyltransferase-like 3 (METTL3)/methyltransferase-like 14 (METTL14) heterodimer that was first discovered by He and colleagues[241] approximately a decade ago, along with several other regulatory proteins such as Wilms tumor 1 associated protein (WTAP) (Figure 1.15). Mechanistically, the METTL3-METTL14 heterodimer forms a positively charged groove for RNA binding. METTL3, which binds to the methyl group doner *S*-adenosyl methionine (SAM), catalyzes m^6A deposition on RNA. On the other hand, METTL14 is crucial for structurally supporting METTL3 to recognize RNA substrates and does not carry catalytic activity[242,243].

Aside from METTL3/METTL14 heterodimer, METTL16 was shown to install m^6A methylation onto the *MAT2A* mRNA that encodes SAM synthetase and U6 snRNA on a consensus motif of CAG, with the central adenosine being methylated. The crystal structure of METTL16 together with *MAT2A* mRNA demonstrated that the enzyme recognizes mRNA looping structure through its N-terminus, suggesting that the consensus sequence alone is insufficient for METTL16 recognition, but instead specific secondary structure is required for methylation[244]. Pendleton et al.[245] demonstrated that METTL16, at low SAM concentration, retains on the 3′UTR of *MAT2A* mRNA due to insufficient methyl group doners. Interestingly, the retention of METTL16 on the *MAT2A* transcript promotes its splicing and the subsequent maturation to allow for cells to synthesize more SAM. Given that SAM is the predominant source of methyl group for most methylation in cells, including DNA, RNA, and proteins, it is not surprising that depletion of METTL16 pronouncedly disrupted celluar regulations. Mendel et al.[246] found massive transcriptome dysregulation when *Mettl16* was depleted in mouse embryos, which disrupted embryonic development starting at ~64-cell blastocyst stage. Thus, although METTL16 is currently known to regulate only a few targets, the methyltransferase is indispensable for cell survival and development.

To achieve dynamic regulations, m^6A could be removed by demethylases in cells (Figure 1.15). Around the same time that the METTL3-METTL14 methyltransferase was discovered, the same group also showed that FTO could remove the m^6A methylation both *in vivo* and *in vitro*[227]. In addition, FTO was identified as a nuclear m^6A demethylase that localizes to the nuclear speckles to process certain newly transcribed small ncRNAs and

pre-mRNAs[240]. However, other studies demonstrated that FTO may localize to cytosol in certain neurons, suggesting the function of FTO might be cell type-dependent[247].

Another identified m^{6}A demethylase is the α-ketoglutarate-dependent dioxygenase alkB homologue 5 (ALKBH5), a orthologue of FTO[248]. Similar to FTO, ALKBH5 also localizes to nuclear speckles to regulate mRNA processing. The results further suggests that ALKBH5 could remove m^{6}A both *in vivo* and *in vitro* and regulate RNA transport. In ALKBH5-deficient cells, mRNA was found to be accumulated in the nucleus. Phenotypically, *Alkbh5*$^{-/-}$ mice exhibit defects in spermatogenesis[248]. Interestingly, overexpressing either FTO or ALKBH5 only subtly decreased global m^{6}A modification, suggesting that the two enzymes might have different RNA substrates[247,248].

1.6.2 m^{6}A Reader Proteins and Cellular Functions

The m^{6}A modification could be recognized by different reader proteins for regulations. In addition, the methylation may also alter the RNA structure to modulate RNA-protein interactions. Liu et al.[249] demonstrated that m^{6}A alters the local structure of long non-coding RNA to promote its association with heterogeneous nuclear ribonucleoprotein C (hnRNPC), a mRNA processing protein that localizes in the nucleus. Notably, hnRNPC did not directly recognize m^{6}A. Instead, the binding is regulated by the m^{6}A-induced change in RNA structure. The resulting RNA-hnRNPC interaction promotes RNA maturation, thereby altering gene expression.

YTH domain-containing 1 (YTHDC1) was identified as the major nuclear m^{6}A reader protein. Xiao et al.[250] provided the first evidence of YTHDC1-dependent RNA regulation in the nucleus. Upon binding to m^{6}A-methylated RNA, YTHDC1 recruits SR

family trans-acting splicing factors such as SRSF3 to process pre-mRNA. In their detailed characterizations, the authors further found that YTHDC1 together with SRSF3 promoted exon inclusion of targeted genes.

Recently, YTHDC1 was found to bridge epitranscriptomic and epigenetic regulations. Xu et al.[251] discovered that, in mouse embryonic stem cells (mESCs), YTHDC1 recognizes METTL3-deposited m^6A on the IAPE-z ncRNA. The YTHDC1 interacts with and recruits METTL3 to chromatin, which further recruits H3K9me3 methyltransferase complex SETDB1-TRIM28. The YTHDC1-METTL3 interaction machinery amplifies the regulations and induces hypermethylation of histone H3K9 to form heterochromatin. In line with this, Li et al.[252] demonstrated that repressive histone mark H3K9me2 could be removed by YTHDC1-recruited lysine-specific demethylase 3B (KDM3B), an H3K9me2 demethylase, to initiate gene expression in differentiated cells.

In the cytosol, the m^6A modification is recognized by the YTH domain family proteins, including YTHDF1, YTHDF2, and YTHDF3, of which the interactions regulate RNA metabolisms. The mRNAs and ncRNAs, when bound with YTHDF2, are susceptible to degradation. Wang et al.[253] demonstrated that YTHDF2 recognizes m^6A-modified RNA through its C-terminus. In addition, the transcripts bound by YTHDF2 exhibited increased lifetime in YTHDF2-depleted cells. Shortly after the discovery of YTHDF2 in promoting RNA destabilization, Du et al.[254] further showed that 3′UTR recruits the CCR4-NOT complex to initiate deadenylation of the m^6A-containing mRNAs, thus leading to mRNA decay.

In addition to RNA destabilization, m^6A in mRNA is also involved in regulating protein translational efficiency. Meyer et al.[228] demonstrated that the m^6A modification in 5′-UTR could promote cap-independent translation. The authors further characterized that eukaryotic initiation factor 3 (eIF3) binds to m^6A in 5′-UTR of mRNAs, of which the mRNA-protein interaction is sufficient to recruit the ribosomal complex for translation initiation even without the aforementioned m^7G cap. On the 3′-UTR, Choe et al.[255] showed that METTL3 installs m^6A modification, during which METTL3 physically interacts with protein translation initiation factor eIF3h to promote the mRNA looping for ribosome recycle. Through the m^6A-dependent protein-protein interaction, the 3′-UTR, where the ribosome complex falls off after translation, is brought to close proximity of the 3′-UTR, allowing for increased ribosome occupancy at the translation initiation sites. Similarly, both YTHDF1[256] and YTHDF3[257] were found to bind to m^6A modification at 3′-UTR and promote translation through mRNA tethering.

Together, m^6A regulates a wide spectrum of RNA metabolisms. In the nucleus, m^6A is recognized by reader protein YTHDC1 and may crosstalk with histone modifications to modulate transcription rate. During post-transcriptional processing, m^6A could recruit splicing factors or alter local RNA structure to modulate RNA-protein interactions. In the cytosol, m^6A-modified RNAs could be recognized by YTHDF family reader proteins to regulate the translational efficiency and stability. Future studies about the molecular mechanisms of the m^6A-centered cellular regulations will bring better molecular insights into RNA biology.

1.6.3 Crosstalk Between Epitranscriptomics and Epigenetics Regulations

Epigenetic regulations modulate gene expression without altering the DNA sequences, allowing for further fine-tuning gene transcription beyond its encoded genetic information (e.g., promoter sequence). Currently, epigenome is divided into three arms, including DNA modifications (e.g., methylation), histone PTMs, and the regulation through non-coding RNAs (ncRNAs)[222,223,258]. DNA methylation promotes the condensation of chromatin for prolonged genes silencing[259]. Acetylation of lysine residues on histone proteins, in contrast, neutralizes the positive charges on histone tails, thus open the chromatin structure to increase DNA accessibility to transcription factors that promote gene expression[260]. Furthermore, histone modifications could also be recognized by reader proteins, where they further recruit transcription regulatory proteins.

ncRNAs, including antisense, intergenic, pseudogene and retrotransposon regions, occupy a large quantity of mammalian genome. Those ncRNAs, although not translated into protein products, play important roles in regulating gene expression[258]. For instance, the Xist (X chromosome inactivation is initiated X-inactive-specific transcript) long non-coding RNA coats the chromosome, which subsequently promote heterochromatin formation. Interestingly, independent reports confirmed that epitranscriptomic regulations such as m^6A modification on ncRNAs could crosstalk with epigenetics regulations. Liu et al.[261] discovered that METTL3/METTL14 methyltransferase installed m^6A in chromatin-associated RNAs (caRNAs). Shortly after, Xu et al.[251] found that METTL3, during the installation of m^6A onto caRNAs, recruits methyltransferase SETDB1 and the cofactor TRIM28 to catalyze H3K27me3 of the intracisternal A particle (IAP) elements,

subsequently assist the heterochromatin formation in mouse embryonic stem cells (mESCs). In line with that, Wei et al.[262] found that FTO targets long-interspersed element-1 (LINE1) ncRNA, and the resulting m^6A demethylation promotes opening of nearby chromatin for gene expression in mESCs.

The crosstalk between epitranscriptome and epigenome was also found in the differentiated cells. Xu et al.[263] revealed that m^6A methylation of nascent caRNAs, including promoter upstream transcripts (PROMPTs) and enhancer RNAs (eRNA). The resulting m^6A-methylated caRNAs could be recognized by nuclear reader proteins (i.e., YTHDC1 and hnRNPG), where they protect methylated nascent RNA from Integrator complex-promoted premature termination of transcription. In addition, Li et al.[252] demonstrated that the m^6A reader proteins YTHDC1 could recognize m^6A in newly transcribed mRNA to promote gene expression through recruiting histone demethylase KDM3B to promote the demethylation of histone H3 lysine 9, a silencing epigenetic mark. Thus, m^6A methylations provides an additional controls of ncRNA-dependent epigenetic regulations.

Alternatively, m^6A could crosstalk with epigenetic regulations directly through modulating the metabolisms of mRNAs encoding the RWE proteins for histone modifications. Wang et al.[264] demonstrate that m^6A could result in the destabilization of *Kat3a* and *P300* mRNAs, which encodes histone acetyltransferases, to indirectly regulate gene expressions for cell differentiations. Similarly, Chen et al.[265] showed that METTL3 promotes the protein expression of Ezh2, a methyltransferase catalyzing H3K27me3 in adult mouse NSCs. Depletion of *Mettl3* in mice caused deficient neuronal development

and neurogenesis, which could be rescued by EZH2 overexpression[265]. Together, these pioneering studies point to alternative direction of how epitranscriptome crosstalks with histone PTMs.

Aside from histone modifications, m^6A could also regulate DNA 5-methylcytosine (5mC) methylation to suppress gene expression. Deng et al.[266] discovered an inverse correlation of the m^6A level of co-transcriptional RNAs and the methylation level of their nearby DNA, suggesting that m^6A may guide the demethylation of DNA to promote its transcription through reprogramming chromatin accessibility. Mechanistically, METTL3-deposited m^6A was recognized by reader protein FXR1, which recruits the DNA demethylase TET1 to remove the DNA methylation. Alternatively, m^6A was also reported to directly regulate TET1 expression through the YTHDF2-mediated mRNA decay[267]. Thus, the m^6A-centered epitranscriptomic regulations may be an active player in maintaining proper DNA methylation status in cells prolonged modulations of gene expression.

Together, those emerging studies demonstrate that the m^6A-centered epitranscriptomic regulations could crosstalk with epigenetic regulations through modulating histone or DNA modifications, which could be achieved through either m^6A-methylated ncRNA-dependent recruitment of regulatory proteins and/or directly altering mRNA metabolism.

1.7 Mesenchymal Stem Cell Differentiation

Mesenchymal stem cells (MSCs) are multi-potent and can be induced to differentiate into several cell lineages upon environmental stimulations. Friendenstein and

colleagues[268] discovered, approximately sixty years ago, that a subset of cells from the bone marrow could later generate bone tissues after heterotopic transplantation, suggesting the existence of MSCs in the bone marrow. Shortly after, Pittenger et al.[269] further demonstrated that those MSCs could remain stemness in cell culture and be differentiated, under the induction of differentiation cocktails, into adipocytes, chondrocytes, or osteocytes. Since those pioneering discoveries, MSCs have attracted attention in the stem cell field due to their importance in human physiology and their applications in biomedical developments. Interestingly, MSCs could also be found in various other tissues in the human body such as the adipose[270] and skin tissues[271], while the most characterized ones are still from the bone marrow[272]. In adult, MSCs are the common progenitor cells for both adipocytes and osteoblasts[273]. Therefore, the homeostasis between adipogenic and osteogenic differentiation of MSCs is crucial in maintaining functional physiology[274]. In this section of introduction, I will briefly discuss the molecular pathways regulating the differentiation of MSCs toward adipogenic and osteogenic lineages.

The differentiation of MSCs occurs in a two-step processes, which starts at lineage commitment without morphological changes followed by maturation into fully differentiated cells. In the case of adipogenesis, the committed MSCs are characterized by the expression of platelet-derived growth factor α (PDGFRα)[275,276]. Interestingly, PDGFRα is not expressed in mature adipocytes, suggesting its exclusive roles in regulating early adipogenesis[275]. The committed cells subsequently undergo growth arrest followed by an up-regulation of peroxisome proliferator-activated receptor-γ (PPARγ), the master adipogenesis regulator[277]. Several important upstream adipogenesis-inducing signaling

cascades all merge at the node of up-regulated PPARγ, which further induces the expression of adipocyte genes. Insulin, for instance, binds to the insulin-like growth factor 1 (IGF1) receptor on the cell surface to activate the AKT kinase and mTOR, which leads to elevated PPARγ expression[278,279].

PPARγ is a transcription factor that promotes the expression of genes associated with adipogenesis and metabolism[277]. Lefterova et al.[280] conducted chromatin immunoprecipitation-sequencing (ChIP-seq) and found that PPARγ bound to most genes regulating adipogenesis in 3T3-L1 murine MSCs. Interestingly, the promoters of those genes also harbor the binding motif of CCAAT/enhancer-binding protein (C/EBP), a transcription factor of which the expression is induced by PPARγ during adipogenesis. C/EBP is another important adipogenesis transcription factor that subsequently work in synergy with PPARγ to further reprogram the committed stem cells to fully differentiate into adipocytes. Thus, initiated by PPARγ, the committed cells subsequently coordinate transcriptional regulations to fully differentiate into adipocytes.

On the other hand, the MSCs could also be induced to differentiate into osteoblasts, of which the process is important for the constitutive bone remodeling in adults. Similar to adipogenesis, the MSCs first undergo osteogenic lineage commitment upon environmental stimulations[281]. The osteogenesis-committed MSCs express lineage-specific markers such as runt-related transcription factor 2 (Runx2), the master osteogenesis transcription factor that promotes the expression of osteoblast genes[282]. Runx2 could activate bone matrix protein genes through binding with the osteoblast-specific cis-acting element (OSE) during osteogenesis[283] Consistently, Tarkkonen et al.[284] discovered, through ChIP-seq, that Runx2

is enriched on the promoters of highly expressed genes in osteoblasts derived from both the human and mouse origins. In keeping with that, Takarada et al.[285] demonstrated that depletion of *Runx2* significantly impaired ossification in mice, highlighting its indispensable roles for successful osteogenesis.

Similar to adipogenesis, MSCs first committed to osteogenesis and then further mature into fully differentiated osteoblasts. However, unlike adipogenesis of which the committed stem cells arrest cell proliferation, the osteogenesis-committed MSCs actively proliferate into the pre-osteoblast[281,286]. Interestingly, Runx2 expression also decreases during the maturation process, suggesting that it mainly functions during early osteogenesis[287]. Overexpressing Runx2 at the later stage of osteogenesis, interestingly, inhibited osteoblast maturation[287,288]. In the later-stage of osteogenesis, other transcription factors such as osterix (OSX) take over transcriptional regulations to further modulate osteoblast maturation. Zhu et al.[289] discovered that OSX regulates the expression of extracellular matrix (ECM)-associated genes, which are important for the ossification in mature osteoblast. Thus, the osteogenesis of MSCs is a highly dynamic process requiring temporal molecular regulations throughout osteogenic differentiation[290,291].

Together, MSCs has the plasticity to be differentiated into multiple cell lineages including adipocytes and osteoblasts. Upon different extracellular stimulations, the MSCs express lineage-specific master transcription factors to direct the differentiation process[292]. In this vein, dysregulated adipo-osteogenic balance of MSCs, which could arise from physical, chemical and/or biological factors, is associated with the pathophysiology of numerous diseases such as osteogenesis imperfecta[292,293]. Thus, understanding the

molecular regulations of the differentiation of MSCs toward adipocyte and osteoblast will provide insightful rationales in therapeutic approach development.

1.8 Scope of the Dissertation

Advances in bottom-up proteomics in various state-of-the-art sample preparation methods allow for investigation of biological regulations with high specificity, sensitivity, and throughput. We set out to employ MRM-based targeted proteomics to systemically investigate small GTPase expression and to identify novel regulatory proteins in MSC differentiation. In addition, we interrogated the molecular regulations of epitranscriptome, including the protein-protein interactions of the m^6A methyltransferase and the involvements of m^6A RWE proteins in modulating the expression of small GTPase superfamily of proteins.

In Chapter 2, we utilized unbiased MRM-based targeted proteomics coupled with scheduled LC and synthetic stable isotope-labeled peptides to systematically quantify the expression of small GTPase superfamily proteins upon the adipogenesis of murine MSCs. We found that Rab32 was down-regulated during the adipogenic differentiation of both 3T3-L1 and C3H10T1/2 murine MSCs. Overexpressing Rab32 prior to adipogenic induction effectively suppressed the adipogenesis rate, suggesting that Rab32 is an important adipogenesis regulator, of which its diminished expression is a pre-requisite for successful differentiation.

MSCs are multi-potent stem cells and can be differentiated into various lineages aside from adipocyte. Thus, we speculated that the MSCs may systemically regulate the expression of small GTPase superfamily to fine-tune the differentiation process. In Chapter

3, we employed the MRM-based targeted proteomic platform to further investigate the differential expression of small GTPases during the course of osteogenic differentiation of H9 human embryonic stem cells. We found altered expression of a large number of small GTPases accompanied with osteogenic differentiation, especially those involved in autophagy. In addition, we discovered KRAS as a novel osteogenesis regulator, where it regulates the accumulation of extracellular matrix for mineralization through attenuating the activity of secreted matrix metalloproteinase-9 (MMP9).

The m^6A methylation in mRNA regulates its metabolism such as stability and translational efficiency. In Chapter 3, we employed the MRM-based targeted proteomic platform to comprehensively examine the differential expression of small GTPase proteins in cells depleted of METTL3 (the catalytic subunit of m^6A methyltransferase complex), m^6A demethylases ALKBH5 and FTO, and the m^6A reader proteins including YTHDF1, YTHDF2, or YTHDF3. Interestingly, the small GTPases of which the expression is regulated by METTL3 and/or ALKBH5 displayed higher discrepancy between the protein and mRNA expression in paired primary/metastatic melanoma or colorectal cancer cells than those that are not, suggesting that the epitranscriptomic regulations may be involved in their protein expression. We further discovered that the elevated expression of RhoB in *METTL3*$^{-/-}$ cells is due to the increased mRNA stability, suggesting a m^6A-elicited destabilization of the transcript.

To our knowledge, the protein-protein interactions of the METTL3-METTL14 heterodimer, the crucial subunits of the major m^6A methyltransferase complex, remain largely unexplored. In Chapter 4, we employed state-of-the-art APEX-based proximity

labeling, followed by LC-MS/MS analysis, to systematically investigate the interaction proteins of METTL3 and METTL14 in HEK293T cells. We found that a subset of histone RWE proteins are enriched in the proximity proteome of METTL3, including histone acetyltransferase 1 (HAT1). Additionally, we observed that METTL3-calatyzed formation of m^6A in chromatin-associated RNAs (caRNAs) promotes the chromatin occupancy of the long α isoform of HAT1 (HAT1α), thereby stimulating the acetylation of histone H3 at Lys-9 (H3K9Ac). Moreover, we found that the function of HAT1α on H3K9 acetylation requires its nuclear localization and is promoted by YTHDC1, an m^6A reader protein in the nucleus that also interacts with METTL3.

Together, the applications of bottom-up proteomics in conjugation with the advancements of quantitative approaches and sample preparation methods (e.g., APEX-based proximity labeling) allows for systematically investigating biological regulations. In addition, we envision that the biological discoveries reported in this dissertation, including the roles of small GTPase superfamily in MSC differentiation and the crosstalk of epitranscriptome and epigenome, will lay a strong foundation for future investigations in cellular regulations.

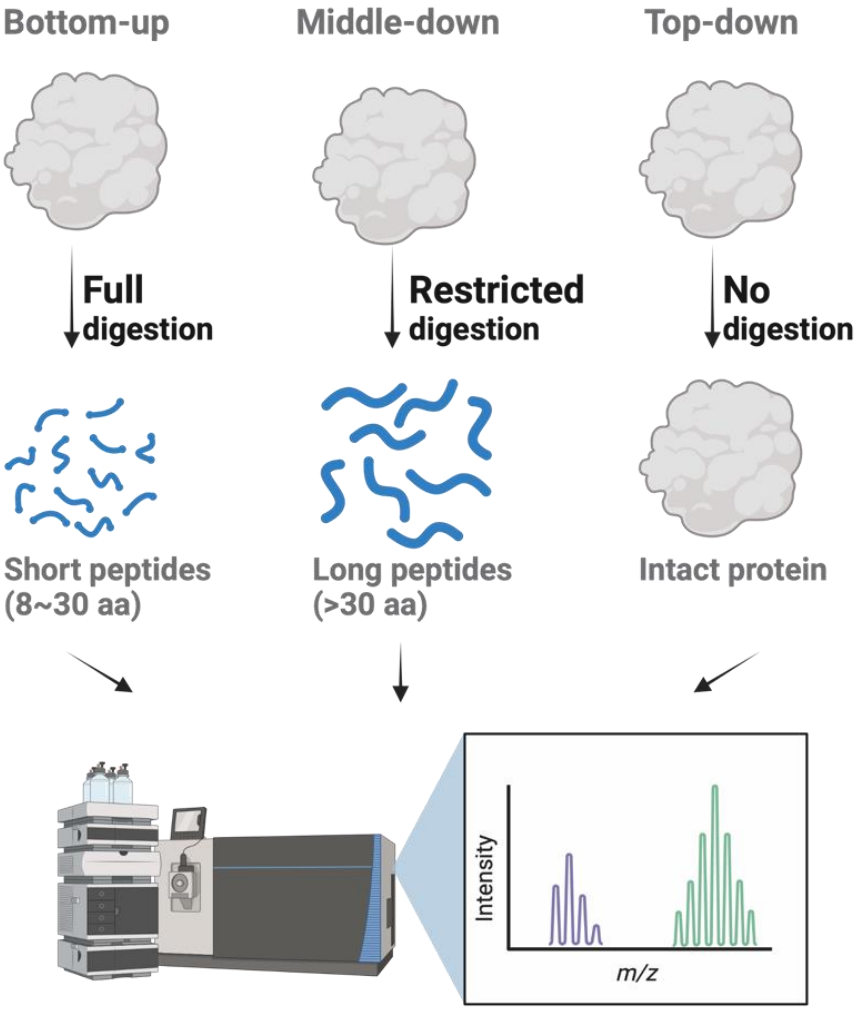


Figure 1.1 Overview of the workflows for common proteomics approaches, including bottom-up, middle-down, and top-down.

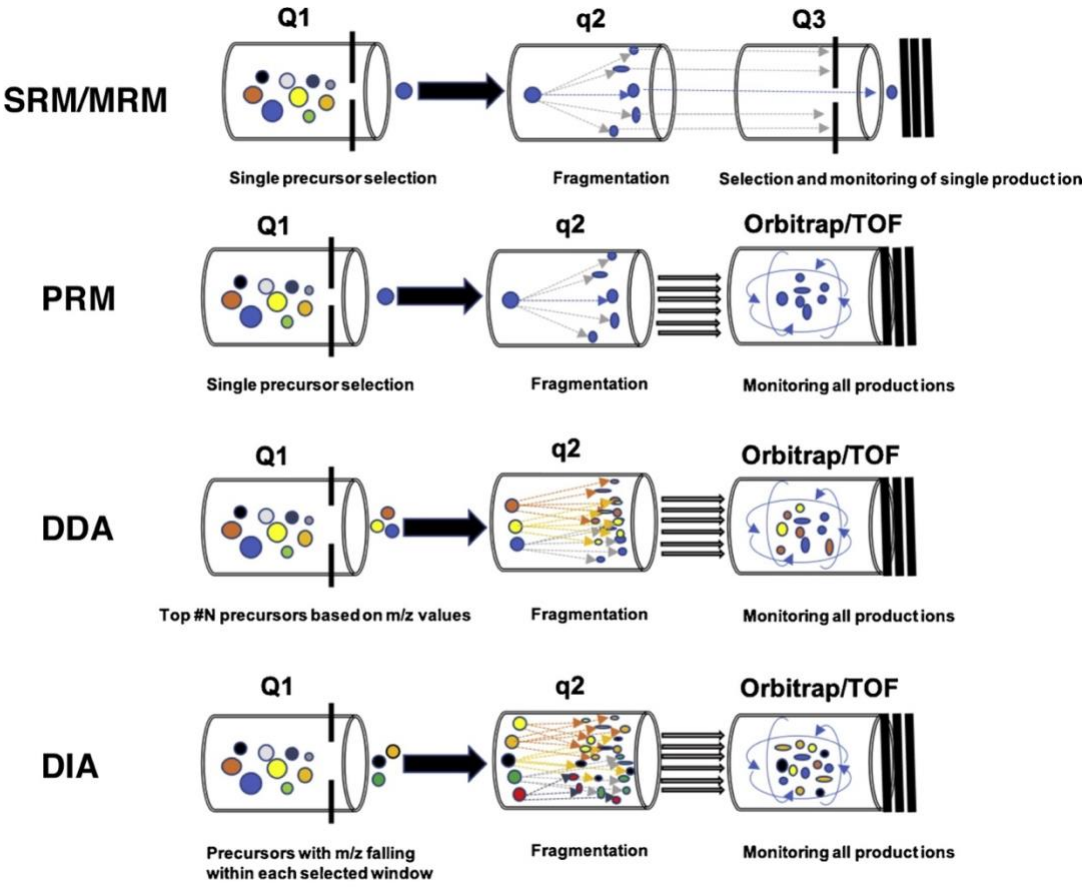


Figure 1.2 A schematic diagram illustrating different bottom-up proteomics approaches, including DDA, DIA, PRM and MRM (a.k.a. SRM).

The first quadrupole mass analyzer is depicted as Q1. The collision cell is depicted as q2. Q3 represents the second quadrupole mass analyzer in a triple-quadrupole mass spectrometer.

The figure is adapted from Li, J. *et al.*, Data-independent acquisition (DIA): An emerging proteomics technology or analysis of drug-metabolizing enzymes and transporters. *Drug Discov. Today Technol.* 39 (2021).

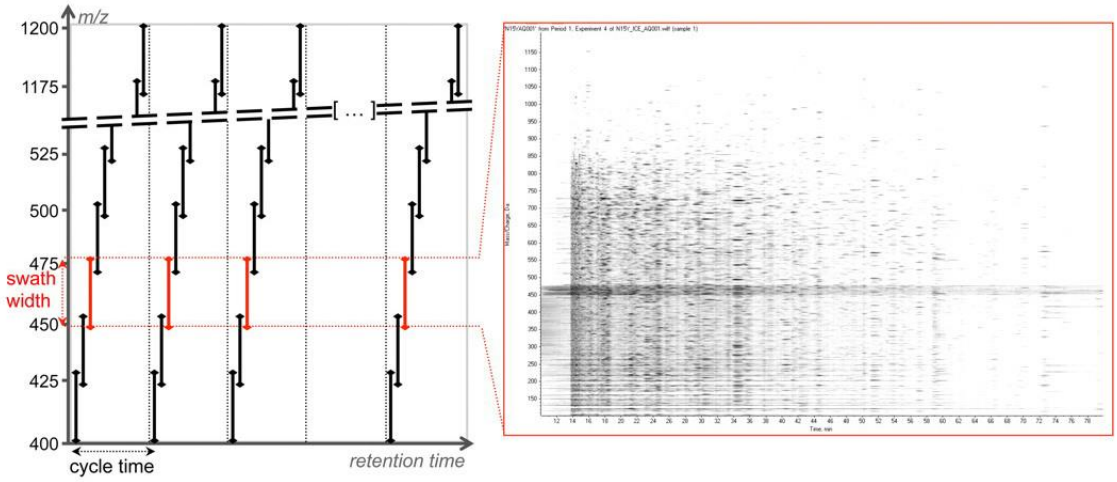


Figure 1.3 A schematic diagram of SWATH-based DIA acquisition.

The MS^2 scanning is cycled through 32 isolation widths of 25-Th each (double arrows) across the *m/z* range of 400 to 1200.

The figure is adapted from Gillet 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* 11, 6 (2012).

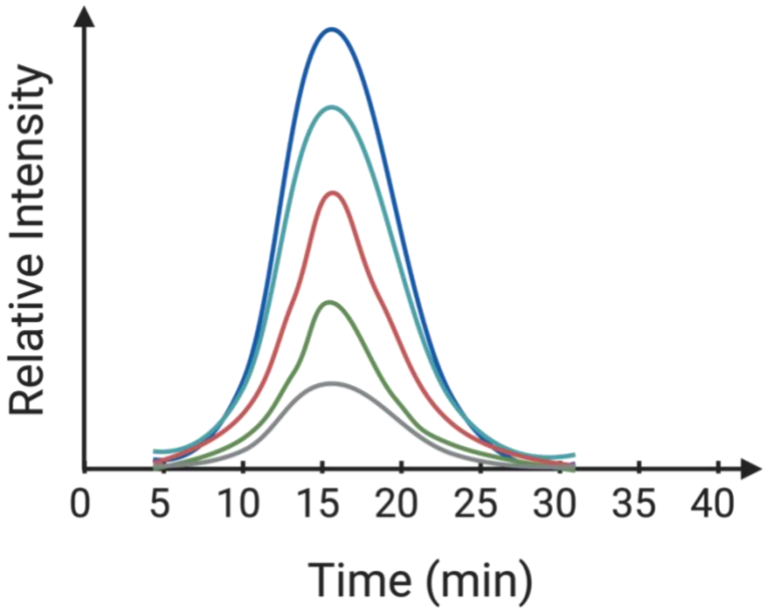


Figure 1.4 Concurrent transitions in targeted proteomics.

A schematic diagram showing the extracted-ion chromatogram (XIC) of concurrent fragment ions monitored in targeted proteomics analysis.

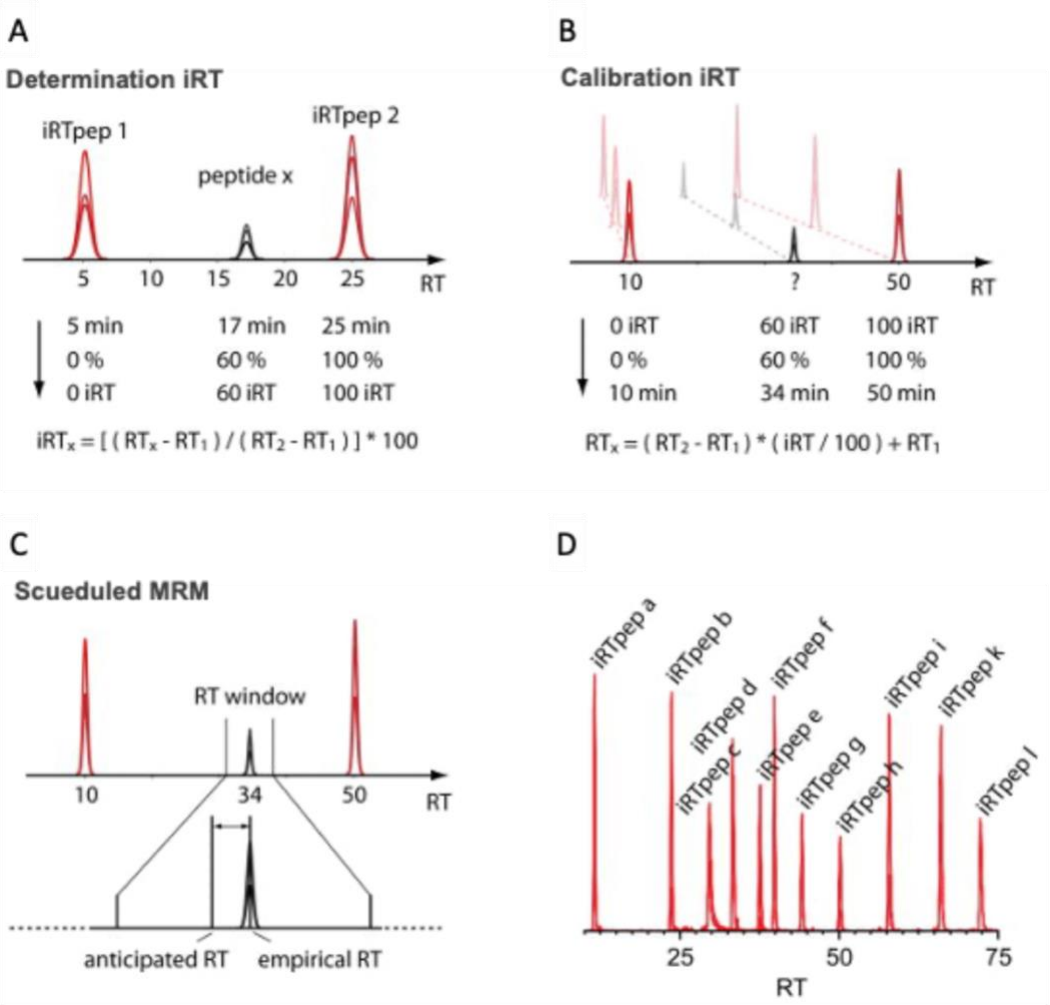


Figure 1.5 Experimental design of scheduled LC.

(A) Two of the standard peptides (i.e., iRTpep1 and iRTpep2) are given iRT values as 0 and 100, respectively. The iRT of the peptide of interest is assigned based on its relative retention time to the two selected peptides. (B) In the calibration run of standard peptides, the iRT-RT linear relationship is derived to predict the RT of the peptide of interest. (C-D) A schematic diagram of the schedule LC implemented in MRM analysis.

The figure is adapted and modified from Escher et al., Using iRT, a normalized retention time for more targeted measurement of peptides. *Proteomics* 12, 8 (2012).

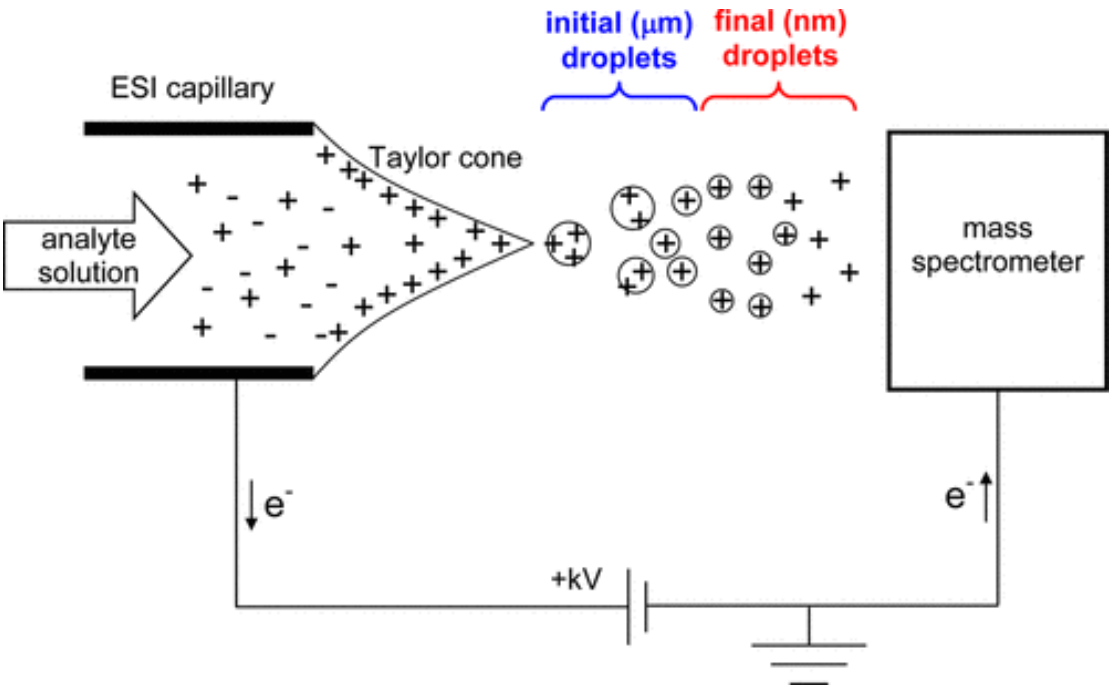


Figure 1.6 A schematic diagram showing the process of nanoelectrospray ionization operated in the positive-ion mode.

The figure is adapted from Konermann et al., Unraveling the mechanism of electrospray ionization. *Anal. Chem.* 85, 1 (2012).

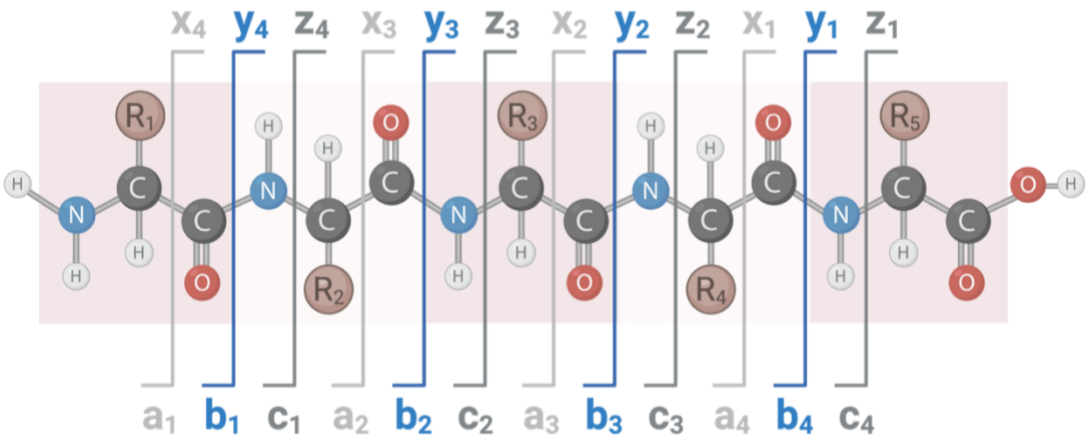


Figure 1.7 A schematic diagram showing bond cleavages accompanied with peptide fragmentation.

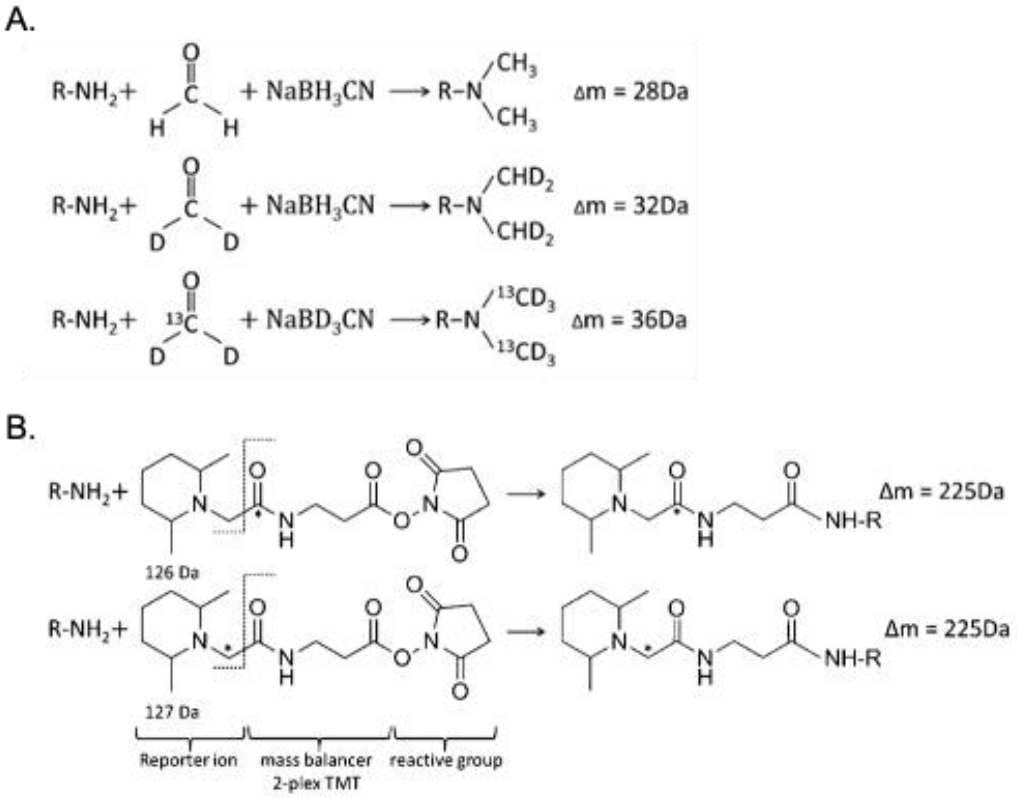


Figure 1.8 Labeling reactions that target free amino groups.

(A) Triplex dimethyl labeling that induces mass shift of precursor for MS^1-based quantification. (B) Duplex TMT-based isobaric labeling. The probe is composed of reporter ion, reactive group, and a mass balancer. The bond susceptible to cleavage during fragmentation to release the reporter ion is indicated by a dashed line.

The figure is adapted from Zhang et al., Protein analysis by shotgun/bottom-up proteomics. *Chem. Rev.* 113 (2013).

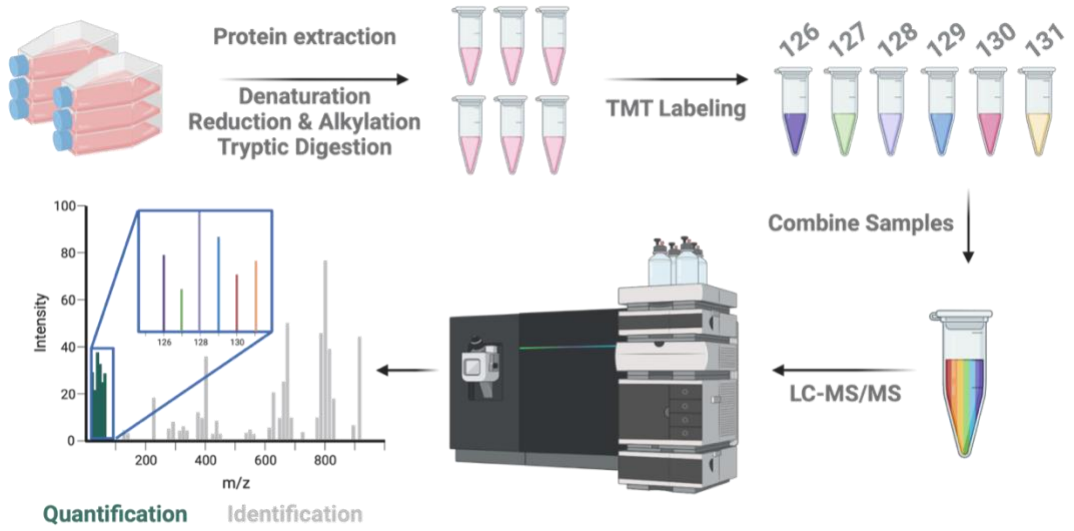


Figure 1.9 A schematic diagram showing workflow of TMT-based quantitative proteomics.

(A) Proteins extracted from cells were digested by trypsin, and the resulting peptides were labeled by 6-plex TMT reagents. (B) Precursor ion is isolated for fragmentation. The low mass range reporter ions fall off and are monitored at MS2 scan for quantification.

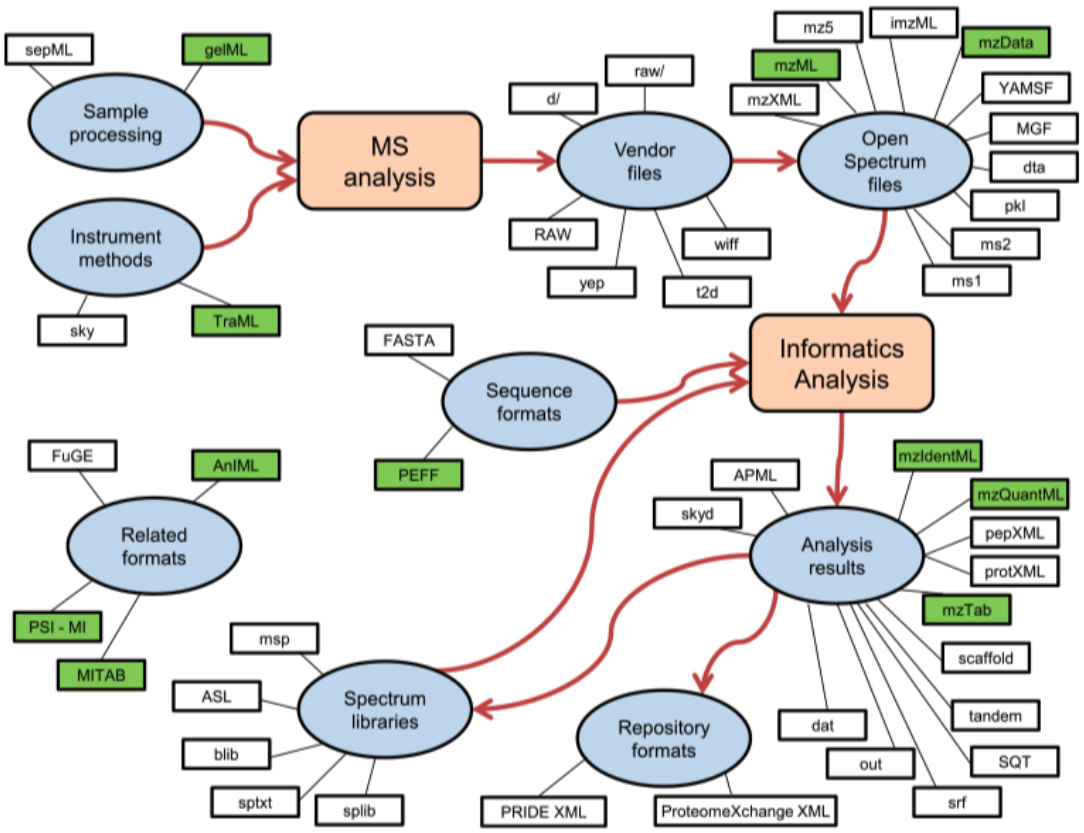


Figure 1.10 Overview of common formats of proteomics data.

The purple ovals represent the data type categories and the associated rectangle represent the currently available file formats. Those rectangle boxes colored in green represents the well-recognized formats.

The figure is adapted from Deutsch E.W., File formats commonly used in mass spectrometry. *Mol. Cell. Proteomics.* 11, 12 (2012).

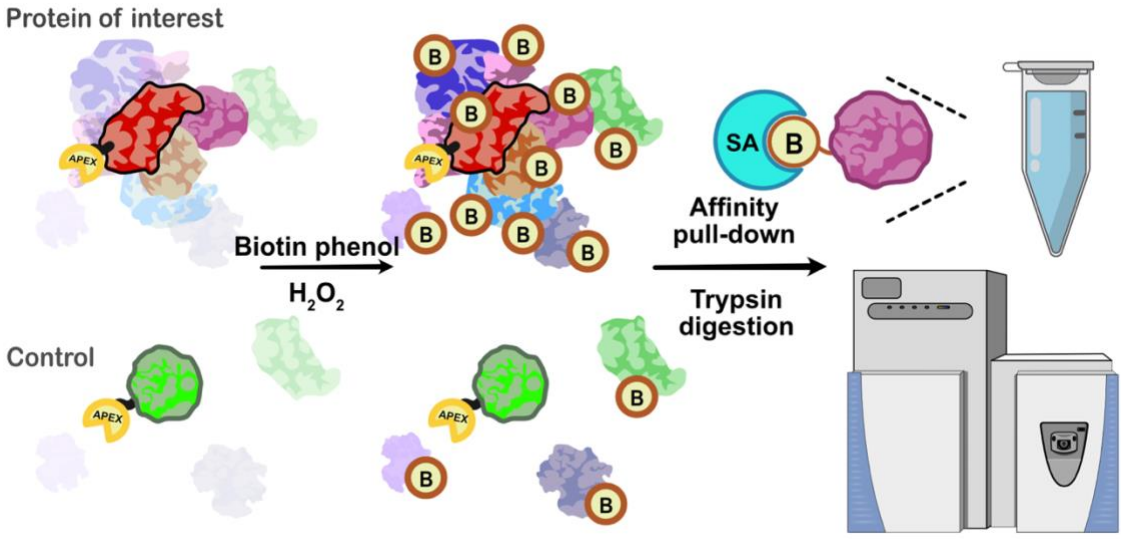


Figure 1.11 A schematic diagram showing the APEX-based proximity labeling approach.

The protein of interest is fused with the engineered APEX. Following biotin phenol feeding and hydrogen peroxide induction, proteins in close proximity are biotinylated. Affinity purification with harsh binding and washing conditions is employed to enrich the labeled proteins with high specificity.

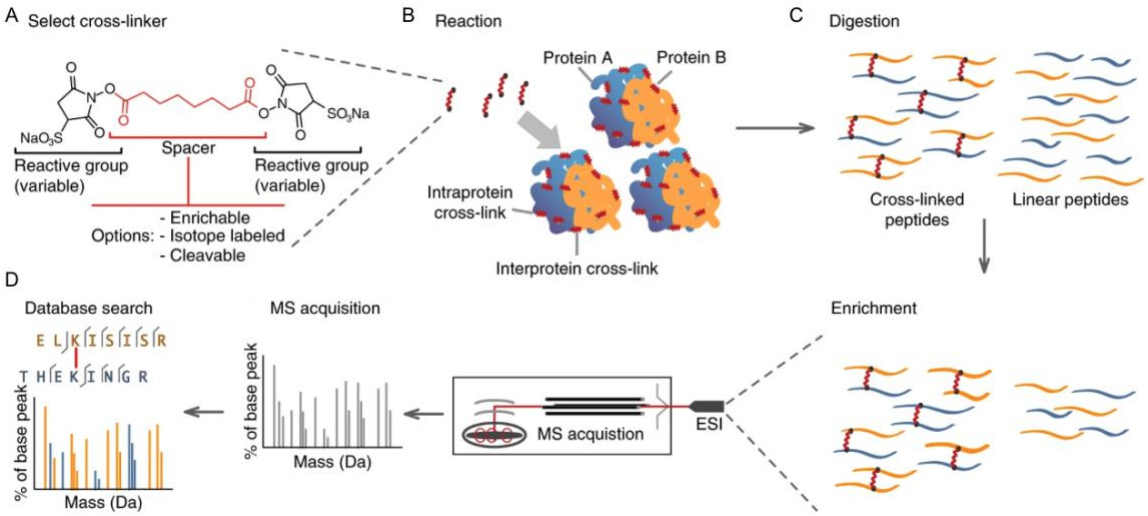


Figure 1.12 Workflow of protein cross-linking followed by LC-MS/MS analysis.

(A) A cross-linker that contains reactive moiety on both ends for reacting with amino acids. A spacer that comprises of user-selected length and chemical prosperities joins the two reactive group. (B) Amino acids within the pre-defined length are cross-linked. (C-D) Following cross-linking reaction, the proteins are digested into peptides and the cross-linked products are enriched for LC/MS-MS analysis.

The figure is adapted from Oreilly F.J., and Rappsilber J. Cross-linking mass spectrometry: methods and applications in structural, molecular and systems biology. *Nat. Struct. Mol. Biol.* 25 (2018).

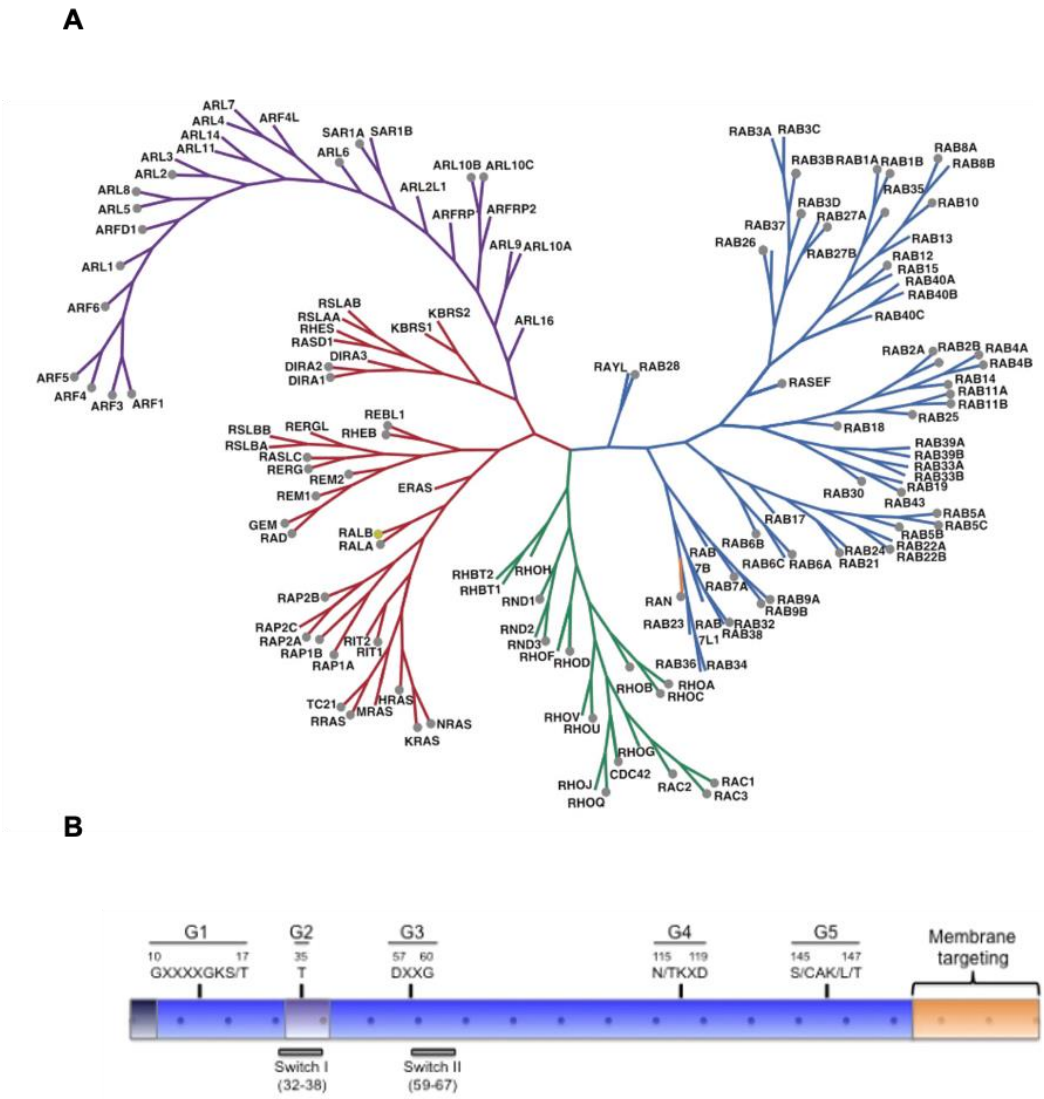


Figure 1.13 Evolution and structure of small GTPases.

(A) Phylogenetic tree of the Ras superfamily. The blue, green, red, purple, and orange braches represent the Rab, Rho, Ras, Arf and the Ran small GTPase families, respectively. (B) A diagram depicting the general domain structure of small GTPase proteins.

The figure A is adapted from Gray L. J. et al., Targeting the small GTPase superfamily through their regulatory proteins. *Angew. Chem. Int. Ed. Engl.* 16, 59 (2020).

The figure B is adapted from Wennerberg K. and Der C. J., Rho-family GTPases: it's not only Rac and Rho (and I like it). *J. Cell Sci.* 15, 117 (2004).

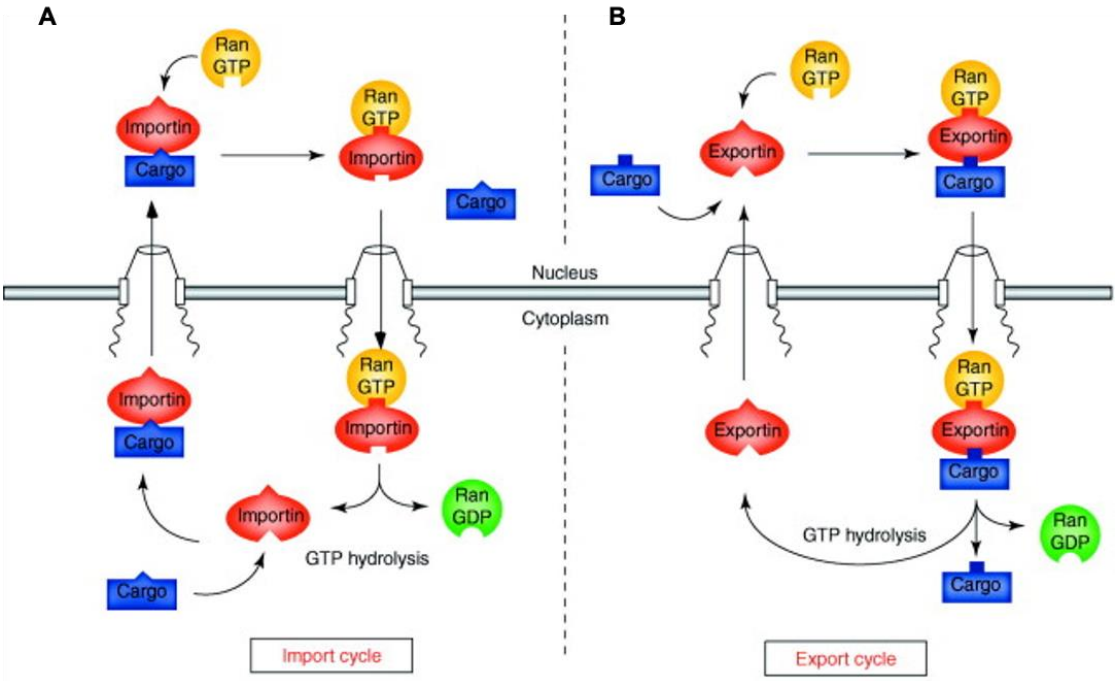


Figure 1.14 A schematic diagram showing the canonical Ran-centered nucleocytoplasmic molecule transport.

(A) Importin binds to the cargo in the cytosol and transports it into the nucleus. In the nucleus, the GTP-bound Ran binds to importin, leading to the release of cargo and the recycle of importin into the cytosol. (B) To transport molecule out from the nucleus, exportin loads the cargo upon binding to GTP-bound Ran. Once in the cytosol, the GTP is hydrolyzed into GDP and the Ran-Exportin-cargo complex dissociates.

The figure is adapted from Scott K. et al., Nucleocytoplasmic transport: Ran, beta and beyond. *Trends Cell Biol.*11, 12 (2001).

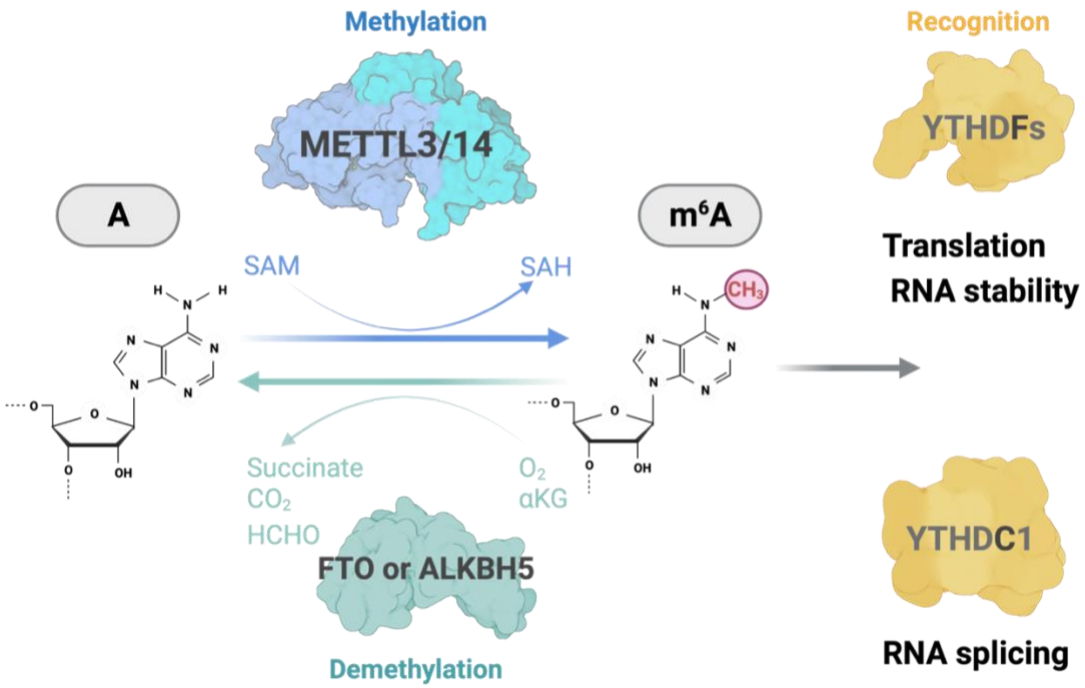


Figure 1.15 A schematic diagram showing m^{6}A methylation on RNA.

The adenosine at the center of GAC or AAC motif could be recognized by the METTL3/METTL14 methyltransferase complex to form m^{6}A modification. The methylation could be removed by demethylases, including FTO and ALKBH5. The m^{6}A-modified RNA could be recognized by reader proteins to further modulate RNA metabolisms such as translational efficiency or stability.

1.9 References

1. Pruitt, K.D., Tatusova, T. & Maglott, D.R. NCBI reference sequences (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. *Nucleic Acids Res.* **35**, D61-5 (2007).

2. Smith, L.M. & Kelleher, N.L. Proteoform: a single term describing protein complexity. *Nat. Methods* **10**, 186-7 (2013).

3. Aebersold, R. et al. How many human proteoforms are there? *Nat. Chem. Biol.* **14**, 206-214 (2018).

4. Cagney, G., Amiri, S., Premawaradena, T., Lindo, M. & Emili, A. In silico proteome analysis to facilitate proteomics experiments using mass spectrometry. *Proteome Sci.* **1**, 5 (2003).

5. Latterich, M., Abramovitz, M. & Leyland-Jones, B. Proteomics: new technologies and clinical applications. *Eur. J. Cancer* **44**, 2737-41 (2008).

6. Ferkingstad, E. et al. Large-scale integration of the plasma proteome with genetics and disease. *Nat. Genet.* **53**, 1712-1721 (2021).

7. Cai, W. et al. Top-down proteomics of large proteins up to 223 kDa enabled by serial size exclusion chromatography strategy. *Anal. Chem.* **89**, 5467-5475 (2017).

8. Gault, J. et al. Combining native and 'omics' mass spectrometry to identify endogenous ligands bound to membrane proteins. *Nat. Methods* **17**, 505-508 (2020).

9. Eng, J.K., McCormack, A.L. & Yates, J.R. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *J. Am. Soc. Mass Spectrom.* **5**, 976-89 (1994).

10. Wilson, S.R., Vehus, T., Berg, H.S. & Lundanes, E. Nano-LC in proteomics: recent advances and approaches. *Bioanalysis* **7**, 1799-815 (2015).

11. Kalli, A. & Hess, S. Effect of mass spectrometric parameters on peptide and protein identification rates for shotgun proteomic experiments on an LTQ-orbitrap mass analyzer. *Proteomics* **12**, 21-31 (2012).

12. Wong, C.C., Cociorva, D., Venable, J.D., Xu, T. & Yates, J.R., 3rd. Comparison of different signal thresholds on data dependent sampling in Orbitrap and LTQ mass spectrometry for the identification of peptides and proteins in complex mixtures. *J. Am. Soc. Mass Spectrom.* **20**, 1405-14 (2009).

13. Zhang, Y., Wen, Z., Washburn, M.P. & Florens, L. Effect of dynamic exclusion duration on spectral count based quantitative proteomics. *Anal. Chem.* **81**, 6317-26 (2009).

14. Pelletier, A.R. et al. MealTime-MS: A machine learning-guided real-time mass spectrometry analysis for protein identification and efficient dynamic exclusion. *J. Am. Soc. Mass Spectrom.* **31**, 1459-1472 (2020).

15. Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc.* **11**, 2301-2319 (2016).

16. Kong, A.T., Leprevost, F.V., Avtonomov, D.M., Mellacheruvu, D. & Nesvizhskii, A.I. MSFragger: ultrafast and comprehensive peptide identification in mass spectrometry-based proteomics. *Nat. Methods* **14**, 513-520 (2017).

17. Orsburn, B.C. Proteome discoverer-A community enhanced data processing suite for protein informatics. *Proteomes* **9**, 15 (2021).

18. Elias, J.E. & Gygi, S.P. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat. Methods* **4**, 207-14 (2007).

19. Sinitcyn, P., Rudolph, J.D. & Cox, J. Computational methods for understanding mass spectrometry–based shotgun proteomics data. *Annu. Rev. Biomed. Data Sci.*, 207-234 (2018).

20. Bruderer, R. et al. Extending the limits of quantitative proteome profiling with data-independent acquisition and application to acetaminophen-treated three-dimensional liver microtissues. *Mol. Cell. Proteomics* **14**, 1400-10 (2015).

21. Gillet, L.C. 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* **11**, O111 016717 (2012).

22. Krasny, L. & Huang, P.H. Data-independent acquisition mass spectrometry (DIA-MS) for proteomic applications in oncology. *Mol. Omics* **17**, 29-42 (2021).

23. Zhang, F., Ge, W., Ruan, G., Cai, X. & Guo, T. Data-independent acquisition mass spectrometry-based proteomics and software tools: a glimpse in 2020. *Proteomics* **20**, 1900276 (2020).

24. Hu, A., Noble, W.S. & Wolf-Yadlin, A. Technical advances in proteomics: new developments in data-independent acquisition. *F1000Research* **5**(2016).

25. Koopmans, F., Ho, J.T.C., Smit, A.B. & Li, K.W. Comparative analyses of data independent acquisition mass spectrometric approaches: DIA, wisim-DIA, and untargeted DIA. *Proteomics* **18**, 1700304 (2018).

26. Ting, Y.S. et al. PECAN: library-free peptide detection for data-independent acquisition tandem mass spectrometry data. *Nat. Methods* **14**, 903-908 (2017).

27. Rost, H.L. et al. OpenSWATH enables automated, targeted analysis of data-independent acquisition MS data. *Nat. Biotechnol.* **32**, 219-23 (2014).

28. Egertson, J.D. et al. Multiplexed MS/MS for improved data-independent acquisition. *Nat. Methods* **10**, 744-6 (2013).

29. Egertson, J.D., MacLean, B., Johnson, R., Xuan, Y. & MacCoss, M.J. Multiplexed peptide analysis using data-independent acquisition and Skyline. *Nat. Protoc.* **10**, 887-903 (2015).

30. Lange, V., Picotti, P., Domon, B. & Aebersold, R. Selected reaction monitoring for quantitative proteomics: a tutorial. *Mol. Syst. Biol.* **4**, 222 (2008).

31. Percy, A.J., Parker, C.E. & Borchers, C.H. Pre-analytical and analytical variability in absolute quantitative MRM-based plasma proteomic studies. *Bioanalysis* **5**, 2837-56 (2013).

32. Huang, M., Qi, T.F., Li, L., Zhang, G. & Wang, Y. A targeted quantitative proteomic approach assesses the reprogramming of small GTPases during melanoma metastasis. *Cancer Res.* **78**, 5431-5445 (2018).

33. Kisiala, A., Kambhampati, S., Stock, N.L., Aoki, M. & Emery, R.J.N. Quantification of cytokinins using high-resolution accurate-mass orrbitrap mass spectrometry and parallel reaction monitoring (PRM). *Anal. Chem.* **91**, 15049-15056 (2019).

34. Bourmaud, A., Gallien, S. & Domon, B. Parallel reaction monitoring using quadrupole-Orbitrap mass spectrometer: Principle and applications. *Proteomics* **16**, 2146-59 (2016).

35. Rauniyar, N. Parallel reaction monitoring: a targeted experiment performed using high resolution and high mass accuracy mass spectrometry. *Int. J. Mol. Sci.* **16**, 28566-81 (2015).

36. Huang, M., Darvas, M., Keene, C.D. & Wang, Y. Targeted quantitative proteomic approach for high-throughput quantitative profiling of small GTPases in brain tissues of Alzheimer's disease patients. *Anal. Chem.* **91**, 12307-12314 (2019).

37. Qi, T.F., Miao, W. & Wang, Y. Targeted profiling of epitranscriptomic reader, writer, and eraser proteins accompanied with radioresistance in breast cancer cells. *Anal. Chem.* **94**, 1525-1530 (2022).

38. Escher, C. et al. Using iRT, a normalized retention time for more targeted measurement of peptides. *Proteomics* **12**, 1111-21 (2012).

39. Gallien, S. et al. Highly multiplexed targeted proteomics using precise control of peptide retention time. *Proteomics* **12**, 1122-33 (2012).

40. Gallien, S., Bourmaud, A., Kim, S.Y. & Domon, B. Technical considerations for large-scale parallel reaction monitoring analysis. *J. Proteomics* **100**, 147-59 (2014).

41. Sobsey, C.A. et al. Targeted and untargeted proteomics approaches in biomarker development. *Proteomics* **20**, e1900029 (2020).

42. Wilm, M. & Mann, M. Analytical properties of the nanoelectrospray ion source. *Anal. Chem.* **68**, 1-8 (1996).

43. Smoluch, M.T., Sallberg, M. & Silberring, J. Nanospray mass spectrometry for identification of peptides. Application of a novel interface. *Acta. Biochim. Pol.* **46**, 475-81 (1999).

44. Yin, R., Burnum-Johnson, K.E., Sun, X., Dey, S.K. & Laskin, J. High spatial resolution imaging of biological tissues using nanospray desorption electrospray ionization mass spectrometry. *Nat. Protoc.* **14**, 3445-3470 (2019).

45. Mostafa, M.E., Grinias, J.P. & Edwards, J.L. Supercritical Fluid Nanospray Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **33**, 1825-1832 (2022).

46. Hillenkamp, F. & Karas, M. Mass spectrometry of peptides and proteins by matrix-assisted ultraviolet laser desorption/ionization. *Methods Enzymol.* **193**, 280-95 (1990).

47. Preisler, J., Hu, P., Rejtar, T., Moskovets, E. & Karger, B.L. Capillary array electrophoresis-MALDI mass spectrometry using a vacuum deposition interface. *Anal. Chem.* **74**, 17-25 (2002).

48. Ma, B. & Johnson, R. De novo sequencing and homology searching. *Mol Cell Proteomics* **11**, O111 014902 (2012).

49. Michalski, A., Neuhauser, N., Cox, J. & Mann, M. A systematic investigation into the nature of tryptic HCD spectra. *J. Proteome Res.* **11**, 5479-91 (2012).

50. Swaney, D.L. et al. Supplemental activation method for high-efficiency electron-transfer dissociation of doubly protonated peptide precursors. *Anal. Chem.* **79**, 477-85 (2007).

51. Good, D.M., Wirtala, M., McAlister, G.C. & Coon, J.J. Performance characteristics of electron transfer dissociation mass spectrometry. *Mol. Cell. Proteomics* **6**, 1942-51 (2007).

52. Frese, C.K. et al. Improved peptide identification by targeted fragmentation using CID, HCD and ETD on an LTQ-Orbitrap Velos. *J. Proteome Res.* **10**, 2377-88 (2011).

53. Zubarev, R.A. & Makarov, A. Orbitrap mass spectrometry. *Anal. Chem.* **85**, 5288-96 (2013).

54. Aballo, T.J. et al. Ultrafast and reproducible proteomics from small amounts of heart tissue enabled by azo and timsTOF pro. *J. Proteome Res.* **20**, 4203-4211 (2021).

55. Sandow, J.J., Infusini, G., Dagley, L.F., Larsen, R. & Webb, A.I. Simplified high-throughput methods for deep proteome analysis on the timsTOF Pro. *bioRxiv*, 657908 (2021).

56. Meier, F. et al. diaPASEF: parallel accumulation-serial fragmentation combined with data-independent acquisition. *Nat. Methods* **17**, 1229-1236 (2020).

57. Kelly, R.T. Single-cell proteomics: progress and prospects. *Mol. Cell. Proteomics* **19**, 1739-1748 (2020).

58. Chou, C.C. et al. TDP-43 pathology disrupts nuclear pore complexes and nucleocytoplasmic transport in ALS/FTD. *Nat. Neurosci.* **21**, 228-239 (2018).

59. Norris, J.L., Porter, N.A. & Caprioli, R.M. Mass spectrometry of intracellular and membrane proteins using cleavable detergents. *Anal. Chem.* **75**, 6642-7 (2003).

60. Yu, Y.Q., Gilar, M., Lee, P.J., Bouvier, E.S. & Gebler, J.C. Enzyme-friendly, mass spectrometry-compatible surfactant for in-solution enzymatic digestion of proteins. *Anal. Chem.* **75**, 6023-8 (2003).

61. Muller, T. & Winter, D. Systematic eevaluation of protein reduction and alkylation reveals massive unspecific side effects by iodine-containing reagents. *Mol. Cell. Proteomics* **16**, 1173-1187 (2017).

62. Vandermarliere, E., Mueller, M. & Martens, L. Getting intimate with trypsin, the leading protease in proteomics. *Mass. Spectrom. Rev.* **32**, 453-65 (2013).

63. Tsiatsiani, L. & Heck, A.J. Proteomics beyond trypsin. *FEBS J.* **282**, 2612-26 (2015).

64. Leon, I.R., Schwammle, V., Jensen, O.N. & Sprenger, R.R. Quantitative assessment of in-solution digestion efficiency identifies optimal protocols for unbiased protein analysis. *Mol. Cell. Proteomics* **12**, 2992-3005 (2013).

65. Shevchenko, A., Wilm, M., Vorm, O. & Mann, M. Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal. Chem.* **68**, 850-8 (1996).

66. Shevchenko, A., Tomas, H., Havlis, J., Olsen, J.V. & Mann, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* **1**, 2856-60 (2006).

67. Goodman, J.K., Zampronio, C.G., Jones, A.M.E. & Hernandez-Fernaud, J.R. Updates of the in-gel digestion method for protein analysis by mass spectrometry. *Proteomics* **18**, e1800236 (2018).

68. Kim, Y.I. & Cho, J.Y. Gel-based proteomics in disease research: Is it still valuable? *Biochim. Biophys. Acta. Proteins. Proteom.* **1867**, 9-16 (2019).

69. Lu, X. & Zhu, H. Tube-gel digestion: a novel proteomic approach for high throughput analysis of membrane proteins. *Mol. Cell. Proteomics* **4**, 1948-58 (2005).

70. Wisniewski, J.R., Zougman, A., Nagaraj, N. & Mann, M. Universal sample preparation method for proteome analysis. *Nat. Methods* **6**, 359-62 (2009).

71. Sielaff, M. et al. Evaluation of FASP, SP3, and iST protocols for proteomic sample preparation in the low microgram range. *J. Proteome Res.* **16**, 4060-4072 (2017).

72. Ludwig, K.R., Schroll, M.M. & Hummon, A.B. Comparison of in-solution, FASP, and S-Trap based digestion methods for bottom-up proteomic studies. *J. Proteome Res.* **17**, 2480-2490 (2018).

73. Ong, S.E. et al. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* **1**, 376-86 (2002).

74. Wheatley, D.N., Scott, L., Lamb, J. & Smith, S. Single amino acid (arginine) restriction: growth and death of cultured HeLa and human diploid fibroblasts. *Cell. Physiol. Biochem.* **10**, 37-55 (2000).

75. Ong, S.E., Kratchmarova, I. & Mann, M. Properties of 13C-substituted arginine in stable isotope labeling by amino acids in cell culture (SILAC). *J. Proteome Res.* **2**, 173-81 (2003).

76. Lossner, C., Warnken, U., Pscherer, A. & Schnolzer, M. Preventing arginine-to-proline conversion in a cell-line-independent manner during cell cultivation under stable isotope labeling by amino acids in cell culture (SILAC) conditions. *Anal. Biochem.* **412**, 123-5 (2011).

77. Ong, S.E. & Mann, M. A practical recipe for stable isotope labeling by amino acids in cell culture (SILAC). *Nat. Protoc.* **1**, 2650-60 (2006).

78. Wu, C.C., MacCoss, M.J., Howell, K.E., Matthews, D.E. & Yates, J.R., 3rd. Metabolic labeling of mammalian organisms with stable isotopes for quantitative proteomic analysis. *Anal. Chem.* **76**, 4951-9 (2004).

79. Kruger, M. et al. SILAC mouse for quantitative proteomics uncovers kindlin-3 as an essential factor for red blood cell function. *Cell* **134**, 353-64 (2008).

80. Sury, M.D., Chen, J.X. & Selbach, M. The SILAC fly allows for accurate protein quantification in vivo. *Mol. Cell. Proteomics* **9**, 2173-83 (2010).

81. Konzer, A. et al. Stable isotope labeling in zebrafish allows in vivo monitoring of cardiac morphogenesis. *Mol. Cell. Proteomics* **12**, 1502-12 (2013).

82. Geiger, T., Cox, J., Ostasiewicz, P., Wisniewski, J.R. & Mann, M. Super-SILAC mix for quantitative proteomics of human tumor tissue. *Nat. Methods* **7**, 383-385 (2010).

83. Jaffe, J.D. et al. Global chromatin profiling reveals NSD2 mutations in pediatric acute lymphoblastic leukemia. *Nat. Genet.* **45**, 1386-91 (2013).

84. Choi, S., Goswami, N. & Schmidt, F. Comparative proteomic profiling of 3T3-L1 adipocyte differentiation using silac quantification. *J. Proteome Res.* **19**, 4884-4900 (2020).

85. Hung, V. et al. Spatially resolved proteomic mapping in living cells with the engineered peroxidase APEX2. *Nat. Protoc.* **11**, 456-75 (2016).

86. Schwanhausser, B., Gossen, M., Dittmar, G. & Selbach, M. Global analysis of cellular protein translation by pulsed SILAC. *Proteomics* **9**, 205-9 (2009).

87. Tzouros, M. et al. Development of a 5-plex SILAC method tuned for the quantitation of tyrosine phosphorylation dynamics. *Mol. Cell. Proteomics* **12**, 3339-49 (2013).

88. Cox, J. et al. A practical guide to the MaxQuant computational platform for SILAC-based quantitative proteomics. *Nat. Protoc.* **4**, 698-705 (2009).

89. Chen, X., Wei, S., Ji, Y., Guo, X. & Yang, F. Quantitative proteomics using SILAC: Principles, applications, and developments. *Proteomics* **15**, 3175-92 (2015).

90. Sethuraman, M. et al. Isotope-coded affinity tag (ICAT) approach to redox proteomics: identification and quantitation of oxidant-sensitive cysteine thiols in complex protein mixtures. *J. Proteome Res.* **3**, 1228-33 (2004).

91. Gygi, S.P. et al. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **17**, 994-9 (1999).

92. Hsu, J.L., Huang, S.Y., Chow, N.H. & Chen, S.H. Stable-isotope dimethyl labeling for quantitative proteomics. *Anal. Chem.* **75**, 6843-52 (2003).

93. Boersema, P.J., Raijmakers, R., Lemeer, S., Mohammed, S. & Heck, A.J. Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nat. Protoc.* **4**, 484-94 (2009).

94. Boersema, P.J., Aye, T.T., van Veen, T.A., Heck, A.J. & Mohammed, S. Triplex protein quantification based on stable isotope labeling by peptide dimethylation applied to cell and tissue lysates. *Proteomics* **8**, 4624-32 (2008).

95. Paulo, J.A. & Gygi, S.P. mTMT: An alternative, nonisobaric, tandem mass tag allowing for precursor-based quantification. *Anal. Chem.* **91**, 12167-12172 (2019).

96. Kang, U.B., Yeom, J., Kim, H. & Lee, C. Quantitative analysis of mTRAQ-labeled proteome using full MS scans. *J. Proteome Res.* **9**, 3750-8 (2010).

97. Yin, H.R. et al. Hyperplex-MRM: a hybrid multiple reaction monitoring method using mTRAQ/iTRAQ labeling for multiplex absolute quantification of human colorectal cancer biomarker. *J. Proteome Res.* **12**, 3912-9 (2013).

98. Oppermann, F.S. et al. Comparison of SILAC and mTRAQ quantification for phosphoproteomics on a quadrupole orbitrap mass spectrometer. *J. Proteome Res.* **12**, 4089-100 (2013).

99. Chelius, D. & Bondarenko, P.V. Quantitative profiling of proteins in complex mixtures using liquid chromatography and mass spectrometry. *J. Proteome Res.* **1**, 317-23 (2002).

100. Liu, H., Sadygov, R.G. & Yates, J.R., 3rd. A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal. Chem.* **76**, 4193-201 (2004).

101. Griffin, N.M. et al. Label-free, normalized quantification of complex mass spectrometry data for proteomic analysis. *Nat. Biotechnol.* **28**, 83-9 (2010).

102. Bantscheff, M., Schirle, M., Sweetman, G., Rick, J. & Kuster, B. Quantitative mass spectrometry in proteomics: a critical review. *Anal. Bioanal. Chem.* **389**, 1017-31 (2007).

103. Cox, J. et al. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell. Proteomics* **13**, 2513-26 (2014).

104. Schwanhausser, B. et al. Global quantification of mammalian gene expression control. *Nature* **473**, 337-42 (2011).

105. Arike, L. et al. Comparison and applications of label-free absolute proteome quantification methods on Escherichia coli. *J. Proteomics* **75**, 5437-48 (2012).

106. Al Shweiki, M.R. et al. Assessment of Label-Free Quantification in Discovery Proteomics and Impact of Technological Factors and Natural Variability of Protein Abundance. *J. Proteome Res.* **16**, 1410-1424 (2017).

107. Bostanci, N. et al. Targeted proteomics guided by label-free quantitative proteome analysis in saliva reveal transition signatures from health to periodontal disease. *Mol. Cell. Proteomics* **17**, 1392-1409 (2018).

108. Li, Z. et al. Systematic comparison of label-free, metabolic labeling, and isobaric chemical labeling for quantitative proteomics on LTQ Orbitrap Velos. *J. Proteome Res.* **11**, 1582-90 (2012).

109. Neilson, K.A. et al. Less label, more free: approaches in label-free quantitative mass spectrometry. *Proteomics* **11**, 535-53 (2011).

110. Nusinow, D.P. et al. Quantitative Proteomics of the Cancer Cell Line Encyclopedia. *Cell* **180**, 387-402 e16 (2020).

111. Thompson, A. et al. Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal. Chem.* **75**, 1895-904 (2003).

112. Li, J. et al. TMTpro reagents: a set of isobaric labeling mass tags enables simultaneous proteome-wide measurements across 16 samples. *Nat. Methods* **17**, 399-404 (2020).

113. Dayon, L. et al. Relative quantification of proteins in human cerebrospinal fluids by MS/MS using 6-plex isobaric tags. *Anal. Chem.* **80**, 2921-31 (2008).

114. McAlister, G.C. et al. Increasing the multiplexing capacity of TMTs using reporter ion isotopologues with isobaric masses. *Anal. Chem.* **84**, 7469-78 (2012).

115. Wang, Z. et al. 27-plex tandem mass tag mass spectrometry for profiling brain proteome in alzheimer's disease. *Anal. Chem.* **92**, 7162-7170 (2020).

116. Ross, P.L. et al. Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* **3**, 1154-69 (2004).

117. Choe, L. et al. 8-plex quantitation of changes in cerebrospinal fluid protein expression in subjects undergoing intravenous immunoglobulin treatment for Alzheimer's disease. *Proteomics* **7**, 3651-60 (2007).

118. Frost, D.C., Greer, T. & Li, L. High-resolution enabled 12-plex DiLeu isobaric tags for quantitative proteomics. *Anal. Chem.* **87**, 1646-54 (2015).

119. Xiang, F., Ye, H., Chen, R., Fu, Q. & Li, L. N,N-dimethyl leucines as novel isobaric tandem mass tags for quantitative proteomics and peptidomics. *Anal. Chem.* **82**, 2817-25 (2010).

120. Zhang, J., Wang, Y. & Li, S. Deuterium isobaric amine-reactive tags for quantitative proteomics. *Anal. Chem.* **82**, 7588-95 (2010).

121. Rauniyar, N. & Yates, J.R., 3rd. Isobaric labeling-based relative quantification in shotgun proteomics. *J. Proteome Res.* **13**, 5293-309 (2014).

122. Yi, L. et al. Boosting to amplify signal with isobaric labeling (basil) strategy for comprehensive quantitative phosphoproteomic characterization of small populations of cells. *Anal. Chem.* **91**, 5794-5801 (2019).

123. O'Connell, J.D., Paulo, J.A., O'Brien, J.J. & Gygi, S.P. Proteome-Wide Evaluation of Two Common Protein Quantification Methods. *J. Proteome Res.* **17**, 1934-1942 (2018).

124. Brenes, A., Hukelmann, J., Bensaddek, D. & Lamond, A.I. Multibatch TMT Reveals False Positives, Batch Effects and Missing Values. *Mol. Cell. Proteomics* **18**, 1967-1980 (2019).

125. McAlister, G.C. et al. MultiNotch MS3 enables accurate, sensitive, and multiplexed detection of differential expression across cancer cell line proteomes. *Anal. Chem.* **86**, 7150-8 (2014).

126. Savitski, M.M. et al. Delayed fragmentation and optimized isolation width settings for improvement of protein identification and accuracy of isobaric mass tag quantification on Orbitrap-type mass spectrometers. *Anal. Chem.* **83**, 8959-67 (2011).

127. Gerber, S.A., Rush, J., Stemman, O., Kirschner, M.W. & Gygi, S.P. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 6940-5 (2003).

128. Qi, T.F. et al. Targeted quantitative profiling of epitranscriptomic reader, writer, and eraser proteins using stable isotope-labeled peptides. *Anal. Chem.* **94**, 12559-12564 (2022).

129. Chambers, M.C. et al. A cross-platform toolkit for mass spectrometry and proteomics. *Nat. Biotechnol.* **30**, 918-20 (2012).

130. Deutsch, E.W. File formats commonly used in mass spectrometry proteomics. *Mol. Cell. Proteomics* **11**, 1612-21 (2012).

131. Quackenbush, J. Microarray data normalization and transformation. *Nat. Genet.* **32 Suppl**, 496-501 (2002).

132. Bolstad, B.M., Irizarry, R.A., Astrand, M. & Speed, T.P. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19**, 185-93 (2003).

133. Brombacher, E., Schad, A. & Kreutz, C. Tail-robust quantile normalization. *Proteomics* **20**, e2000068 (2020).

134. Callister, S.J. et al. Normalization approaches for removing systematic biases associated with mass spectrometry and label-free proteomics. *J. Proteome Res.* **5**, 277-86 (2006).

135. Kultima, K. et al. Development and evaluation of normalization methods for label-free relative quantification of endogenous peptides. *Mol. Cell. Proteomics* **8**, 2285-95 (2009).

136. Valikangas, T., Suomi, T. & Elo, L.L. A systematic evaluation of normalization methods in quantitative label-free proteomics. *Brief. Bioinform.* **19**, 1-11 (2018).

137. Plubell, D.L. et al. Extended multiplexing of tandem mass tags (TMT) labeling reveals age and high fat diet specific proteome changes in mouse epididymal adipose tissue. *Mol. Cell. Proteomics* **16**, 873-890 (2017).

138. Klann, K., Krause, D. & Münch, C. DynaTMT: A user-friendly tool to process combined SILAC/TMT data. *bioRxiv*, 2021.07.06.451268 (2021).

139. Robinson, M.D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11**, R25 (2010).

140. Yan, K., Yang, Y., Zhang, Y., Zhao, W. & Liao, L. Normalization method utilizing endogenous proteins for quantitative proteomics. *J. Am. Soc. Mass Spectrom.* **31**, 1380-1388 (2020).

141. Berggard, T., Linse, S. & James, P. Methods for the detection and analysis of protein-protein interactions. *Proteomics* **7**, 2833-42 (2007).

142. Gurevich, V.V. & Gurevich, E.V. GPCR Signaling Regulation: The Role of GRKs and Arrestins. *Front. Pharmacol.* **10**, 125 (2019).

143. Liu-Yesucevitz, L. et al. ALS-linked mutations enlarge TDP-43-enriched neuronal RNA granules in the dendritic arbor. *J. Neurosci.* **34**, 4167-74 (2014).

144. Vanden Broeck, L., Callaerts, P. & Dermaut, B. TDP-43-mediated neurodegeneration: towards a loss-of-function hypothesis? *Trends Mol. Med.* **20**, 66-71 (2014).

145. Puig, O. et al. The tandem affinity purification (TAP) method: A general procedure of protein complex purification. *Methods* **24**, 218-229 (2001).

146. Moresco, J.J., Carvalho, P.C. & Yates, J.R., 3rd. Identifying components of protein complexes in C. elegans using co-immunoprecipitation and mass spectrometry. *J. Proteomics* **73**, 2198-204 (2010).

147. Paek, J. et al. Multidimensional tracking of GPCR signaling via peroxidase-catalyzed proximity labeling. *Cell* **169**, 338-349 e11 (2017).

148. Lee, S.Y. et al. APEX fingerprinting reveals the subcellular localization of proteins of interest. *Cell Rep.* **15**, 1837-47 (2016).

149. Dumrongprechachan, V. et al. Cell-type and subcellular compartment-specific APEX2 proximity labeling reveals activity-dependent nuclear proteome dynamics in the striatum. *Nat. Commun.* **12**, 4855 (2021).

150. Marmor-Kollet, H. et al. Spatiotemporal proteomic analysis of stress granule disassembly using apex reveals regulation by sumoylation and links to als pathogenesis. *Mol. Cell* **80**, 876-891 e6 (2020).

151. Ke, M. et al. Spatiotemporal profiling of cytosolic signaling complexes in living cells by selective proximity proteomics. *Nat. Commun.* **12**, 71 (2021).

152. Kim, D.I. et al. Biosite: a method for direct detection and quantitation of site-specific biotinylation. *J. Proteome Res.* **17**, 759-769 (2018).

153. Motani, K. & Kosako, H. BioID screening of biotinylation sites using the avidin-like protein Tamavidin 2-REV identifies global interactors of stimulator of interferon genes (STING). *J. Biol. Chem.* **295**, 11174-11183 (2020).

154. Choi-Rhee, E., Schulman, H. & Cronan, J.E. Promiscuous protein biotinylation by Escherichia coli biotin protein ligase. *Protein Sci.* **13**, 3043-50 (2004).

155. Roux, K.J., Kim, D.I., Raida, M. & Burke, B. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J. Cell Biol.* **196**, 801-10 (2012).

156. Branon, T.C. et al. Efficient proximity labeling in living cells and organisms with TurboID. *Nat. Biotechnol.* **36**, 880-887 (2018).

157. Chua, X.Y., Aballo, T., Elnemer, W., Tran, M. & Salomon, A. Quantitative interactomics of lck-turboid in living human t cells unveils t cell receptor stimulation-induced proximal lck interactors. *J. Proteome Res.* **20**, 715-726 (2021).

158. Uezu, A. et al. Identification of an elaborate complex mediating postsynaptic inhibition. *Science* **353**, 1123-9 (2016).

159. Murata, K. et al. Efficient induction of proximity-dependent labelling by biotin feeding in BMAL1-BioID knock-in mice. *J. Biochem.* **170**, 453-461 (2021).

160. Trinkle-Mulcahy, L. Recent advances in proximity-based labeling methods for interactome mapping. *F1000Res.* **8**(2019).

161. Bar, D.Z. et al. Biotinylation by antibody recognition-a method for proximity labeling. *Nat. Methods* **15**, 127-133 (2018).

162. De Munter, S. et al. Split-BioID: a proximity biotinylation assay for dimerization-dependent protein interactions. *FEBS Lett.* **591**, 415-424 (2017).

163. Cho, K.F. et al. Split-TurboID enables contact-dependent proximity labeling in cells. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 12143-12154 (2020).

164. Han, Y. et al. Directed evolution of split APEX2 peroxidase. *ACS Chem. Biol.* **14**, 619-635 (2019).

165. Han, S. et al. RNA-protein interaction mapping via MS2- or Cas13-based APEX targeting. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 22068-22079 (2020).

166. Tutucci, E. et al. An improved MS2 system for accurate reporting of the mRNA life cycle. *Nat. Methods* **15**, 81-89 (2018).

167. Qin, W., Cho, K.F., Cavanagh, P.E. & Ting, A.Y. Deciphering molecular interactions by proximity labeling. *Nat. Methods* **18**, 133-143 (2021).

168. Choi, C.R. & Rhee, H.W. Proximity labeling: an enzymatic tool for spatial biology. *Trends. Biotechnol.* **40**, 145-148 (2022).

169. Leitner, A. et al. Expanding the chemical cross-linking toolbox by the use of multiple proteases and enrichment by size exclusion chromatography. *Mol. Cell. Proteomics* **11**, M111 014126 (2012).

170. Schmidt, R. & Sinz, A. Improved single-step enrichment methods of cross-linked products for protein structure analysis and protein interaction mapping. *Anal. Bioanal. Chem.* **409**, 2393-2400 (2017).

171. Petrotchenko, E.V., Serpa, J.J. & Borchers, C.H. An isotopically coded CID-cleavable biotinylated cross-linker for structural proteomics. *Mol. Cell. Proteomics* **10**, M110 001420 (2011).

172. Kaake, R.M. et al. A new in vivo cross-linking mass spectrometry platform to define protein-protein interactions in living cells. *Mol. Cell. Proteomics* **13**, 3533-43 (2014).

173. Chavez, J.D. et al. Chemical crosslinking mass spectrometry analysis of protein conformations and supercomplexes in heart tissue. *Cell Syst.* **6**, 136-141 e5 (2018).

174. Rinner, O. et al. Identification of cross-linked peptides from large sequence databases. *Nat. Methods* **5**, 315-8 (2008).

175. Petrotchenko, E.V. & Borchers, C.H. ICC-CLASS: isotopically-coded cleavable crosslinking analysis software suite. *BMC Bioinformatics* **11**, 64 (2010).

176. Iacobucci, C. et al. A cross-linking/mass spectrometry workflow based on MS-cleavable cross-linkers and the MeroX software for studying protein structures and protein-protein interactions. *Nat. Protoc.* **13**, 2864-2889 (2018).

177. Piersimoni, L., Kastritis, P.L., Arlt, C. & Sinz, A. Cross-linking mass spectrometry for investigating protein conformations and protein-protein interactions horizontal line a method for all seasons. *Chem. Rev.* **122**, 7500-7531 (2022).

178. Parada, L.F., Tabin, C.J., Shih, C. & Weinberg, R.A. Human EJ bladder carcinoma oncogene is homologue of Harvey sarcoma virus ras gene. *Nature* **297**, 474-8 (1982).

179. Müller, M.P. & Goody, R.S. Molecular control of Rab activity by GEFs, GAPs and GDI. *Small GTPases* **9**, 5-21 (2017).

180. Traut, T.W. Physiological concentrations of purines and pyrimidines. *Mol. Cell. Biochem.* **140**, 1-22 (1994).

181. Cox, A.D., Der, C.J. & Philips, M.R. Targeting RAS membrane association: Back to the future for anti-RAS drug discovery? *Clin. Cancer Res.* **21**, 1819-1827 (2015).

182. Gray, J.L., Delft, F. & Brennan, P.E. Targeting the small GTPase superfamily through their regulatory proteins. *Angew. Chem. Int. Ed.* **59**, 6342-6366 (2020).

183. De Luca, A., Maiello, M.R., D'Alessio, A., Pergameno, M. & Normanno, N. The RAS/RAF/MEK/ERK and the PI3K/AKT signalling pathways: role in cancer pathogenesis and implications for therapeutic approaches. *Expert Opin. Ther. Targets* **16 Suppl 2**, S17-27 (2012).

184. Ahearn, I.M., Haigis, K., Bar-Sagi, D. & Philips, M.R. Regulating the regulator: post-translational modification of RAS. *Nat. Rev. Mol. Cell Biol.* **13**, 39-51 (2012).

185. Karachaliou, N. et al. KRAS mutations in lung cancer. *Clin. Lung Cancer* **14**, 205-14 (2013).

186. Kulhanek, K.R., Roose, J.P. & Rubio, I. Regulation of the small GTPase Ras and Its relevance to human disease. in *Ras Activity and Signaling: Methods and Protocols* (eds. Rubio, I. & Prior, I.) 19-43 (Springer US, New York, NY, 2021).

187. Pasqualato, S., Menetrey, J., Franco, M. & Cherfils, J. The structural GDP/GTP cycle of human Arf6. *EMBO Rep.* **2**, 234-8 (2001).

188. Deng, Y., Golinelli-Cohen, M.P., Smirnova, E. & Jackson, C.L. A COPI coat subunit interacts directly with an early-Golgi localized Arf exchange factor. *EMBO Rep.* **10**, 58-64 (2009).

189. Xu, R. et al. Arf6 regulates EGF-induced internalization of E-cadherin in breast cancer cells. *Cancer Cell Int.* **15**, 11 (2015).

190. Kon, S., Tanabe, K., Watanabe, T., Sabe, H. & Satake, M. Clathrin dependent endocytosis of E-cadherin is regulated by the Arf6GAP isoform SMAP1. *Exp. Cell Res.* **314**, 1415-1428 (2008).

191. Prigent, M. et al. ARF6 controls post-endocytic recycling through its downstream exocyst complex effector. *J. Cell Biol.* **163**, 1111-21 (2003).

192. Zaoui, K., Rajadurai, C.V., Duhamel, S. & Park, M. Arf6 regulates RhoB subcellular localization to control cancer cell invasion. *J. Cell Biol.* **218**, 3812-3826 (2019).

193. Adarska, P., Wong-Dilworth, L. & Bottanelli, F. ARF GTPases and their ubiquitous role in intracellular trafficking beyond the Golgi. *Front. Cell Dev. Biol.* **9**(2021).

194. Zeigerer, A. et al. Rab5 is necessary for the biogenesis of the endolysosomal system in vivo. *Nature* **485**, 465-70 (2012).

195. Evans, T.M., Ferguson, C., Wainwright, B.J., Parton, R.G. & Wicking, C. Rab23, a negative regulator of hedgehog signaling, localizes to the plasma membrane and the endocytic pathway. *Traffic* **4**, 869-84 (2003).

196. Olkkonen, V.M. et al. Isolation of a mouse cDNA encoding Rab23, a small novel GTPase expressed predominantly in the brain. *Gene* **138**, 207-11 (1994).

197. Eggenschwiler, J.T. & Anderson, K.V. Dorsal and lateral fates in the mouse neural tube require the cell-autonomous activity of the open brain gene. *Dev. Biol.* **227**, 648-60 (2000).

198. Yan, X. & Shen, Y. Rab-like small GTPases in the regulation of ciliary Bardet-Biedl syndrome (BBS) complex transport. *FEBS J.* **289**, 7359-7367 (2022).

199. Huet, D., Blisnick, T., Perrot, S. & Bastin, P. The GTPase IFT27 is involved in both anterograde and retrograde intraflagellar transport. *Elife* **3**, e02419 (2014).

200. Bonet-Ponce, L. & Cookson, M.R. The role of Rab GTPases in the pathobiology of Parkinson' disease. *Curr. Opin. Cell Biol.* **59**, 73-80 (2019).

201. Wallings, R.L. & Tansey, M.G. LRRK2 regulation of immune-pathways and inflammatory disease. *Biochem. Soc. Trans.* **47**, 1581-1595 (2019).

202. Tsukuba, T., Yamaguchi, Y. & Kadowaki, T. Large Rab GTPases: Novel membrane trafficking regulators with a calcium sensor and functional domains. *Int. J. Mol. Sci.* **22**, 7691 (2021).

203. Wheeler, D.B., Zoncu, R., Root, D.E., Sabatini, D.M. & Sawyers, C.L. Identification of an oncogenic RAB protein. *Science* **350**, 211-7 (2015).

204. Kurokawa, K. & Matsuda, M. Localized RhoA activation as a requirement for the induction of membrane ruffling. *Mol. Biol. Cell.* **16**, 4294-303 (2005).

205. Machacek, M. et al. Coordination of Rho GTPase activities during cell protrusion. *Nature* **461**, 99-103 (2009).

206. Bolado-Carrancio, A. et al. Periodic propagating waves coordinate RhoGTPase network dynamics at the leading and trailing edges during cell migration. *Elife* **9**, e58165 (2020).

207. Svensmark, J.H. & Brakebusch, C. Rho GTPases in cancer: friend or foe? *Oncogene* **38**, 7447-7456 (2019).

208. Orgaz, J.L., Herraiz, C. & Sanz-Moreno, V. Rho GTPases modulate malignant transformation of tumor cells. *Small GTPases* **5**, e29019 (2014).

209. Zamboni, V. et al. Rho GTPases in intellectual disability: From genetics to therapeutic opportunities. *Int. J. Mol. Sci.* **19**(2018).

210. Hedrick, N.G. & Yasuda, R. Regulation of Rho GTPase proteins during spine structural plasticity for the control of local dendritic plasticity. *Curr. Opin. Neurobiol.* **45**, 193-201 (2017).

211. Chan, D., Citro, A., Cordy, J.M., Shen, G.C. & Wolozin, B. Rac1 protein rescues neurite retraction caused by G2019S leucine-rich repeat kinase 2 (LRRK2). *J. Biol. Chem.* **286**, 16140-9 (2011).

212. Guiler, W., Koehler, A., Boykin, C. & Lu, Q. Pharmacological modulators of small GTPases of Rho family in neurodegenerative diseases. *Front Cell Neurosci* **15**, 661612 (2021).

213. Spiering, D. & Hodgson, L. Dynamics of the Rho-family small GTPases in actin regulation and motility. *Cell. Adh. Migr.* **5**, 170-80 (2011).

214. Li, H.Y., Cao, K. & Zheng, Y. Ran in the spindle checkpoint: a new function for a versatile GTPase. *Trends. Cell Biol.* **13**, 553-7 (2003).

215. Monecke, T. et al. Crystal structure of the nuclear export receptor CRM1 in complex with Snurportin1 and RanGTP. *Science* **324**, 1087-91 (2009).

216. Kuersten, S., Ohno, M. & Mattaj, I.W. Nucleocytoplasmic transport: Ran, beta and beyond. *Trends Cell Biol.* **11**, 497-503 (2001).

217. Bischoff, F.R. & Ponstingl, H. Catalysis of guanine nucleotide exchange on Ran by the mitotic regulator RCC1. *Nature* **354**, 80-2 (1991).

218. Bischoff, F.R., Krebber, H., Kempf, T., Hermes, I. & Ponstingl, H. Human RanGTPase-activating protein RanGAP1 is a homologue of yeast Rna1p involved in mRNA processing and transport. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 1749-53 (1995).

219. Macara, I.G. Transport into and out of the Nucleus. *Microbiol. Mol. Biol. Rev.* **65**, 570-594 (2001).

220. Dasso, M. The Ran GTPase: theme and variations. *Curr. Biol.* **12**, R502-8 (2002).

221. Boudhraa, Z., Carmona, E., Provencher, D. & Mes-Masson, A.M. Ran GTPase: A Key player in tumor progression and metastasis. *Front. Cell Dev. Biol.* **8**, 345 (2020).

222. Moore, L.D., Le, T. & Fan, G. DNA methylation and its basic function. *Neuropsychopharmacology* **38**, 23-38 (2013).

223. Felsenfeld, G. A brief history of epigenetics. *Cold Spring Harb. Perspect. Biol.* **6**, a018200 (2014).

224. Berdasco, M. & Esteller, M. DNA methylation in stem cell renewal and multipotency. *Stem Cell Res. Ther.* **2**, 42 (2011).

225. Liu, C. et al. Epigenetic regulation of miR-184 by MBD1 governs neural stem cell proliferation and differentiation. *Cell Stem Cell* **6**, 433-44 (2010).

226. Cohn, W.E. & Volkin, E. Nucleoside-5'-phosphates from ribonucleic acid. *Nature* **167**, 483-484 (1951).

227. Jia, G. et al. N^6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat. Chem. Biol.* **7**, 885-7 (2011).

228. Meyer, Kate D. et al. 5´ UTR m^6A Promotes cap-independent translation. *Cell* **163**, 999-1010 (2015).

229. He, C. Grand challenge commentary: RNA epigenetics? *Nat. Chem. Biol.* **6**, 863-5 (2010).

230. Holley, R.W. et al. Structure of a ribonucleic acid. *Science* **147**, 1462-1465 (1965).

231. Helm, M. & Alfonzo, J.D. Posttranscriptional RNA Modifications: playing metabolic games in a cell's chemical Legoland. *Chem. Biol.* **21**, 174-85 (2014).

232. Suzuki, T. The expanding world of tRNA modifications and their disease relevance. *Nat. Rev. Mol. Cell Biol.* **22**, 375-392 (2021).

233. Han, L. & Phizicky, E.M. A rationale for tRNA modification circuits in the anticodon loop. *RNA* **24**, 1277-1284 (2018).

234. Motorin, Y. & Helm, M. tRNA stabilization by modified nucleotides. *Biochemistry* **49**, 4934-44 (2010).

235. Sloan, K.E. et al. Tuning the ribosome: The influence of rRNA modification on eukaryotic ribosome biogenesis and function. *RNA Biol.* **14**, 1138-1152 (2017).

236. Ramanathan, A., Robb, G.B. & Chan, S.H. mRNA capping: biological functions and applications. *Nucleic Acids Res.* **44**, 7511-26 (2016).

237. Lewis, J.D. & Izaurralde, E. The role of the cap structure in RNA processing and nuclear export. *Eur. J. Biochem.* **247**, 461-9 (1997).

238. Andreev, D.E. et al. Differential contribution of the m7G-cap to the 5' end-dependent translation initiation of mammalian mRNAs. *Nucleic Acids Res.* **37**, 6135-47 (2009).

239. Furuichi, Y., LaFiandra, A. & Shatkin, A.J. 5'-Terminal structure and mRNA stability. *Nature* **266**, 235-9 (1977).

240. Meyer, K.D. & Jaffrey, S.R. The dynamic epitranscriptome: N^6-methyladenosine and gene expression control. *Nat. Rev. Mol. Cell Biol.* **15**, 313-26 (2014).

241. Liu, J. et al. A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. *Nat. Chem. Biol.* **10**, 93-5 (2014).

242. Wang, X. et al. Structural basis of N^6-adenosine methylation by the METTL3-METTL14 complex. *Nature* **534**, 575-8 (2016).

243. Wang, P., Doxtader, K.A. & Nam, Y. Structural basis for cooperative function of Mettl3 and Mettl14 methyltransferases. *Mol. Cell* **63**, 306-317 (2016).

244. Doxtader, K.A. et al. Structural basis for regulation of METTL16, an s-adenosylmethionine homeostasis factor. *Mol. Cell* **71**, 1001-1011 e4 (2018).

245. Pendleton, K.E. et al. The U6 snRNA m_6A methyltransferase METTL16 regulates SAM synthetase intron retention. *Cell* **169**, 824-835 e14 (2017).

246. Mendel, M. et al. Methylation of structured RNA by the m^{6}A writer METTL16 Is essential for mouse embryonic development. *Mol. Cell* **71**, 986-1000.e11 (2018).

247. Hess, M.E. et al. The fat mass and obesity associated gene (Fto) regulates activity of the dopaminergic midbrain circuitry. *Nat. Neurosci.* **16**, 1042-8 (2013).

248. Zheng, G. et al. ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol. Cell* **49**, 18-29 (2013).

249. Liu, N. et al. *N*^{6}-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature* **518**, 560-4 (2015).

250. Xiao, W. et al. Nuclear m^{6}A reader YTHDC1 regulates mRNA splicing. *Mol. Cell* **61**, 507-519 (2016).

251. Xu, W. et al. METTL3 regulates heterochromatin in mouse embryonic stem cells. *Nature* **591**, 317-321 (2021).

252. Li, Y. et al. *N*$^{6-}$Methyladenosine co-transcriptionally directs the demethylation of histone H3K9me2. *Nat. Genet.* **52**, 870-877 (2020).

253. Wang, X. et al. N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature* **505**, 117-20 (2014).

254. Du, H. et al. YTHDF2 destabilizes m^{6}A-containing RNA through direct recruitment of the CCR4-NOT deadenylase complex. *Nat. Commun.* **7**, 12626 (2016).

255. Choe, J. et al. mRNA circularization by METTL3-eIF3h enhances translation and promotes oncogenesis. *Nature* **561**, 556-560 (2018).

256. Wang, X. et al. *N*^{6}-methyladenosine modulates messenger RNA translation efficiency. *Cell* **161**, 1388-99 (2015).

257. Shi, H. et al. YTHDF3 facilitates translation and decay of *N*^{6}-methyladenosine-modified RNA. *Cell Res.* **27**, 315-328 (2017).

258. Lee, J.T. Epigenetic regulation by long noncoding RNAs. *Science* **338**, 1435-9 (2012).

259. Hansen, K.H. et al. A model for transmission of the H3K27me3 epigenetic mark. *Nat. Cell Biol.* **10**, 1291-300 (2008).

260. Verdone, L., Agricola, E., Caserta, M. & Di Mauro, E. Histone acetylation in gene regulation. *Brief Funct. Genomic Proteomic* **5**, 209-21 (2006).

261. Liu, J. et al. N^6-methyladenosine of chromosome-associated regulatory RNA regulates chromatin state and transcription. *Science* **367**, 580-586 (2020).

262. Wei, J. et al. FTO mediates LINE1 m^6A demethylation and chromatin regulation in mESCs and mouse development. *Science* **376**, 968-973 (2022).

263. Xu, W. et al. Dynamic control of chromatin-associated m^6A methylation regulates nascent RNA synthesis. *Mol. Cell* **82**, 1156-1168 e7 (2022).

264. Wang, Y. et al. N^6-methyladenosine RNA modification regulates embryonic neural stem cell self-renewal through histone modifications. *Nat. Neurosci.* **21**, 195-206 (2018).

265. Chen, J. et al. m^6A regulates neurogenesis and neuronal development by modulating histone methyltransferase Ezh2. *Genom. Proteom. Bioinform* **17**, 154-168 (2019).

266. Deng, S. et al. RNA m^6A regulates transcription via DNA demethylation and chromatin accessibility. *Nat. Genet.* **54**, 1427-1437 (2022).

267. Yang, X. et al. Interactive regulation of DNA demethylase gene TET1 and m^6A methyltransferase gene METTL3 in myoblast differentiation. *Int. J. Biol. Macromol.* **223**, 916-930 (2022).

268. Friedenstein, A.J., Piatetzky-Shapiro, I.I. & Petrakova, K.V. Osteogenesis in transplants of bone marrow cells. *Development* **16**, 381-390 (1966).

269. Pittenger, M.F. et al. Multilineage potential of adult human mesenchymal stem cells. *Science* **284**, 143-7 (1999).

270. Zuk, P.A. et al. Human adipose tissue is a source of multipotent stem cells. *Mol. Biol. Cell* **13**, 4279-95 (2002).

271. Jo, H. et al. Applications of mesenchymal stem cells in skin regeneration and rejuvenation. *Int. J. Mol. Sci.* **22**, 2410 (2021).

272. Charbord, P. Bone marrow mesenchymal stem cells: historical overview and concepts. *Hum. Gene Ther.* **21**, 1045-56 (2010).

273. Cook, D. & Genever, P. Regulation of mesenchymal stem cell differentiation. in *Transcriptional and translational regulation of stem cells* (eds. Hime, G. & Abud, H.) 213-229 (Springer Netherlands, Dordrecht, 2013).

274. Akune, T. et al. PPARgamma insufficiency enhances osteogenesis through osteoblast formation from bone marrow progenitors. *J. Clin. Invest.* **113**, 846-55 (2004).

275. Berry, R. & Rodeheffer, M.S. Characterization of the adipocyte cellular lineage *in vivo*. *Nat. Cell Biol.* **15**, 302-8 (2013).

276. Lee, Y.H., Petkova, A.P., Mottillo, E.P. & Granneman, J.G. *In vivo* identification of bipotential adipocyte progenitors recruited by beta3-adrenoceptor activation and high-fat feeding. *Cell Metab.* **15**, 480-91 (2012).

277. Siersbaek, R., Nielsen, R. & Mandrup, S. PPARgamma in adipocyte differentiation and metabolism--novel insights from genome-wide studies. *FEBS Lett.* **584**, 3242-9 (2010).

278. Yeh, W.C., Bierer, B.E. & McKnight, S.L. Rapamycin inhibits clonal expansion and adipogenic differentiation of 3T3-L1 cells. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 11086-90 (1995).

279. Summers, S.A., Whiteman, E.L. & Birnbaum, M.J. Insulin signaling in the adipocyte. *Int. J. Obes. Relat. Metab. Disord.* **24 Suppl 4**, S67-70 (2000).

280. Lefterova, M.I. et al. PPARgamma and C/EBP factors orchestrate adipocyte biology via adjacent binding on a genome-wide scale. *Genes Dev.* **22**, 2941-52 (2008).

281. Infante, A. & Rodriguez, C.I. Osteogenesis and aging: lessons from mesenchymal stem cells. *Stem Cell Res. Ther.* **9**, 244 (2018).

282. Tseng, P.C. et al. Resveratrol promotes osteogenesis of human mesenchymal stem cells by upregulating RUNX2 gene expression via the SIRT1/FOXO3A axis. *J. Bone Miner. Res.* **26**, 2552-63 (2011).

283. Ducy, P. & Karsenty, G. Two distinct osteoblast-specific cis-acting elements control expression of a mouse osteocalcin gene. *Mol. Cell Biol.* **15**, 1858-69 (1995).

284. Tarkkonen, K., Hieta, R., Kytola, V., Nykter, M. & Kiviranta, R. Comparative analysis of osteoblast gene expression profiles and Runx2 genomic occupancy of mouse and human osteoblasts in vitro. *Gene* **626**, 119-131 (2017).

285. Takarada, T. et al. An analysis of skeletal development in osteoblast-specific and chondrocyte-specific runt-related transcription factor-2 (Runx2) knockout mice. *J. Bone Miner. Res.* **28**, 2064-9 (2013).

286. Komori, T. Regulation of proliferation, differentiation and functions of osteoblasts by Runx2. *Int. J. Mol. Sci.* **20**, 1694 (2019).

287. Komori, T. Regulation of bone development and maintenance by Runx2. *Front. Biosci.* **13**, 898-903 (2008).

288. Liu, W. et al. Overexpression of Cbfa1 in osteoblasts inhibits osteoblast maturation and causes osteopenia with multiple fractures. *J. Cell Biol.* **155**, 157-66 (2001).

289. Zhu, F., Friedman, M.S., Luo, W., Woolf, P. & Hankenson, K.D. The transcription factor osterix (SP7) regulates BMP6-induced human osteoblast differentiation. *J. Cell Physiol.* **227**, 2677-85 (2012).

290. Khatiwala, C.B., Peyton, S.R., Metzke, M. & Putnam, A.J. The regulation of osteogenesis by ECM rigidity in MC3T3-E1 cells requires MAPK activation. *J. Cell Physiol.* **211**, 661-72 (2007).

291. Khatiwala, C.B., Kim, P.D., Peyton, S.R. & Putnam, A.J. ECM compliance regulates osteogenesis by influencing MAPK signaling downstream of RhoA and ROCK. *J. Bone Miner. Res.* **24**, 886-98 (2009).

292. Chen, Q. et al. Fate decision of mesenchymal stem cells: adipocytes or osteoblasts? *Cell Death Differ.* **23**, 1128-39 (2016).

293. Ghaben, A.L. & Scherer, P.E. Adipogenesis and metabolic health. *Nat. Rev. Mol. Cell Biol.* **20**, 242-258 (2019).

2 Chapter 2. Targeted Proteomic Analysis of Small GTPases in Murine Adipogenesis

2.1 Introduction

Adipocytes play a crucial role in energy homeostasis, not only passively serving as the primary energy depot but also actively participating in various physiological processes that regulate metabolism.[1] However, excessive adiposity, or obesity, is a major risk factor for metabolic syndromes, which is manifested by the observation that overweight patients are often comorbid with type 2 diabetes, hyperlipidemia, and cardiovascular diseases.[2-4] Epidemiological studies also identified obesity as a significant risk factor for various cancers.[5,6] A recent report from the World Health Organization shows that globally more than 1.9 billion adults are overweight or obese, and such high prevalence poses a great threat to worldwide public health.[7] Thus, unveiling novel adipogenic pathways may provide a molecular basis for targeting the primary risk factors for these prevalent non-communicable diseases.

Adipogenesis is a strictly regulated process under the control of multiple intrinsic and extrinsic factors.[8] The precursor stem cells first commit into a transitional stage called preadipocytes, which lose their multipotent potential. Later the preadipocytes enter the final stage of adipogenesis and start expressing adipocyte-specific genes to activate the machinery for lipid synthesis and storage.[8] Current research in the molecular regulation of adipogenesis focuses on transcriptional activation of adipocytic genes by peroxisome proliferator-activated receptor γ (PPAR-γ), CCAAT-enhancer-binding proteins (C/EBPs) and Krüppel-like factors (KLFs).[9] However, the upstream regulation and signal

transduction cascades for these nuclear transcription factors remain largely unexplored. Previous studies also documented that adipogenesis induction by 3-isobutyl-1-methyl xanthine (IBMX) leads to the accumulation of cyclic adenosine monophosphate (cAMP), thereby activating protein kinase A (PKA). Nevertheless, not much is known about the downstream proteins involved in this adipogenic signaling cascade.[10]

Interestingly, previous studies have shown that some small GTPases assume important roles in PKA-mediated adipogenic differentiation.[11] Small GTPases are monomeric guanine nucleotide-binding proteins that comprise several subfamilies, including Ras, Rab, Rho, Arf, and Ran.[12] In response to environmental stimulation such as hormones, these proteins function as pivotal hubs of various cellular processes including cell division, differentiation, and signal transduction.[13-16] Several small GTPases have been previously characterized in adipocytes. For instance, Arl15 mediates metabolic regulation in adipose tissues[17] and the Rho subfamily was found to be involved in adipogenesis.[18,19] Nevertheless, there has been no comprehensive analysis about the entire family of small GTPases in regulating adipogenesis and adipocyte physiology.

We recently developed a targeted proteomic method for interrogating the small GTPase proteome,[20,21] which exploits their narrow molecular weight distribution (15-37 kDa) for their enrichment by SDS-PAGE followed by in-gel digestion of the enriched proteins. In this study, we employed this method, in conjunction with scheduled multiple-reaction monitoring (MRM) and synthetic stable isotope-labeled peptides, to assess the altered protein expression of small GTPases accompanied with adipogenic differentiation.

We observed overall diminished expression of a large number of small GTPases and uncovered a previously unknown role of Rab32 in promoting adipogenesis.

2.2 Materials and Methods

2.2.1 Cell Culture and Adipocyte Differentiation

3T3-L1 murine preadipocyte (ATCC CL-173) and C3H10T1/2 murine mesenchymal stem cells (MSC) (ATCC CCL-226) were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% newborn calf serum (Gibco) and fetal bovine serum (Fisher), respectively, and penicillin/streptomycin (100 IU/mL, GE Healthcare). Both lines of cells were sub-cultured at 80% confluence and maintained in a humidified atmosphere containing 5% CO_2. All experiments were conducted within 25 and 15 passages for 3T3-L1 and C3H10T1/2 cells, respectively.

Adipogenesis was induced as described previously.[22,23] Briefly, differentiation was initiated at 2 days post-confluence in complete DMEM containing 1 µM dexamethasone (Sigma), 0.5 mM 3-isobutyl-1-methylxanthine (Sigma), and 175 nM (for 3T3-L1 cells) or 1750 nM (for C3H10T1/2 cells) insulin (Sigma). After 48 h, the cells were refreshed with maintenance medium containing complete DMEM and 10 nM insulin. Maintenance medium was changed in every 48 h and the cells were harvested for protein extraction at 96 h post-differentiation. For overexpression, approximately 1.5×10^6 cells in a 6-well plate were first transiently transfected with 1.5 µg of plasmid encoding human RAB32 (Addgene plasmid #49611) using transIT-2020 (Mirus mir5400), and adipogenesis was induced 48 h later.

2.2.2 Cell Lysis and Protein Digestion

Cells were trypsinized and washed once with phosphate-buffered saline (PBS). Total proteins were extracted by incubating with CelLytic M cell lysis reagent (Sigma) containing 1% protease inhibitor cocktail (Sigma) on-ice for 30 min, followed by centrifugation at 16,000 *g* for 30 min at 4°C. Lysates were transferred into a pre-chilled tube by using a 1-mL syringe with a 26G needle without disturbing the floating lipid layer. Protein concentrations were determined by Quick Start Bradford Protein Assay (Bio-Rad).

For LC-MRM analysis, the cell lysates were concentrated with a 10 kDa MWCO centrifugal filter (VWR), and 100 µg of protein was mixed with 4× Laemmli SDS loading buffer to a final volume of 40 µl. Total proteins were resolved using a 10% SDS-PAGE gel and stained with Coomassie Brilliant Blue R-250. Gel bands in the molecular weight range of 15-37 kDa were excised for tryptic digestion. Briefly, the gel was cut into 1 mm^3 cubes and destained sequentially with 25% and 50% acetonitrile (ACN) in 50 mM ammonium bicarbonate (pH 7.8). The proteins were subjected to reduction with 10 mM dithiothreitol (DTT) at 37°C for 1 h and alkylation with 55 mM iodoacetamide at room temperature in dark for 20 min. The proteins were then digested in-gel with trypsin at 37°C for 18 h with an enzyme/protein ratio of 1:100. Peptides were extracted from the gel pieces with 50% ACN/5% acetic acid (HOAc), concentrated by Speed-Vac and desalted by using C18 ZipTip (Agilent). A crude pool of synthetic small GTPase peptides (New England Peptide, Inc.) with a C-terminal [$^{15}N_2$, $^{13}C_6$]-labeled lysine (+8.0 Da) or [$^{15}N_4$, $^{13}C_6$]-labeled arginine (+10 Da) (~4.0 fmol each) were spiked into approximately 4% of the digestion mixture, and the resulting mixture was subjected to LC-MRM analysis.[24]

2.2.3 Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS)

LC-MS/MS experiments were conducted on a TSQ Altis triple-quadrupole mass spectrometer operated in the MRM mode or an Q Exactive Plus hybrid quadrupole-Orbitrap mass spectrometer operated in the DDA mode. Peptides were separated with an UltiMate 3000 UPLC (Thermo Fisher Scientific, San Jose, CA) prior to electrospray ionization. Peptide mixtures were loaded onto a 4-cm capillary column (150 μm i.d.) packed with C18 resin (5 μm in particle size and 120 Å in pore size, Dr. Maisch GmbH HPLC) at a flow rate of 3 μL/min. The peptides were separated on a ~25 cm analytical column (75 μm i.d.) packed with C18 resin (3 μm in particle size and 120 Å in pore size, Dr. Maisch GmbH HPLC) with a 90-min gradient of 2-35% ACN in 0.1% formic acid at a flow rate of 300 nL/min. Ions were isolated with an FWHM of 0.7 in both Q1 and Q3 with a cycle time of 3 s. The collision energies for fragmenting the precursor ions were obtained from the default setting in Skyline.[25]

An MRM library for 138 unique peptides derived from murine small GTPases was generated with Skyline using LC-MS/MS data acquired from shotgun proteomic analyses.[21,24] The precursor → product ion transitions were monitored in a 3-min retention time window converted from its normalized retention time (iRT) and the iRT-RT standard curve of 10 standard peptides from the tryptic digestion mixture of BSA.[26] All targeted peptides were manually inspected and filtrated at dot-product (dotp) > 0.7 to ensure that the peak was correctly picked for the transitions. The relative levels of the targeted peptides were calculated from the ratios of sum of peak areas for the transitions of sample over the corresponding heavy stable isotope-labeled peptides. Each sample was analyzed twice, and

the values from the two technical replicates were averaged to represent one biological replicate. A total of two and three biological replicates were conducted for 3T3-L1 and C3H10T1/2 cells, respectively.

2.2.4 Oil Red O Staining

Oil Red O staining was performed as described previously.[27] Briefly, cells were washed once with PBS and fixed in 3.7% formaldehyde in PBS at room temperature for 1 h. The fixed cells were stained with 3 mg/mL Oil Red O (Sigma) in 60% isopropanol at room temperature for 30 min. Excess dye was subsequently washed off with water. Bright field images were recorded with a BZ-X710 Automated Microscope (Keyence).

2.2.5 RNA Extraction and RT-PCR

Total RNA was extracted with Total RNA Kit I (Omega) and purified with HiBind RNA mini columns (VWR). cDNA was generated from 1 μg of total RNA. Approximately 66 ng of cDNA was mixed with Luna qPCR master mixture (New England Biolabs) in a 96-well optical reaction plate and the reaction was monitored with Bio-rad CFX Connect. The primer sequences used for qPCR were: Rab32 forward, TACGTGCACCAGCTCTTCTCC; Rab32 reverse, TCGAGTCATGTTGCCAAACCG; PPARγ forward, CGCTGATGCACTGCCTATGA; PPARγ reverse, AGAGGTCCACAGAGCTGATTCC; GAPDH forward, CATCACTGCCACCCGAAGACTG; GAPDH reverse, ATGCCAGTGAGCTTCCCGTTCAG.

2.3 Results

2.3.1 Scheduled MRM Profiling of Differentially Expressed Small GTPases During Adipogenic Differentiation

Augmented adipogenesis contributes to metabolic diseases such as obesity and diabetes, yet the proteins factors involved in adipocyte differentiation from adipose-derived stem cells remain largely unexplored. In the present study, we aim to systemically characterize the alterations in expression of small GTPase proteins from adipose-derived stem cells that have committed to the adipocyte lineage. To improve the coverage of the small GTPase proteome and the throughput of the method, we employed scheduled MRM analysis coupled with gel fractionation (Figure 2.1).[20] The high-throughput of method is reflected by the fact that we were able to monitor 138 small GTPases, which represent approximately 80% of the murine small GTPase proteome,[28] in a single 90-min LC-MS/MS run. The method is capitalized on the fact that the molecular weights of most, if not all, small GTPases are in the range of 15-37 kDa; thus, we enriched the small GTPase proteome using SDS-PAGE and excised the gel bands in the corresponding molecular weight range.

To achieve sensitive and reliable detection of small GTPase peptides, we selected one unique peptide for each small GTPase in our library and chose the three most abundant y-ions for each peptide based on MS/MS acquired from shotgun proteomic analyses. Together, we incorporated peptides derived from 138 murine small GTPases into our Skyline MRM library (Figure S2.1). Prior to injection, the samples were mixed with the synthetic stable isotope-labeled heavy peptides of identical amino acid sequences to our

target analytes. It is noteworthy that the heavy peptides and the unlabeled sample peptides had nearly identical retention times and thus they were ionized at the same time. In LC-MRM analyses, the mass spectrometer was scheduled to monitor a subset of transitions in a predefined 3-min retention time window calculated from an iRT algorithm.[26]

3T3-L1 murine embryonic fibroblastic cells and C3H10T1/2 murine mesenchymal stem cells (MSC) are pluripotent stem cells that can be differentiated into a subset of cell types including adipocyte.[29-31] Numerous pivotal metabolic networks, e.g. the molecular regulation of glucose uptake upon insulin stimulation, have been discovered in adipocytes derived from these two cell lines.[23,32,33] Thus, these cell lines are ideal systems to reveal the molecular determinants of adipogenic differentiation. Herein, we induced the differentiation of 3T3-L1 and C3H10T1/2 cells into adipocytes, and the differentiation efficiency was quantified by Oil Red O staining (Figure 2.2A). Notably, while the majority of 3T3-L1 fibroblasts were almost fully differentiated into adipocytes, the differentiation efficiency for C3H10T1/2 cells was only approximately 20%, with the rest being undifferentiated precursor cells, which is consistent with the previous findings.[23]

A total of 100 µg of protein lysate was separated by SDS-PAGE, and the 15-37 kDa protein fraction was digested in-gel with trypsin. Approximately 4% of the ensuing digestion mixture was injected for LC-MS/MS analysis. Together, we quantified 55 and 49 small GTPases in 3T3-L1 and C3H10T1/2 cells, respectively, and among them 41 were commonly detected in both cell lines (Figures 2.2B and C, S2.1 and S2.2). To evaluate the sensitivity of the scheduled MRM method, we also analyzed the same samples in data-dependent acquisition (DDA) mode. In DDA mode, precursor ions for acquiring MS/MS

are selected on the basis of their relative abundances, which leads to inherently poor reproducibility in the identification and quantification of low-abundance co-eluting peptides. In contrast, MRM selectively monitors the peptides of interest and thus increases the detection sensitivity and reproducibility. Indeed, our results showed that the scheduled MRM analysis provides ~8 and 11-fold increases in total numbers of quantified small GTPases in 3T3-L1 and C3H10T1/2 cells, respectively, as compared with DDA analysis (Figure S2.3).

Interestingly, we found an overall diminished expression of small GTPases upon differentiation to adipocytes (Figure 2.3A and B). Rho small GTPases are by far the most well characterized family of small GTPases in adipogenesis. Excessive activities of Rho GTPases (e.g. RhoA) redirects the differentiation of precursor stem cells and inhibits adipogenesis,[18] and RhoA also promotes osteogenesis while suppressing adipogenesis.[19] Our results revealed decreased expression of RhoC in both cell lines and RhoA in 3T3-L1 upon differentiation to adipocytes. Hence, suppressing the activities or protein expressions of certain small GTPases might be crucial for adipogenic differentiation from mesenchymal stem cells.

2.3.2 Validation of Differentially Expressed Small GTPases With Western Blot Analysis

Our proteomic results revealed an overall down-regulation of small GTPases during adipogenesis. To determine the key regulators that orchestrate adipogenic differentiation, we applied a cutoff of 1.5-fold change and found that, among all the quantified small GTPases, 2 and 35 proteins were up- and down-regulated, respectively, in 3T3-L1 cells

(Figure S2.4A and B). In contrast, only 6 small GTPases were found decreased with the same threshold in C3H10T1/2 cells (Figure S2.4B), which can be rationalized from the relatively low differentiation rate of this cell line (Figure 2.2A). As the majority of the C3H10T1/2 cells remained undifferentiated, we can expect a higher background in discovering the differentially expressed small GTPases accompanied with adipogenesis. In this regard, by lowering the cutoff threshold to 1.3-fold for C3H10T1/2 cells, we were able to identify 17 small GTPases exhibited differential expression, and 13 out of these 17 proteins were also down-regulated in 3T3-L1 cells upon differentiation into adipocytes (Figure S2.5).

Strikingly, we observed larger alterations in quantification results for the expression levels of small GTPases from C3H10T1/2 than 3T3-L1 cells (Table S2.1 and S2.2). This can be attributed to the greater heterogeneity of C3H10T1/2 cells.[34] As a result, the C3H10T/2 cells are more multipotent with respect to cell differentiation. This variability is also highlighted by the advice that C3H10T1/2 cells are cultured within 15 passages,[35,36] suggesting that the cell physiology might drift progressively in later passages.

Although the two precursor cell lines employed for the adipogenic differentiation differ from each other in various physiological aspects, we expect that general regulatory small GTPases will exhibit similar trends in expression levels upon differentiation. Indeed, a comparison of the data from the two cell lines showed that Rab32 was consistently down-regulated in adipocytes relative to their precursor stem cells. To further confirm the findings made from the scheduled MRM analysis, we employed Western blot analysis for several small GTPases using commercially available antibodies (Figures 2.4A and B). The

Western blot results confirmed the quantification data obtained from MRM analysis, thereby supporting the accuracy and robustness of scheduled MRM in quantifying the alterations in expression levels of small GTPases during adipogenic differentiation.

2.3.3 Rab32 is a Novel Suppressor of Adipogenesis

Previous studies revealed Rab32 as an important regulator for metabolism, where cells expressing Rab32 dominant-negative mutants failed to accumulate lipid droplets.[37] Results from our scheduled MRM analysis prompted us to posit that diminished expression of Rab32 may promote adipogenic differentiation. To test this hypothesis, we overexpressed Rab32 in 3T3-L1 and C3H10T1/2 cells. The degree of differentiation was again examined with Oil Red O staining and the results demonstrated that augmented expression of Rab32 significantly suppressed adipogenesis in both lines of cells (Figure 2.5A and B).

While the protein level of Rab32 decreased upon differentiation, the molecular mechanism through which it is down-regulated during differentiation remains unclear. To further investigate the regulation of Rab32 expression, we used real-time PCR to assess the changes in mRNA level of Rab32 during differentiation. Our results showed that, while the mRNA level of the master driver for adipocyte differentiation, i.e., PPARγ, was increased, manifesting a successful induction, the differentiation did not alter the mRNA expression level of Rab32 (Figure S2.6). Thus, the precursor stem cells may orchestrate attenuated Rab32 expression via a post-transcriptional mechanism.

2.4 Discussion

To our knowledge, this is the first scheduled LC-MRM, coupled with synthetic stable isotope-labeled peptides, for the quantitative analysis of differential expression of the murine small GTPase proteome. With the high-throughput method presented herein, we were able to accurately quantify 55 and 49 small GTPases in 3T3-L1 and C3H10T1/2 cells, respectively, upon differentiation into adipocytes. To evaluate the performance of the MRM method, we also performed DDA analysis for the same tryptic peptide mixtures, and the results demonstrated that scheduled MRM afforded superior reproducibility and sensitivity in the quantitative assessment of the small GTPase proteome.[38,39] Notably, Huang et al.,[20,21,40] by employing a similar approach with stable isotope labeling by amino acids in cell culture (SILAC) technique, successfully quantified approximately 90 small GTPases in human cells, which offered significantly increased coverage of small GTPase proteome when compared to analysis conducted in the DDA mode. In this vein, while DDA mode is commonly used in exploratory proteomic analysis, targeted proteomics such as MRM can provide more sensitive and reproducible quantification for a subset of proteins of interest such as disease biomarkers.[41,42] Together, we demonstrated a reliable and efficient platform in revealing differential expression of small GTPases in cultured murine cells. It is noteworthy that the heavy isotope-labeled peptides were spiked into the samples prior to injection for LC-MS/MS analyses, which shows the adaptability and flexibility of the method to any unlabeled specimens such as clinical samples.[24]

To date, the involvement of small GTPases in cell fate determination in adipogenesis remains largely unexplored. Here, we systemically examined the differential

expression of the small GTPase proteome during adipogenesis. Our results validated that the known suppressors for adipogenesis, i.e. the Rho family small GTPases (e.g. RhoA), were down-regulated upon adipogenesis.[18,43] Among the quantified small GTPases, we found that the expression of Rab32 was consistently lower in the differentiated adipocytes relative to the undifferentiated precursor cells. Functionally, Rab32 is a known regulatory factor in lipid accumulation. Dominant-negative Rab32 mutation prevents lipid droplet formation in fruit flies.[37,44] Although Rab32 is known to modulate lipid metabolism, its role in adipogenesis has not yet been characterized. Here, we found that Rab32 may negatively regulate adipogenesis, as overexpression of Rab32 could suppress the adipogenic differentiation of 3T3-L1 and C3H10T1/2 cells.

Rab32 was shown to interact with protein kinase A (PKA) and regulate intracellular transport of vesicles in a cAMP-dependent manner.[45,46] Interestingly, the same pathway can also be activated by a phosphodiesterase inhibitor IBMX,[47] a component of the adipogenesis induction cocktail used herein. Such activation of the cAMP/PKA pathway leads to the expression of adipogenesis driver, PPARγ, thereby orchestrating the differentiation of mesenchymal stem cells toward adipocytes. Thus, Rab32 may mediate adipogenesis via PKA, though the underlying mechanism awaits further study.

Our real-time PCR results showed that during adipogenesis, while Rab32 protein levels were down-regulated after induction of adipogenesis, its mRNA levels remained unchanged. Our findings suggested that attenuated level of Rab32 protein during adipocyte differentiation may arise from a yet-to-be-identified factor that suppresses Rab32 translation via a RNA modification-dependent mechanism.[48,49]

In conclusion, we have demonstrated the accuracy and sensitivity of the scheduled MRM method in studying differentially expressed small GTPases in murine cells. Our results showed that the method allowed for high-throughput detection of small GTPase proteins in murine cells. With the method, we discovered approximately 20 candidate small GTPase regulators of adipogenesis. In addition, we documented Rab32 as a suppressor for adipogenic differentiation. To our knowledge, this is the first comprehensive study of the reprogramming of the small GTPases during adipogenesis.

2.5 Acknowledgements

The authors would like to thank Prof. Alan Saltiel and Shaochen Chen at the University of California San Diego for providing the 3T3-L1 and C3H10T1/2 cells, respectively. This work was supported by the National Institutes of Health (R01 CA210072).

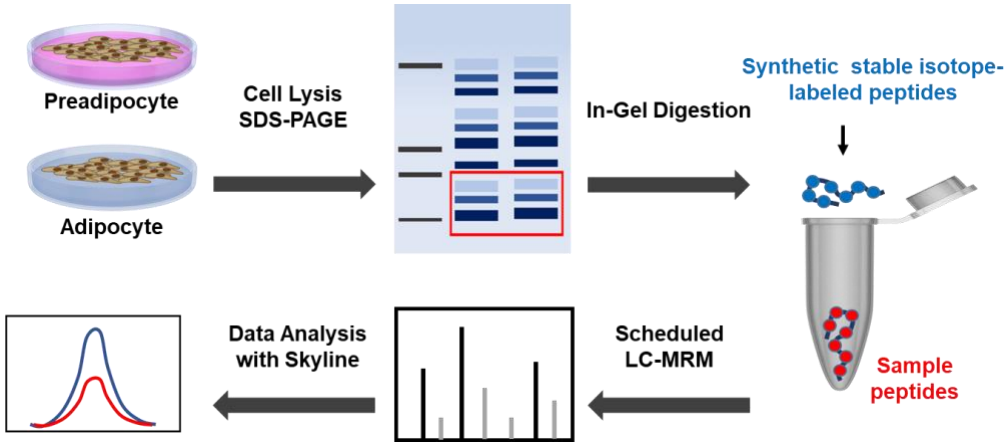


Figure 2.1 A schematic diagram illustrating the scheduled MRM-based targeted proteomic workflow for assessing the alterations in small GTPase proteome during adipogenic differentiation.

Cells were induced to differentiate to adipocytes and total lysates were collected for SDS-PAGE fractionation and in-gel tryptic digestion. Synthetic stable isotope-labeled peptides corresponding to 138 murine small GTPases were added to samples and the mixtures were subjected to scheduled MRM analysis. The results were analyzed with Skyline.

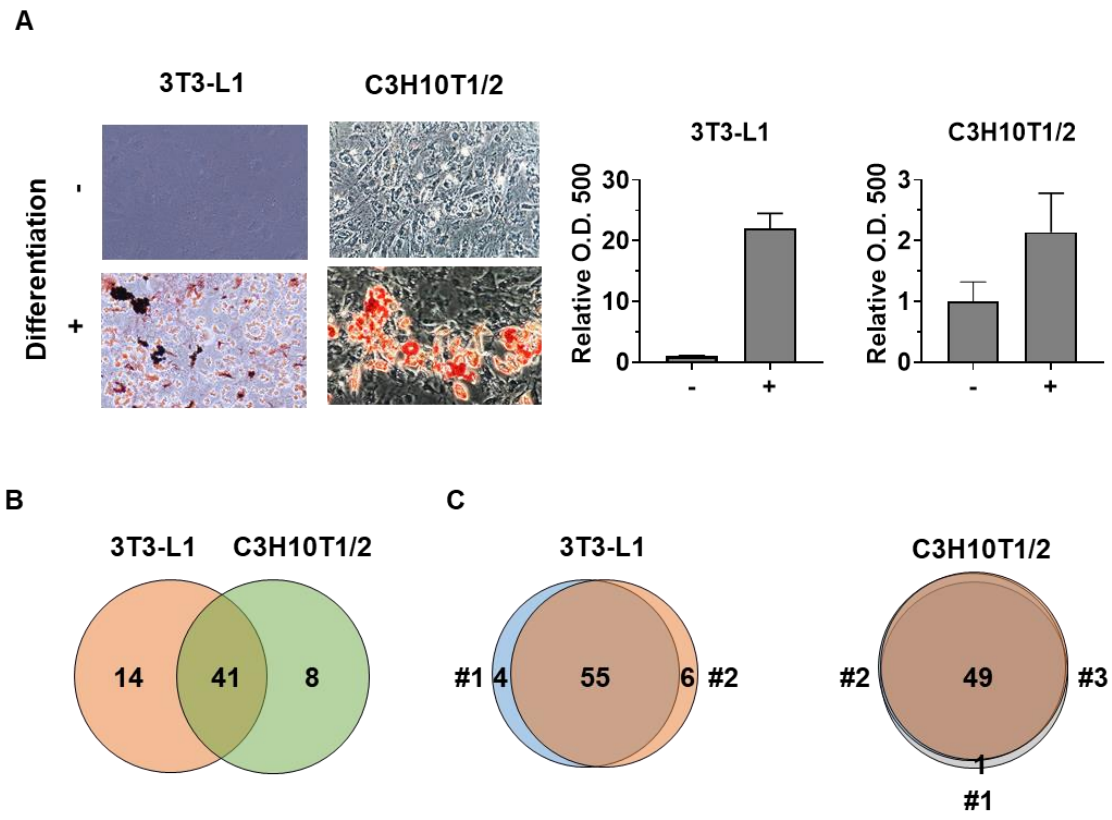


Figure 2.2 Numbers of small GTPases quantified in 3T3-L1 and C3H10T1/2 murine cells during adipocyte differentiation.

(A) 3T3-L1 murine fibroblast and C3H10T1/2 murine mesenchymal stem cells (-) were induced to differentiation into adipocytes (+). Adipogenesis efficiency was confirmed by Oil Red O staining. The lipid dye was later extracted and its UV absorbance at 500 nm was measured. The data represent the mean ± standard deviation (S.D.) of quantification results (n = 3). (B) A Venn diagram showing the numbers of small GTPases in 3T3-L1 and C3H10T1/2 cells quantified with scheduled MRM analysis. (C) Venn diagrams showing the overlap of quantified small GTPases from the two biological replicates in 3T3-L1 cells and three biological replicates in C3H10T1/2 cells.

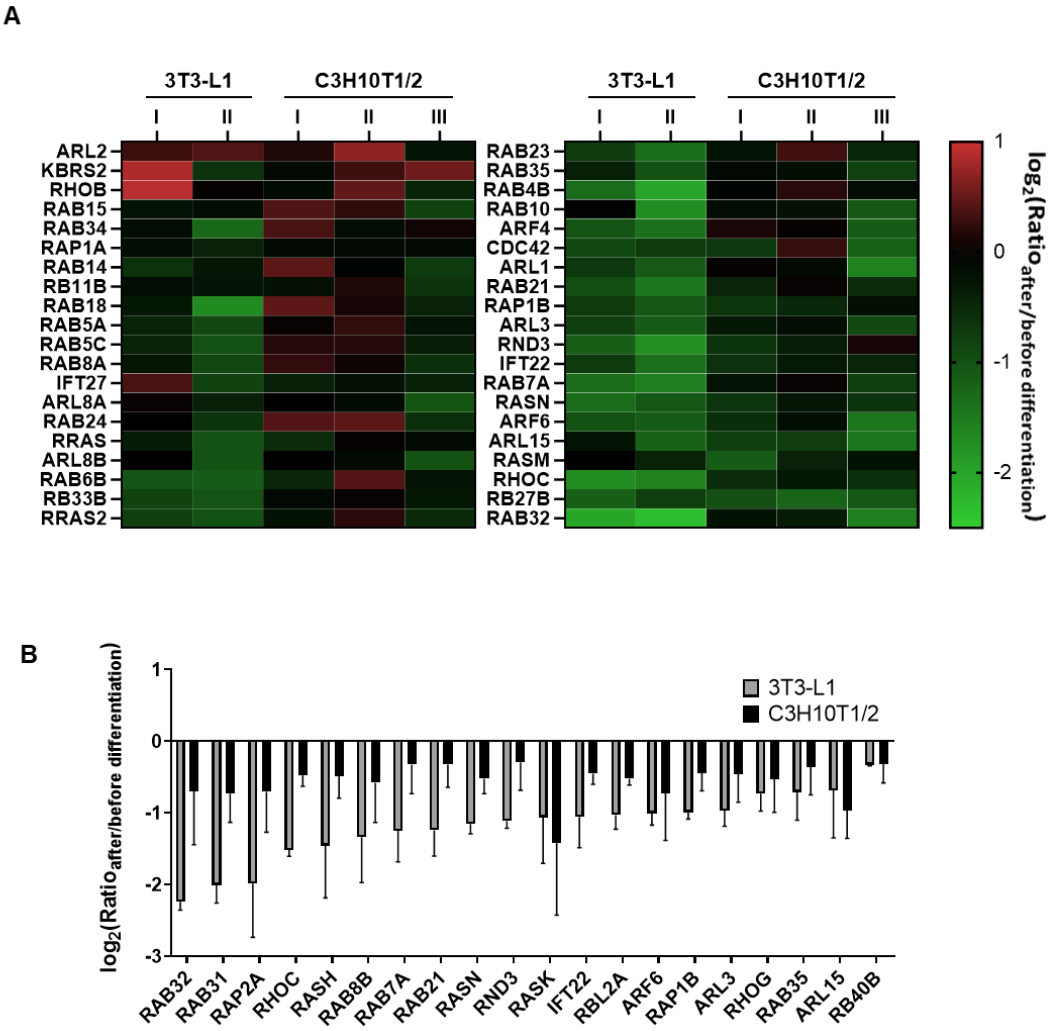


Figure 2.3 Differentially expressed small GTPases upon adipocyte differentiation.

(A) A heat map depicting the quantified small GTPases in 3T3-L1 and C3H10T1/2 cells upon differentiation. (B) Quantification data of top 20 differentially expressed small GTPases in 3T3-L1 and C3H10T1/2 cells. The data represent mean ± S.D of results obtained from 3 biological replicates.

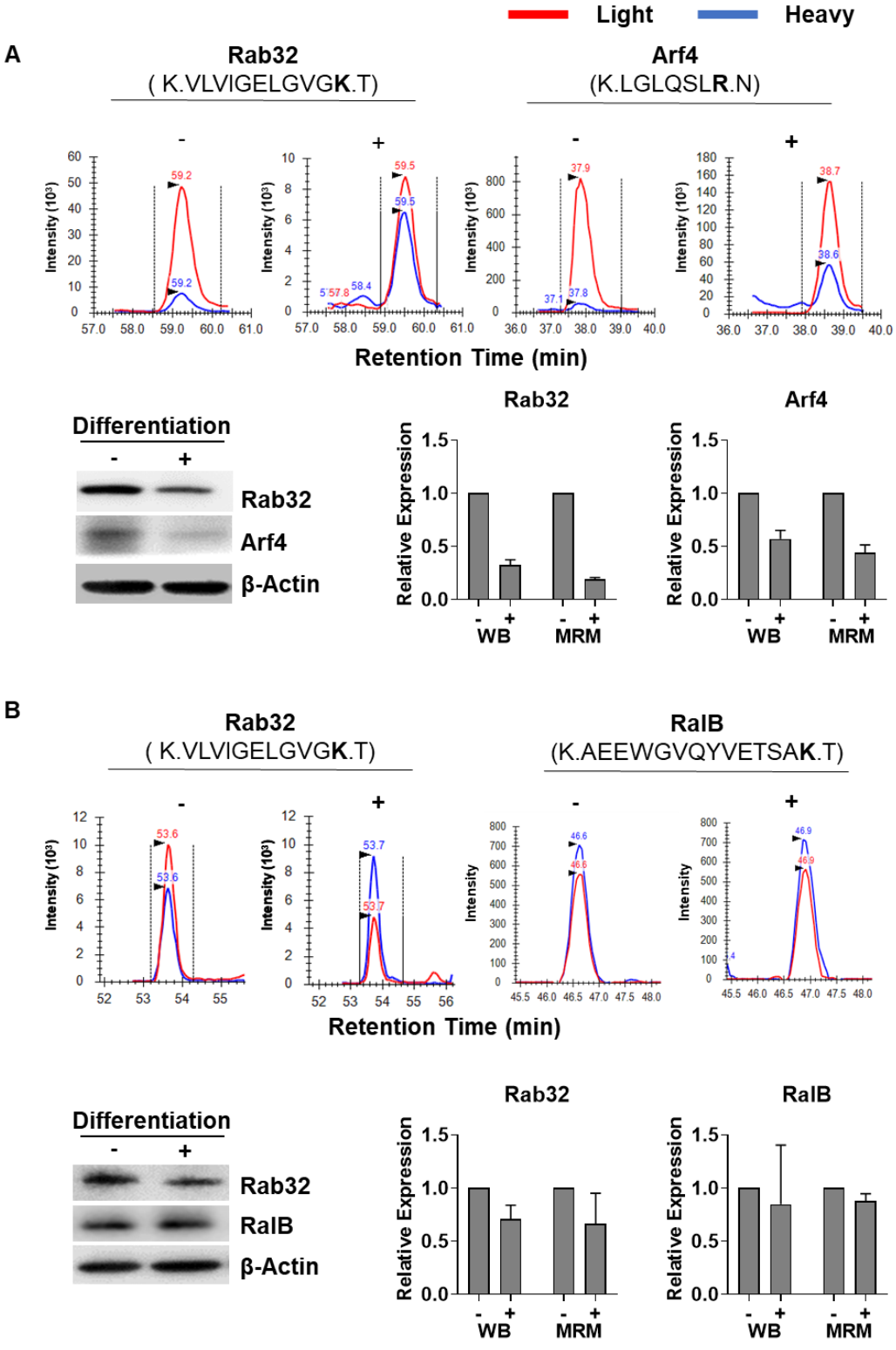

Light
Heavy
A
Rab32
(K.VLVIGELGVGK.T)
Arf4
(K.LGLQSLR.N)
Intensity (10³)
Retention Time (min)
Differentiation
Rab32
Arf4
β-Actin
Relative Expression
WB
MRM
B
Rab32
(K.VLVIGELGVGK.T)
RalB
(K.AEEWGVQYVETSAK.T)
Intensity
Retention Time (min)
Differentiation
Rab32
RalB
β-Actin
Relative Expression
WB
MRM

Figure 2.4 Western blot validation of the quantification results obtained from scheduled MRM analysis.

Shown are selected-ion chromatograms (upper panel) for monitoring tryptic peptides of representative small GTPases before (-) and after (+) differentiation of 3T3-L1 (A) and (B) C3H10T1/2 cells into adipocytes. The traces for the unlabeled peptides derived from cell lysates are shown in red and the corresponding traces for the spiked-in heavy isotope-labeled peptide are shown in blue. Western blot analysis (lower panel) was conducted to validate the MRM results of selected small GTPases. Band intensity of small GTPases was normalized against that of β-actin, and further normalized against the ratio of the undifferentiated cells.

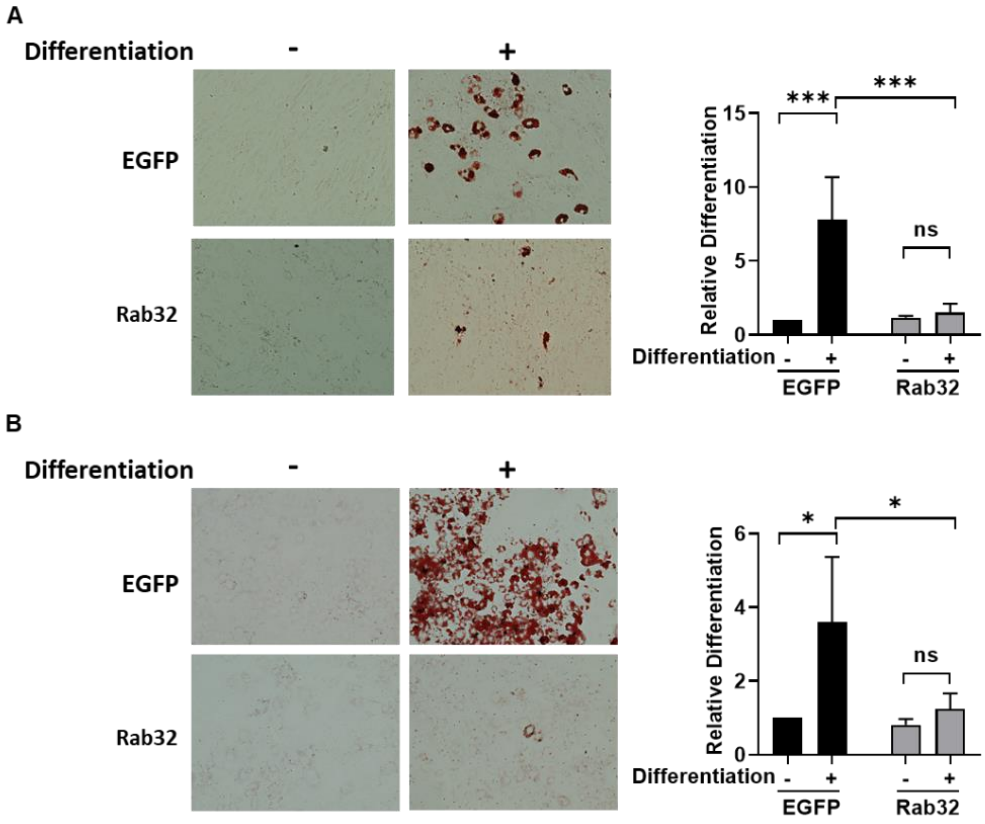


Figure 2.5 The role of Rab32 in adipocyte differentiation.

(A) 3T3-L1 and (B) C3H10T1/2 cells were transfected with control (EGFP) or EGFP-Rab32 plasmid followed by adipogenesis induction. The undifferentiated (-) and differentiated (+) cells were stained with Oil Red O and the adipogenesis efficiency was quantified by UV adsorption at 500 nm.

2.6 References

1. Rosen, E.D. & Spiegelman, B.M. Adipocytes as regulators of energy balance and glucose homeostasis. *Nature* **444**, 847-53 (2006).

2. Lavie, C.J., Milani, R.V. & Ventura, H.O. Obesity and cardiovascular disease: risk factor, paradox, and impact of weight loss. *J. Am. Coll. Cardiol.* **53**, 1925-32 (2009).

3. Jeffery, E., Church, C.D., Holtrup, B., Colman, L. & Rodeheffer, M.S. Rapid depot-specific activation of adipocyte precursor cells at the onset of obesity. *Nat. Cell. Biol.* **17**, 376-85 (2015).

4. Bhupathiraju, S.N. & Hu, F.B. Epidemiology of Obesity and Diabetes and Their Cardiovascular Complications. *Circ. Res.* **118**, 1723-35 (2016).

5. van Kruijsdijk, R.C., van der Wall, E. & Visseren, F.L. Obesity and cancer: the role of dysfunctional adipose tissue. *Cancer Epidem. Biomar.* **18**, 2569-78 (2009).

6. Yoshimoto, S. et al. Obesity-induced gut microbial metabolite promotes liver cancer through senescence secretome. *Nature* **499**, 97-101 (2013).

7. Swinburn, B.A. et al. The global obesity pandemic: shaped by global drivers and local environments. *Lancet* **378**, 804-14 (2011).

8. Rosen, E.D., Walkey, C.J., Puigserver, P. & Spiegelman, B.M. Transcriptional regulation of adipogenesis. *Genes Dev.* **14**, 1293-307 (2000).

9. Moseti, D., Regassa, A. & Kim, W.K. Molecular Regulation of Adipogenesis and Potential Anti-Adipogenic Bioactive Molecules. *Int. J. Mol. Sci.* **17**, 124 (2016).

10. Meng, W. et al. Rheb Inhibits Beiging of White Adipose Tissue via PDE4D5-Dependent Downregulation of the cAMP-PKA Signaling Pathway. *Diabetes* **66**, 1198-1213 (2017).

11. Otani, T. et al. Signaling pathway for adiponectin expression in adipocytes by osteocalcin. *Cell. Signal.* **27**, 532-44 (2015).

12. Mishra, A.K. & Lambright, D.G. Invited review: Small GTPases and their GAPs. *Biopolymers* **105**, 431-48 (2016).

13. van Dam, T.J., Bos, J.L. & Snel, B. Evolution of the Ras-like small GTPases and their regulators. *Small GTPases* **2**, 4-16 (2011).

14. Ridley, A.J. Rho GTPase signalling in cell migration. *Curr. Opin. Cell. Biol.* **36**, 103-12 (2015).

15. Villarroel-Campos, D., Bronfman, F.C. & Gonzalez-Billault, C. Rab GTPase signaling in neurite outgrowth and axon specification. *Cytoskeleton* **73**, 498-507 (2016).

16. Yang, J.Q. et al. RhoA orchestrates glycolysis for TH2 cell differentiation and allergic airway inflammation. *J. Allergy. Clin. Immunol.* **137**, 231-245 e4 (2016).

17. Rocha, N. et al. The metabolic syndrome- associated small G protein ARL15 plays a role in adipocyte differentiation and adiponectin secretion. *Sci .Rep.* **7**, 17593 (2017).

18. Sordella, R., Jiang, W., Chen, G.-C., Curto, M. & Settleman, J. Modulation of Rho GTPase Signaling Regulates a Switch between Adipogenesis and Myogenesis. *Cell* **113**, 147-158 (2003).

19. Strzelecka-Kiliszek, A. et al. Functions of Rho family of small GTPases and Rho-associated coiled-coil kinases in bone cells during differentiation and mineralization. *Biochim. Biophys. Acta.* **1861**, 1009-1023 (2017).

20. Huang, M. & Wang, Y. Targeted quantitative proteomic approach for probing altered protein expression of small GTPases associated with colorectal cancer metastasis. *Anal. Chem.* **91**, 6233-6241 (2019).

21. Huang, M., Qi, T.F., Li, L., Zhang, G. & Wang, Y. A targeted quantitative proteomic approach assesses the reprogramming of small gtpases during melanoma metastasis. *Cancer Res.* **78**, 5431-5445 (2018).

22. Merkestein, M. et al. FTO influences adipogenesis by regulating mitotic clonal expansion. *Nat. Commun.* **6**, 6792 (2015).

23. Tang, Q.Q., Otto, T.C. & Lane, M.D. Commitment of C3H10T1/2 pluripotent stem cells to the adipocyte lineage. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 9607-11 (2004).

24. Huang, M., Darvas, M., Keene, C.D. & Wang, Y. Targeted Quantitative Proteomic Approach for High-Throughput Quantitative Profiling of Small GTPases in Brain Tissues of Alzheimer's Disease Patients. *Anal. Chem.* **91**, 12307-12314 (2019).

25. Pino, L.K. et al. The Skyline ecosystem: Informatics for quantitative mass spectrometry proteomics. *Mass Spectrom Rev.* (2017).

26. Escher, C. et al. Using iRT, a normalized retention time for more targeted measurement of peptides. *Proteomics* **12**, 1111-21 (2012).

27. Dani, C. et al. Differentiation of embryonic stem cells into adipocytes in vitro. *J. Cell Sci.* **110**, 1279-1285 (1997).

28. Wennerberg, K., Rossman, K.L. & Der, C.J. The Ras superfamily at a glance. *J. Cell Sci.* **118**, 843-6 (2005).

29. Grigoriadis, A.E., Heersche, J.N. & Aubin, J.E. Differentiation of muscle, fat, cartilage, and bone from progenitor cells present in a bone-derived clonal cell population: effect of dexamethasone. *J. Cell Biol.* **106**, 2139 (1988).

30. Chen, Q. et al. Fate decision of mesenchymal stem cells: adipocytes or osteoblasts? *Cell Death Differ.* **23**, 1128-39 (2016).

31. Ogino, Y. et al. RhoA-Mediated Functions in C3H10T1/2 Osteoprogenitors Are Substrate Topography Dependent. *J. Cell. Physiol.* **231**, 568-75 (2016).

32. Skorobogatko, Y. et al. RalA controls glucose homeostasis by regulating glucose uptake in brown fat. *Proc. Natl. Acad. Sci. U S A* **115**, 7819-7824 (2018).

33. Vernochet, C. et al. Adipose-specific deletion of TFAM increases mitochondrial oxidation and protects mice against obesity and insulin resistance. *Cell Metab.* **16**, 765-76 (2012).

34. Ho, A.D., Wagner, W. & Franke, W. Heterogeneity of mesenchymal stromal cell preparations. *Cytotherapy* **10**, 320-30 (2008).

35. Reznikoff, C.A., Bertram, J.S., Brankow, D.W. & Heidelberger, C. Quantitative and qualitative studies of chemical transformation of cloned C3H mouse embryo cells sensitive to postconfluence inhibition of cell division. *Cancer Res.* **33**, 3239-49 (1973).

36. Khayat, G., Rosenzweig, D.H. & Quinn, T.M. Low frequency mechanical stimulation inhibits adipogenic differentiation of C3H10T1/2 mesenchymal stem cells. *Differentiation* **83**, 179-184 (2012).

37. Wang, C., Liu, Z. & Huang, X. Rab32 is important for autophagy and lipid storage in Drosophila. *PLOS ONE* **7**, e32086 (2012).

38. Ji, Y.-H. et al. Quantitative proteomics analysis of chondrogenic differentiation of C3H10T1/2 mesenchymal stem cells by iTRAQ labeling coupled with on-line two-dimensional LC/MS/MS. *Mol. Cell. Proteomics* **9**, 550 (2010).

39. Jiang, Y. et al. Proteome Profiling of Mitotic Clonal Expansion during 3T3-L1 Adipocyte Differentiation Using iTRAQ-2DLC-MS/MS. *J. Proteome Res.* **13**, 1307-1314 (2014).

40. Huang, M. & Wang, Y. Roles of Small GTPases in Acquired Tamoxifen Resistance in MCF-7 Cells Revealed by Targeted, Quantitative Proteomic Analysis. *Anal. Chem.* **90**, 14551-14560 (2018).

41. Ang, C.S. & Nice, E.C. Targeted in-gel MRM: a hypothesis driven approach for colorectal cancer biomarker discovery in human feces. *J. Proteome Res.* **9**, 4346-55 (2010).

42. Percy, A.J., Chambers, A.G., Yang, J., Hardie, D.B. & Borchers, C.H. Advances in multiplexed MRM-based protein biomarker quantitation toward clinical utility. *Biochim. Biophys. Acta.* **1844**, 917-26 (2014).

43. Meyers, V.E., Zayzafoon, M., Douglas, J.T. & McDonald, J.M. RhoA and cytoskeletal disruption mediate reduced osteoblastogenesis and enhanced adipogenesis of human mesenchymal stem cells in modeled microgravity. *J. Bone. Miner. Res.* **20**, 1858-66 (2005).

44. Li, Q., Wang, J., Wan, Y. & Chen, D. Depletion of Rab32 decreases intracellular lipid accumulation and induces lipolysis through enhancing ATGL expression in hepatocytes. *Biochem. Biophys. Res. Commun.* **471**, 492-6 (2016).

45. Alto, N.M., Soderling, J. & Scott, J.D. Rab32 is an A-kinase anchoring protein and participates in mitochondrial dynamics. *J. Cell Biol.* **158**, 659-668 (2002).

46. Park, M., Serpinskaya, A.S., Papalopulu, N. & Gelfand, V.I. Rab32 Regulates Melanosome Transport in Xenopus Melanophores by Protein Kinase A Recruitment. *Curr. Biol.* **17**, 2030-2034 (2007).

47. Yang, D.-C. et al. cAMP/PKA Regulates Osteogenesis, Adipogenesis and Ratio of RANKL/OPG mRNA Expression in Mesenchymal Stem Cells by Suppressing Leptin. *PLOS ONE* **3**, e1540 (2008).

48. Ries, R.J. et al. m6A enhances the phase separation potential of mRNA. *Nature* **571**, 424-428 (2019).

49. Meyer, K.D. et al. 5' UTR m^6A promotes cap-independent translation. *Cell* **163**, 999-1010 (2015).

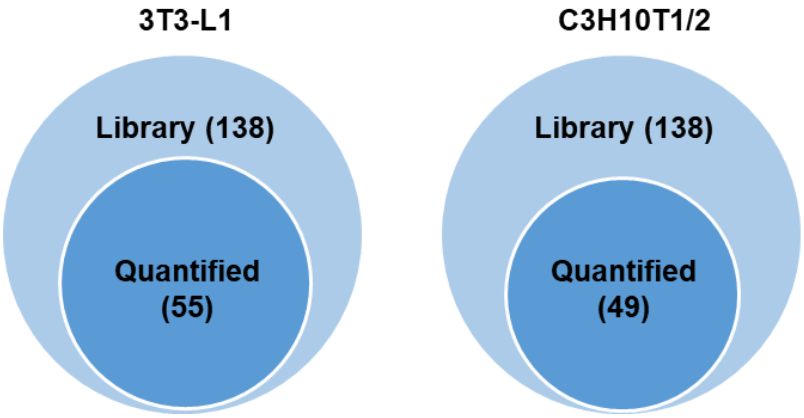


Figure S2.1 Scheduled MRM analysis of small GTPases upon adipogenesis.

(A) A Venn diagram showing the small GTPase proteins identified and quantified in both cell lines studied herein.

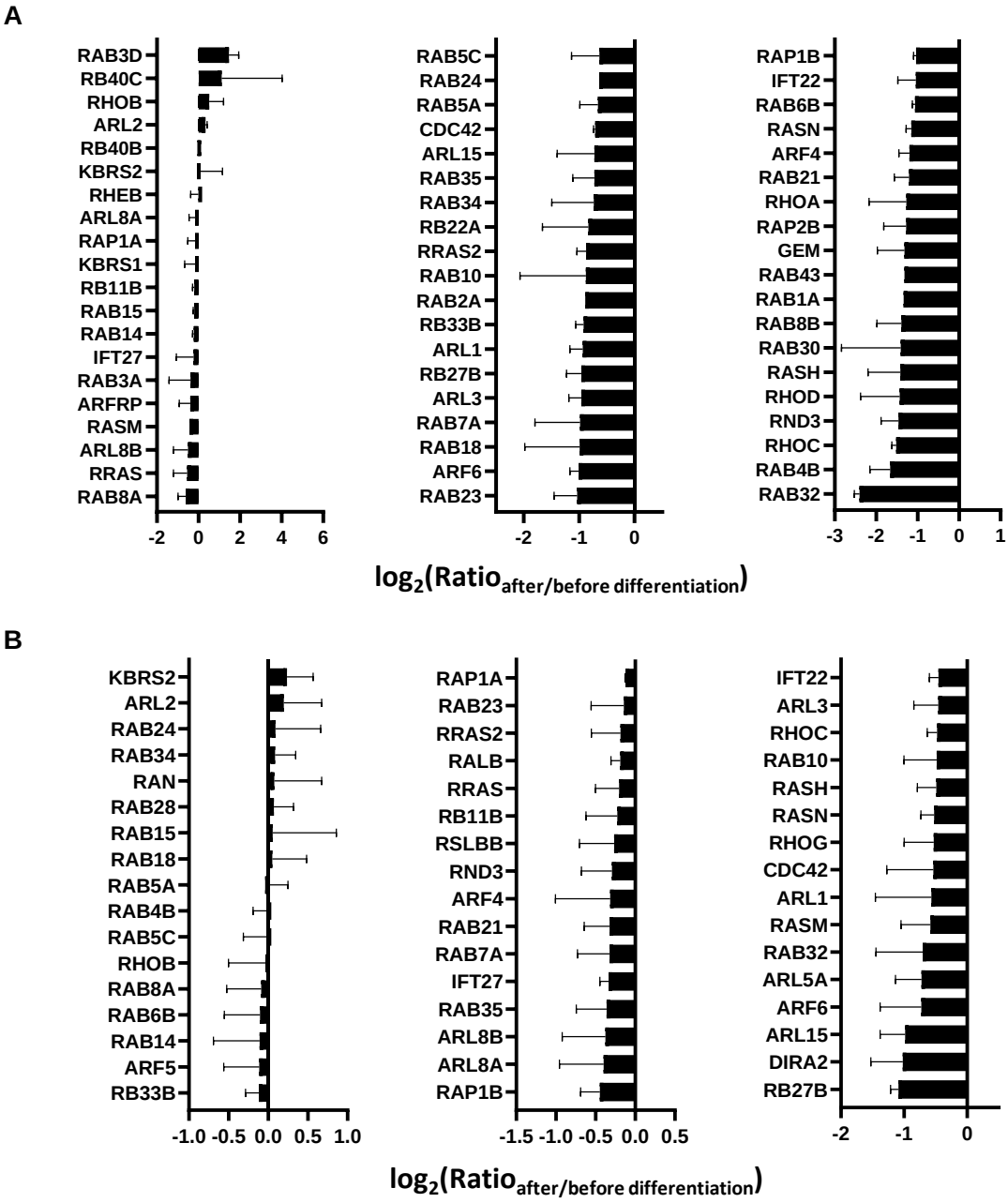


Figure S2.2 Quantification results for small GTPases in 3T3-L1.

(A) and C3H10T1/2 (B) cells during adipogenic differentiation. The data represent mean ± standard deviation (S.D.) of results from 2 and 3 independent experiments in 3T3-L1 and C3H10T1/2 cell lines, respectively.

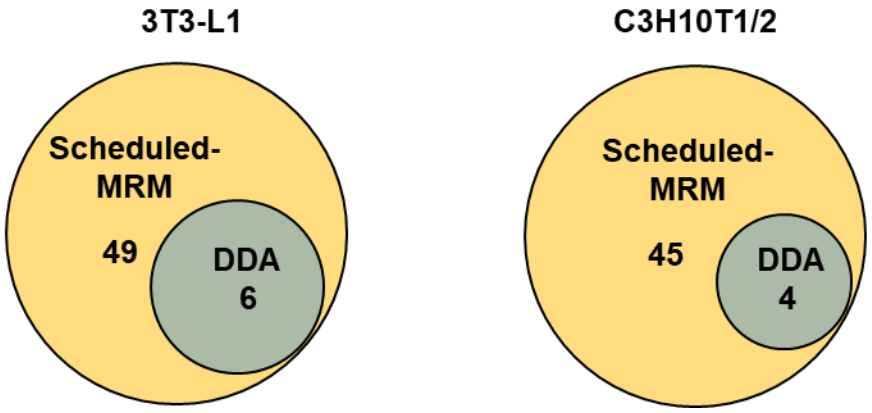


Figure S2.3 Venn diagrams showing the numbers of small GTPases detected in 3T3-L1 and C3H10T1/2 cells with scheduled MRM and DDA analyses.

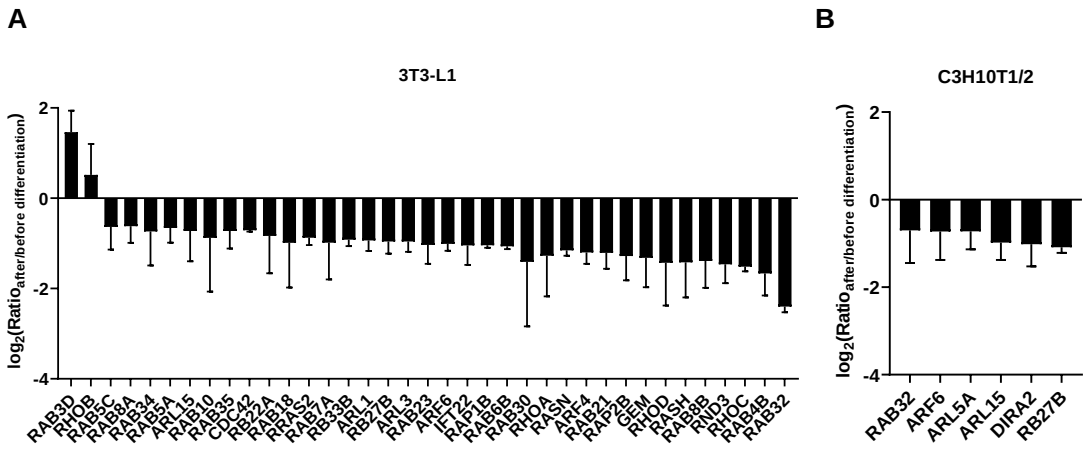


Figure S2.4 Bar chart showing the quantification results for those small GTPases exhibiting >1.5-fold changes during adipogenesis.

The quantification results of (A) 3T3-L1 and (B) C3H10T1/2 cells. The data represent the mean ± standard deviation (S.D.) of results obtained from two (for 3T3-L1 cells) or three (for C3H10T1/2 cells) biological replicates.

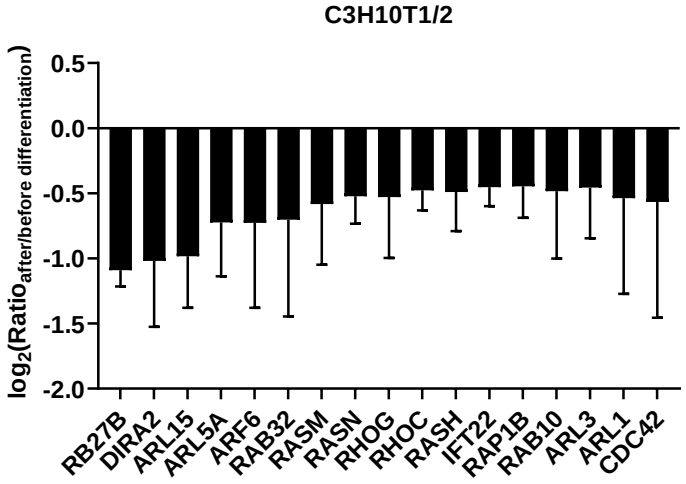


Figure S2.5 Bar chart showing the quantification results for those small GTPases exhibiting >1.3-fold changes during adipogenesis.

The quantification results of C3H10T1/2 cells. The data represent the mean ± standard deviation (S.D.) of results obtained from two (for 3T3-L1 cells) or three (for C3H10T1/2 cells) biological replicates.

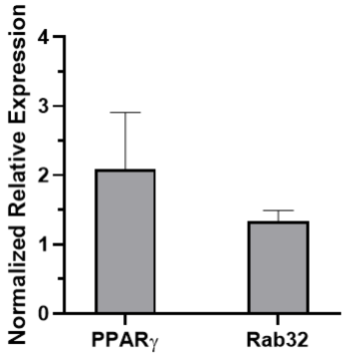


Figure S2.6 The expression of *Rab32* mRNA was unaltered during adipogenesis.

Real-time PCR to show the mRNA expression levels of *Pparg* and *Rab32* genes, which encode PPARγ and Rab32 proteins, respectively, in C3H10T1/2 cells after and before differentiation into adipocytes. The data represent the mean ± standard deviation (S.D.) of results obtained from three biological replicates.

Table S2.1. A complete list of all small GTPases quantified in the scheduled LC-MRM analysis in the 3T3-L1 cells.

Uniprot Accession	Gene	Peptide	Subfamily	Log2		
				Biological #1	Biological #2	Mean Ratio (Adipocyte/Preadipocyte)
P18085	ARF4	LGLQSLR	ARF	-1.02	-1.38	-1.20
P62330	ARF6	FNVWDVGGQDK	ARF	-0.90	-1.12	-1.01
Q13795	ARFRP	DCLTQACSALTGK	ARF	-0.03	-0.77	-0.40
P40616	ARL1	ILILGLDGAGK	ARF	-0.77	-1.10	-0.94
Q9NXU5	ARL15	ELGGADNIR	ARF	-0.25	-1.20	-0.72
P36404	ARL2	ELQSLLVEER	ARF	0.27	0.39	0.33
P36405	ARL3	ILLLGLDNAGK	ARF	-0.80	-1.12	-0.96
Q96BM9	ARL8A	LWDIGGQPR	ARF	0.04	-0.37	-0.16
Q9NVJ2	ARL8B	IWDIGGQPR	ARF	-0.02	-1.00	-0.51
Q9H7X7	IFT22	ILFVGPCESGK	RAB	-0.74	-1.35	-1.05
Q9BW83	IFT27	CILAGDPAVGK	RAB	0.34	-0.82	-0.24
P61026	RAB10	FHTITTSYYR	RAB	-0.04	-1.72	-0.88
P61106	RAB14	TGENVEDAFLEAAK	RAB	-0.19	-0.27	-0.23
P59190	RAB15	IQIWDTAGQER	RAB	-0.24	-0.17	-0.21
Q9NP72	RAB18	TLTPSYYR	RAB	-0.29	-1.69	-0.99
P62820	RAB1A	MGPGATAGGAEK	RAB	-1.34		-1.34
Q9UL25	RAB21	VVLLGEGCVGK	RAB	-0.96	-1.46	-1.21
Q9ULC3	RAB23	TIGVDFLER	RAB	-0.74	-1.33	-1.03
Q969Q5	RAB24	AQLFETSSK	RAB		-0.64	-0.64
P61019	RAB2A	MITIDGK	RAB		-0.89	-0.89
Q15771	RAB30	IVLIGNAGVGK	RAB	-0.40	-2.42	-1.41
Q13637	RAB32	VLVIGELGVGK	RAB	-2.49	-2.31	-2.40
Q9BZG1	RAB34	AEYWAVSSLTGENVR	RAB	-0.20	-1.27	-0.73
Q15286	RAB35	LLIIGDSGVGK	RAB	-0.45	-1.00	-0.73
P20336	RAB3A	TYSWDNAQVLLVGNK	RAB	-1.11	0.34	-0.38
O95716	RAB3D	LLLIGNSSVGK	RAB	1.80	1.13	1.46
Q86YS6	RAB43	VATELIMR	RAB	-1.32		-1.32
P61018	RAB4B	FLVIGSAGTGK	RAB	-1.32	-2.01	-1.67
P20339	RAB5A	QASPNIVIALSGNK	RAB	-0.43	-0.89	-0.66
P51148	RAB5C	QASPNIVIALAGNK	RAB	-0.28	-0.99	-0.64
Q9NRW1	RAB6B	QITIEEGEQR	RAB	-1.02	-1.11	-1.07
P51149	RAB7A	FQSLGVAFYR	RAB	-0.41	-1.56	-0.98
P61006	RAB8A	LALDYGIK	RAB	-0.36	-0.88	-0.62
Q92930	RAB8B	NIEEHASSDVER	RAB	-0.97	-1.81	-1.39
Q15907	RB11B	NILTEIYR	RAB	-0.15	-0.26	-0.20
Q9UL26	RB22A	TVQYQNELHK	RAB	-0.25	-1.42	-0.84
O00194	RB27B	LLALGDSGVGK	RAB	-1.15	-0.77	-0.96
Q9H082	RB33B	SAIQVPTDLAQK	RAB	-0.82	-1.02	-0.92
Q12829	RB40B	EIDEHAPGVPK	RAB		0.12	0.12
Q96S21	RB40C	VFSLQDLCCR	RAB	3.17	-0.95	1.11
Q9NYS0	KBRS1	TLIEPFTLLASK	RAS	-0.52	0.18	-0.17
Q9NYR9	KBRS2	VVVCGQASVGK	RAS	0.85	-0.63	0.11
P62834	RAP1A	QWCNCAFLESSAK	RAS	0.08	-0.41	-0.16
P61224	RAP1B	QWNNCAFLESSAK	RAS	-1.00	-1.08	-1.04
P61225	RAP2B	ASVDELFAEIVR	RAS	-1.66	-0.90	-1.28
P01112	RASH	SFEDIHQYR	RAS	-0.87	-1.97	-1.42
O14807	RASM	LVVVGDGGVGK	RAS		-0.44	-0.44
P01111	RASN	SFADINLYR	RAS	-1.24	-1.07	-1.15
Q15382	RHEB	ALAESWNAAFLESSAK	RAS	-0.27	0.26	-0.01
P10301	RRAS	LFTQILR	RAS	-0.08	-1.01	-0.55
P62070	RRAS2	QVTQEEGQQLAR	RAS	-0.76	-0.99	-0.87
P55040	GEM	VVLIGEQGVGK	RGK	-0.86	-1.78	-1.32
P60953	CDC42	TCLLISYTTNK	RHO	-0.69	-0.73	-0.71
P61586	RHOA	IGAFGYMECSAK	RHO	-0.64	-1.91	-1.28
P62745	RHOB	IQAYDYLECSAK	RHO	1.00	0.03	0.51
P08134	RHOC	ISAFGYLECSAK	RHO	-1.44	-1.59	-1.51
O00212	RHOD	SVGAVAYLECSAR	RHO	-0.76	-2.10	-1.43
P61587	RND3	IVVVGDSQCGK	RHO	-1.17	-1.76	-1.46

Table S2.2. A complete list of all small GTPases quantified in the scheduled LC-MRM analysis in the C3H10T1/2 cells.

Uniprot Accession	Gene	Peptide	Subfamily	Log2 (Ratio)			
				Biological #1	Biological #2	Biological #3	Mean Ratio (Adipocyte/Preadipocyte)
P18085	ARF4	LGLQSLR	ARF	-1.11	0.03	0.13	-0.32
P84085	ARF5	LGLQHLR	ARF	-0.63	0.12	0.17	-0.11
P62330	ARF6	FNVWDVGGQDK	ARF	-1.45	-0.18	-0.55	-0.72
P40616	ARL1	ILILGLDGAGK	ARF	-1.59	-0.13	0.02	-0.56
Q9NXU5	ARL15	ELGGADNIR	ARF	-1.44	-0.75	-0.76	-0.98
P36404	ARL2	ELQSLLVEER	ARF	-0.26	0.69	0.16	0.20
P36405	ARL3	ILLLGLDNAGK	ARF	-0.90	-0.17	-0.30	-0.46
Q9Y689	ARL5A	AGLLIFANK	ARF	-1.19	-0.58	-0.40	-0.72
Q96BM9	ARL8A	LWDIGGQPR	ARF	-1.04	-0.14	-0.02	-0.40
Q9NVJ2	ARL8B	IWDIGGQPR	ARF	-1.00	-0.13	0.00	-0.38
Q9H7X7	IFT22	ILFVGPCESGK	RAB	-0.45	-0.31	-0.60	-0.45
Q9BW83	IFT27	CILAGDPAVGK	RAB	-0.40	-0.21	-0.40	-0.33
P61026	RAB10	FHTITTSYYR	RAB	-1.08	-0.21	-0.16	-0.49
P61106	RAB14	TGENVEDAFLEAAK	RAB	-0.72	-0.05	0.44	-0.11
P59190	RAB15	IQIWDTAGQER	RAB	-0.27	0.97	-0.53	0.06
Q9NP72	RAB18	TLTPSYYR	RAB	-0.41	0.11	0.45	0.05
Q9UL25	RAB21	VVLLGEGCVGK	RAB	-0.56	0.04	-0.45	-0.33
Q9ULC3	RAB23	TIGVDFLER	RAB	-0.50	0.31	-0.24	-0.15
Q969Q5	RAB24	AQLFETSSK	RAB	-0.57	0.44	0.40	0.09
P51157	RAB28	VAAEILGIK	RAB	0.34	0.02	-0.15	0.07
Q13637	RAB32	VLVIGELGVGK	RAB	-1.56	-0.32	-0.23	-0.70
Q9BZG1	RAB34	AEYWAVSSLTGENVR	RAB	0.08	-0.17	0.35	0.09
Q15286	RAB35	LLIIGDSGVGK	RAB	-0.80	-0.18	-0.10	-0.36
P61018	RAB4B	FLVIGSAGTGK	RAB	-0.15	0.20	-0.07	-0.01
P20339	RAB5A	QASPNIVIALSGNK	RAB	-0.25	0.23	0.04	0.01
P51148	RAB5C	QASPNIVIALAGNK	RAB	-0.36	0.17	0.17	-0.01
Q9NRW1	RAB6B	QITIEEGEQR	RAB	-0.24	0.40	-0.47	-0.10
P51149	RAB7A	FQSLGVAFYR	RAB	-0.76	0.04	-0.25	-0.33
P61006	RAB8A	LALDYGIK	RAB	-0.58	0.07	0.24	-0.09
Q15907	RB11B	NILTEIYR	RAB	-0.63	0.16	-0.21	-0.22
O00194	RB27B	LLALGDSGVGK	RAB	-1.08	-1.22	-0.97	-1.09
Q9H082	RB33B	SAIQVPTDLAQK	RAB	-0.29	0.04	-0.11	-0.12
Q12829	RB40B	EIDEHAPGVPK	RAB	-0.11			-0.11
P62826	RAN	FNVWDTAGQEK	RAN	-0.61	0.46	0.38	0.08
Q96HU8	DIRA2	VAVFGAGGVGK	RAS	-1.01	-0.52	-1.53	-1.02
Q9NYR9	KBRS2	VVVCGQASVGK	RAS	0.53	0.30	-0.12	0.24
P11234	RALB	AEEWGVQYVETSAK	RAS	-0.31	-0.18	-0.08	-0.19
P62834	RAP1A	QWCNCAFLESSAK	RAS	-0.12	-0.13	-0.12	-0.12
P61224	RAP1B	QWNNCAFLESSAK	RAS	-0.19	-0.48	-0.67	-0.45
P01112	RASH	SFEDIHQYR	RAS	-0.72	-0.15	-0.60	-0.49
O14807	RASM	LVVVGDGGVGK	RAS	-0.23	-0.41	-1.11	-0.59
P01111	RASN	SFADINLYR	RAS	-0.64	-0.28	-0.65	-0.52
P10301	RRAS	LFTQILR	RAS	-0.10	0.03	-0.54	-0.21
P62070	RRAS2	QVTQEEGQQLAR	RAS	-0.53	0.20	-0.23	-0.19
Q9BPW5	RSLBB	TSLIPRPK	RAS	-0.18	-0.74	0.12	-0.27
P60953	CDC42	TCLLISYTTNK	RHO	-1.19	0.26	-0.68	-0.54
P62745	RHOB	IQAYDYLECSAK	RHO	-0.44	0.47	-0.14	-0.04
P08134	RHOC	ISAFGYLECSAK	RHO	-0.59	-0.30	-0.54	-0.48
P84095	RHOG	YLECSALQQDGVK	RHO	-0.97	-0.04	-0.58	-0.53
P61587	RND3	IVVVGDSQCGK	RHO	0.12	-0.36	-0.64	-0.29

3 Chapter 3. Targeted Proteomic Profiling Revealed Roles of Small GTPases during Osteogenic Differentiation

3.1 Introduction

Bone remodeling is a continuous and active process throughout adulthood that maintains the integrity of skeletal system.[1,2] Optimal osteogenic differentiation is essential for health; excessive and inadequate osteogenesis, however, can result in heterotopic ossification and osteoporosis, respectively.[2,3] At the cellular level, osteoblasts are derived from mesenchymal stem cells (MSCs) in adults,[4,5] where osteogenesis of MSCs could be activated by circulating hormones and locally generated cytokines and/or growth factors, e.g., the secreted transforming growth factor β (TGF-β).[6,7] Notably, depending on extracellular stimuli and downstream signaling events, the MSCs can also be differentiated into other lineages, e.g., adipocytes, chondroblasts, and myoblasts.[8] Thus, understanding the molecular mechanisms of osteoblast differentiation and maturation is crucial for understanding the pathogenesis of bone diseases.

Small GTPases of the Ras superfamily are molecular switches that undergo conformational changes between the guanosine triphosphate (GTP)-associated active state and the guanosine diphosphate (GDP)-bound inactive state.[9-11] When bound with GTP, these proteins interact with downstream effector proteins to modulate diverse cellular processes, including cell differentiation, growth, and proliferation.[12,13] Previous studies revealed the involvement of small GTPases in regulating osteogenesis. CDC42, for instance, was found to promote osteogenesis in adult bone marrow-derived MSCs.[14] In addition, the intraflagellar transport protein 80 (IFT80) is upregulated during osteogenic

differentiation *in vivo*, where genetic depletion of IFT80 diminished the expression of osteoblast regulatory genes (e.g., *Runx2*) upon osteogenesis of murine MSCs.[15]

Despite the earlier evidence supporting the roles of small GTPases in modulating osteogenesis, the global picture of their functions in osteogenic differentiation remains poorly investigated. We reason that a systematic analysis of the expression of the entire small GTPase proteome during osteogenesis could offer mechanistic insights into osteogenic differentiation. Unlike the conventional data-dependent acquisition (DDA)-based discovery proteomic analysis, targeted proteomic analysis on triple-quadrupole mass spectrometers operated in the multiple-reaction monitoring (MRM) mode provides excellent sensitivity, reproducibility, and accuracy in protein quantification.[16,17] Targeted proteomics was found to afford increased depth in proteome coverage when compared to the DDA mode.[18]

Here, we applied an MRM-based targeted proteomic approach, together with the use of synthetic stable isotope-labeled peptides as internal standards, to investigate systematically the dynamic changes in expression of the small GTPase proteome during the course of osteogenesis of H9 human embryonic stem cells (hESCs) and C3H10T1/2 murine mesenchymal progenitor cells (mMPCs).[18,19] We were able to quantify more than 50% of the small GTPase proteome in both H9 and C3H10T1/2 cells. Our results suggest that stem cells coordinate the expression of small GTPases to modulate autophagy, thereby promoting osteogenic differentiation. We also found that KRAS regulates osteogenesis through attenuating matrix metalloproteinase 9 (MMP9)-mediated degradation of extracellular matrix (ECM).

3.2 Materials and Methods

3.2.1 Cell Culture and Transfection

C3H10T1/2 cells were cultured in Dulbecco's modified Eagle's medium (DMEM, Thermo Fisher) supplemented with 10% fetal bovine serum (FBS, Thermo Fisher) and 1% penicillin-streptomycin solution (100 IU/mL, GE Healthcare). The cells were maintained at 37°C in a humidified chamber supplemented with 5% CO_2. Osteogenesis was induced as described elsewhere[20]. Briefly, confluent C3H10T1/2 cells were refreshed with osteogenesis medium (complete DMEM supplemented with 50 µg/mL sodium ascorbate, 10 mM β-glycerophosphate, and 200 ng/mL recombinant BMP2 protein) every 72 hr for 7 days. All experiments were conducted for cells within 15 passages.

H9 hESCs (WiCell) were cultured on a Matrigel (BD Bioscience)-coated surface in mTeSR plus medium (Stem Cell Technologies). The cells were maintained at 37°C in a humidified chamber supplemented with 5% CO_2. To subculture, cell colonies were treated with accutase (Innovative Cell Technologies) for 30 sec followed by scraping. The colonies were seeded into fresh culture medium at a 1:6 ratio (v/v). Differentiation was initiated as described previously.[21] Briefly, cells at confluence were incubated in DMEM supplemented with 15% FBS (Atlanta Biologicals), 0.5% penicillin-streptomycin (Gibco), 1% non-essential amino acids (Gibco), and 0.1 mM β-mercaptoethanol (Sigma) for 5 days. Osteogenesis was further induced by supplementing the cells with 50 µg/mL sodium ascorbate (Sigma), 10 mM β-glycerophosphate, and 0.12 µM vitamin D_3 for 15 days. Confluent cells prior to differentiation induction were collected at day-0 as the

undifferentiated control and the cells with differentiation/osteogenesis induction were collected at the indicated time points.

3.2.2 Lentivirus Preparation and Stable Cell Line Generation

HEK293T cells were transfected with pHAGE-KRAS (Addgene #116755) plasmid together with pLTR-G (Addgene #17532) envelope plasmid and pCMV-dR8.2 dvpr (Addgene #8455) package plasmid using TransIT-X2 dynamic transfection system (Mirus). At 18 hr following transfection, the cells were refreshed with DMEM (Gibco) supplemented with 15% FBS (Atlanta Biologicals) and 0.5% penicillin-streptomycin (Gibco) for H9 cells, and DMEM (Thermo Fisher) supplemented with 10% FBS (Thermo Fisher) for C3H10T1/2 cells. Cell culture supernatant containing viral particles was collected at 24 hr after medium change and filtered through a 0.45-μm sterile filter.

For the transduction of H9 cells, virus-containing medium was further supplemented with the aforementioned differentiation/osteogenesis induction cocktails. The cells were subsequently cultured in the differentiation medium containing 1 μg/mL puromycin. C3H10T1/2 cells were transfected with lentivirus for 48 hr, screened with 1 μg/mL of puromycin for a week, and cultured in complete DMEM supplemented with 1 μg/mL of puromycin.

3.2.3 Calcium Assay

Calcium assay was conducted as described previously.[22,23] Cells were cultured in a 48-well plate and 300 μL of EDTA-free RIPA buffer was added to each well to lyse the cells at the indicated days. The cells in RIPA lysis buffer were scraped off and the solution was centrifuged at 16,000 ×g for 30 min at 4°C. The supernatant was collected as RIPA

lysate, which was used to monitor soluble calcium. The RIPA lysate was mixed with Arsenazo III at a 1:3 ratio (v/v) and the mixture was subsequently subjected to UV-vis absorption measurement at 750 nm. To measure the RIPA-insoluble calcium precipitates, 100 µL of 1.0 M HCl was added to the 48-well plate, and the mixture was incubated with shaking at 4°C for 18 hr. The HCl lysate was then transferred to the insoluble pellet collected from the above-mentioned RIPA lysis procedure. The mixture was further incubated with shaking at 4°C for 18 hr and then centrifuged at 16,000 ×g at 4°C for 5 min. The supernatant was mixed with Arsenazo III at a ratio of 1:29 (v/v). The mixture was subjected to UV-vis spectrophotometry analysis at 750 nm. The absorbance values were converted into the quantity of calcium through standard curves, and the amounts of the RIPA-soluble calcium and insoluble calcium in HCl lysate were combined to represent the total calcium amount. The total calcium amount was divided by the protein concentration obtained from Lowry assay (Bio-Rad), and the resulting value was further normalized against that of day-0 to represent relative calcification upon osteogenic differentiation.

3.2.4 Alizarin Red S Staining

Alizarin Red S staining was conducted as described previously.[20] Briefly, cells were washed with phosphate-buffered saline (PBS) twice and fixed in 3.7% formaldehyde in PBS at room temperature for 1 hr. The fixed cells were washed twice with water and stained in Alizarin Red S staining solution (20 mg/mL of Alizarin Red S in water, pH adjusted to 4.2 with NH_4OH, and filtered through a 0.45-µm sterile filter) at room temperature for 45 min. Excess dye was subsequently washed off with water, and the cells were submerged in PBS for imaging.

For quantification, Alizarin Red S was extracted from the stained cells by rocking in 10% acetic acid at room temperature for 30 min. The cells were subsequently scraped off and the cell suspension was incubated at 85°C for 10 min, followed by centrifugation at 20,000 ×g for 15 min at room temperature. The supernatant was collected, mixed with concentrated NH_4OH (30% in water) at a ratio of 1: 0.375 (v/v), and the resulting mixture was subjected to UV-vis spectrophotometry analysis at 405 nm.

3.2.5 Cell Lysis and Proteomic Sample Preparation

To extract total proteins, cells were lysed in ice-chilled CelLytic M cell lysis reagent (Sigma) containing 1% protease inhibitor cocktail (Sigma). The cell lysate was subsequently centrifuged at 16,000 ×g for 30 min at 4°C to remove debris. Supernatants containing total proteins were quantified using Quick Start Bradford Protein Assay (Bio-Rad). Total proteins were mixed with 4× Laemmli SDS loading buffer (Bio-Rad), and the resulting mixtures were boiled for 5 min. Approximately 50 µg of total proteins were resolved on a 14% SDS-PAGE gel and subsequently stained with Coomassie Brilliant Blue R-250.

Proteins were digested in-gel as described previously[24]. Briefly, gel bands including the molecular weight range of 15-37 kDa were excised and cut into 1 mm^3 cubes. The gel pieces were destained sequentially with 25% and 50% acetonitrile in 50 mM ammonium bicarbonate (pH 7.8). Cysteines in proteins were reduced and alkylated by incubating the gel pieces in 10 mM dithiothreitol (DTT) at 37°C for 1 hr and 55 mM iodoacetamide at room temperature in the dark for 20 min, respectively. The proteins were digested in-gel with trypsin at 37°C for 18 hr at an enzyme/protein ratio of 1:100. Peptides were extracted

from the gel pieces by shaking sequentially in 5% and 50% acetonitrile containing 5% acetic acid. The resulting tryptic peptides were desalted with C18 ZipTip (Agilent) and reconstituted in 0.1% formic acid. Approximately 8% of the digestion mixture was spiked-in with a crude pool of synthetic small GTPase peptides (New England Peptide, Inc.; 4 fmol each) with a C-terminal [$^{15}N_2$,$^{13}C_6$]-labeled lysine (+8.0 Da) or [$^{15}N_4$,$^{13}C_6$]-labeled arginine (+10 Da).

3.2.6 Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS)

The MRM-based LC-MS/MS experiments were performed on a TSQ Altis triple-quadrupole mass spectrometer (Thermo) coupled with an UltiMate 3000 UPLC (Thermo) and a Flex nanoelectrospray ion source (Thermo). The spiked-in sample was loaded onto a trapping column (3 cm, 150 µm ID) packed in-house with C18 resin (5 µm in particle size and 120 Å in pore size, Dr. Maisch GmbH HPLC) at a flow rate of 3 µL/min. The peptides were eluted from the trapping column and resolved on a ~25 cm analytical column (75 µm i.d.) packed with C18 resin (3 µm in particle size and 120 Å in pore size, Dr. Maisch GmbH HPLC) using a gradient of 12-40% buffer B (0.1% formic acid and 80% acetonitrile). The flow rate of elution was 0.3 µL/min. The spray voltage was 2 kV and the ion transfer tube temperature was set at 320 °C. The resolution for Q1 and Q3 was set at 0.7 m/z in full-width at half-maximum (FWHM). Collision-induced dissociation (CID) in Q2 was conducted in 1.5 mTorr argon with the relative collision energy calculated from default settings in Skyline (version 19.1.0.193).

The mass spectrometer was scheduled to monitor precursor-product ion transitions for each of the 147 and 137 human and mouse small GTPase peptides, respectively. The

three most abundant y ions detected in MS/MS obtained from data-dependent acquisition (DDA) mode were selected for the quantification of each peptide. Cycle time was 3 s and the retention time (RT) window for each peptide was 3.5 min. The retention times of the eluted peptides were predicted from their normalized retention times (iRT) and the iRT-RT linear regression determined from 9 tryptic peptides of BSA.[25]

The DDA analysis was conducted on a Fusion Lumos Orbitrap tribrid mass spectrometer (Thermo) coupled with an Easy 1000 nLC (Thermo) and a Flex nanoelectrospray ion source (Thermo). The mass spectrometer was equipped with a high-field asymmetric-waveform ion mobility spectrometry (FAIMS), where the compensation voltages (CV) were set at -40, -60, and -80 V. The carrier gas flow was set at 4.2 L/min. The cycle time was 3 sec with each CV scanned for 1 sec. The LC conditions were the same as those described for LC-MRM analysis except that a 160-min linear gradient of 12-34% buffer B was employed. The spray voltage was set as 2 kV and the ion transfer tube temperature was 320°C. MS was acquired with Orbitrap at a resolution of 60,000. Fragmentation was conducted with higher-energy collisional dissociation (HCD) at a relative collisional energy of 30. The MS/MS were acquired in the linear ion trap, where the scan rate was set as rapid.

3.2.7 MRM Data Analysis

All extracted-ion chromatograms (XICs) were manually inspected and filtered with dot-product (dotp) being > 0.7. The peak areas of the product ions were summed and subsequently normalized against those for the corresponding spiked-in heavy isotope-labeled peptide. The resulting ratio of the differentiated cells was further normalized

against that of the undifferentiated cells for the relative quantifications of small GTPases upon osteogenic differentiation.

For H9 hESCs, the protein lysates from each biological replicate were processed (i.e., with SDS-PAGE fractionation, in-gel digestion, and stable isotope-labeled peptide addition) on two separate days, and the ratios obtained from the two technical replicates were averaged to represent the ratio for the biological replicate. In addition, samples from each technical replicate were analyzed by LC-MRM twice, and the ratios obtained from the two LC-MRM runs were averaged to yield the protein ratio for the technical replicate. A total of three and two biological replicates were conducted for the C3H10T1/2 and H9 cells, respectively.

The mass spectrometry proteomics data are deposited to the ProteomeXchange Consortium via the PRIDE[26] partner repository with the dataset identifier (PXD038467) and Panorama (https://panoramaweb.org/pTMsXN.url).

3.2.8 RNA Extraction and RT-qPCR

Total RNA was extracted from approximately 1×10^6 cells with Total RNA Kit I (Omega) and further purified with HiBind RNA minicolumns (VWR). Reverse transcription was subsequently conducted with 1 µg of total RNA, and 50 ng of the resulting cDNA was mixed with Luna qPCR Master Mix (New England Biolabs) and primer sequences (Table S3.1) in a 96-well optical reaction plate (Bio-Rad). Real-time quantitative PCR (RT-qPCR) was conducted on a CFX real-time PCR detection system (Bio-Rad). The sequences of qPCR primers are listed in table S3.1.

3.2.9 Western Blot

Approximately 20 µg of protein lysate was denatured by boiling in Laemmli loading buffer and resolved by SDS-PAGE. The proteins were subsequently transferred from the gel onto a nitrocellulose membrane and blocked with 5% blocker (Bio-Rad). The membrane was hybridized sequentially with primary and horseradish peroxidase (HRP)-conjugated secondary antibody (Table S3.2). HRP signal was detected with Pierce ECL Western Blotting Substrate (Thermo Fisher) on an Odyssey Imaging System (LI-COR Biosciences).

3.2.10 Gelatin Zymography Assay

Cells were washed twice with PBS and then cultured in FBS-free media overnight. The conditioned medium was collected and centrifuged at 16,000 ×g for 30 min at 4°C. Proteins in the supernatant were concentrated by using a centrifugal filter with 10 kDa molecular weight cutoff (VWR) and subsequently mixed with non-reducing sample buffer (125 mM Tris-HCl, pH 6.8, 4% SDS, 20% glycerol, 0.01% bromophenol blue). The resulting protein mixture was resolved on a 7.5% polyacrylamide gel doped with 4 mg/mL gelatin. The gel was washed twice in a washing buffer (50 mM Tris HCl, pH 7.5, 2.5% Triton X-100, 5 mM $CaCl_2$, 1 µM $ZnCl_2$), and subsequently incubated with rocking in a buffer containing 50 mM Tris-HCl (pH 7.5), 5 mM $CaCl_2$, 1 µM $ZnCl_2$ and 1% Triton X-100 at room temperature for 30 min, followed by incubation at 37°C for 24 hr. The resulting gel was stained with Coomassie Blue, and gel images were recorded on an Odyssey Imaging System at the wavelength of 700 nm.

3.2.11 Data Analysis

Statistics were conducted with Prism (version 9.4.1). Two-tailed Student's *t*-test was used unless specified. Hierarchical clustering was conducted by clustermap from the seaborn package (version 0.11.2) in Python. Distance metric used in clustering was set as default. Principal component analysis (PCA) was conducted by PCA method from the Scikit-Learn decomposition package (version 1.1.2) in Python. The relative expression of small GTPases was first processed by StandardScaler on Scikit-Learn package and the resulting scaled relative expression was fit into the PCA method with n-components being set at three. The PCA results were employed for generating 3D scatterplot with Matplotlib package (version 3.5.1) in Python.

3.3 Results

3.3.1 Scheduled LC-MRM Analysis Revealed Altered Expression of Small GTPases During Osteogenic Differentiation of H9 hESCs

Previous studies revealed the roles of several small GTPases, e.g., CDC42 and IFT80, in osteogenesis.[12,27] Nevertheless, the global picture of the involvements of small GTPases in osteogenic differentiation remains largely unexplored. Here, we aim to investigate comprehensively the differential expression of small GTPases during the entire time course of osteogenic differentiation. To achieve high coverage of the small GTPase proteome and to improve quantification accuracy, we implemented a scheduled MRM-based targeted proteomic approach coupled with synthetic stable isotope-labeled standard peptides (Figure 3.1A).[19]

ESCs can be differentiated progressively into osteoblasts (Figure S3.1A).[28,29] To examine the dynamic expression of small GTPases during osteogenic differentiation, we collected cells on different days throughout the differentiation time course (i.e., on days 3, 5, 6, 8, 12, 15, and 20 following osteogenesis induction) for LC-MRM analysis. Together, we identified 93 small GTPases, which represent approximately 60% of the human small GTPase proteome (Figure 3.1B), in a single injection. The majority of the identified small GTPases were detected in both biological replicates, with 83 proteins being commonly quantified (Figure 3.1C). Moreover, the quantification results are highly consistent between the two biological replicates (Figure 3.2A), suggesting the excellent reproducibility of the method.

To better characterize the dynamic expression of small GTPases and reveal their roles in regulating osteogenic differentiation, we conducted hierarchical clustering analysis based on their relative expression with respect to undifferentiated cells (i.e., day-0). The expression profiles for day-3 and day-5 are similar, but are distinct from cells collected after day-6 (Figure 3.2B). Interestingly, this distinction is consistent with the critical time points in differentiation (Figure S3.1A). In this vein, H9 cells were withdrawn from pluripotency upon differentiation induction at day-0 and the osteogenesis was further induced starting at day-5. Our results showed that the expression of a large subset of small GTPases was modulated during the initial differentiation and further systemically altered upon osteogenic induction. To confirm the proteomic results, we conducted Western blot analysis for selected small GTPases (Figure 3.2C-D). The quantification results obtained

from the two orthogonal methods are highly consistent with an $R^2 > 0.8$ (Figure 3.2E), underscoring the quantification accuracy of the LC-MRM method.

Among the quantified small GTPases, RAB32 was continuously up-regulated during the entire time course of osteogenic differentiation (Figure 3.2A-B). On the other hand, we observed previously diminished expression of RAB32 upon adipogenic differentiation of mMPCs, including 3T3-L1 and C3H10T1/2 cells,[12] suggesting an important role of this small GTPase in modulating cell fate decision. The expression of KRAS, on the other hand, was increased during germ layer differentiation, but diminished at later time points during osteogenic differentiation from MPCs (Figure 3.2A), suggesting that the protein may assume multiple roles in different phases of the differentiation process. Other small GTPases, e.g., RAB27B and RAB33A, are up- or down-regulated, respectively (Figure 3.2A-B), throughout the course of differentiation from ESCs to osteoblasts. The results suggest that different small GTPases may assume distinct roles in the osteogenic differentiation of stem cells.

3.3.2 Validation of Small GTPase Expression in Murine C3H10T1/2 mMPCs

Our aforementioned proteomic results demonstrated the potential roles of small GTPases in the osteogenic differentiation of H9 human hESCs. To identify small GTPases with general roles in osteogenesis, we further conducted LC-MRM analysis to explore the osteogenesis-induced differential expression of small GTPases in C3H10T1/2 mMPCs.[20] We confirmed the successful differentiation of C3H10T1/2 cells to osteoblasts by Alizarin Red S staining and by quantifying Alizarin Red S dye extracted from the stained cells (Figure S3.3A). The successful osteogenesis of C3H10T1/2 cells was also manifested by

the mRNA expression levels of osteogenesis markers, including *Runx2* and *Ocn* (Figure S3.3B).

We next monitored the relative expression levels of small GTPases in differentiated and undifferentiated C3H10T1/2 cells using LC-MRM, and our results led to the quantification of 74 small GTPases, representing > 50% of the murine small GTPase proteome (Figure 3.3A). Among the quantified small GTPases, 23 were differentially expressed by at least 1.5-fold based on at least 2 biological replicates (Figure 3.3B-C), suggesting their potential involvement in regulating osteogenic differentiation from mMPCs. We next conducted Western blot analysis for selected small GTPases (Figure S3.4A-B), and the quantification results are again highly consistent with those obtained from the LC-MRM analysis (Figure S3.4C). Among the commonly identified proteins in H9 hESCs and C3H10T1/2 mMPCs, KRAS was consistently down-regulated upon osteogenic differentiation from mesenchymal stage (i.e., day-6 to day-20 for H9 cells) in both cell lines (Figure 3.2A and 3.3C). On the other hand, RAB32 was commonly up-regulated, suggesting its potential roles as a master regulator for mMPC differentiation.[12]

3.3.3 Autophagy-Regulating Small GTPases are Differentially Expressed During Osteogenesis of H9 hESCs

Autophagy, which recycles non-essential or damaged macromolecules and organelles, is a conserved degradation pathway in cells upon encountering stress.[30] Previous studies showed that autophagy is a necessary process in osteoblast differentiation.[31,32] Interestingly, many small GTPases are known to regulate autophagy. For instance, RAB27B and RAB33A were documented to promote and inhibit autophagy,

respectively.[33-35] In line with this notion, our MRM results revealed up- and down-regulations of RAB27B and RAB33A during osteogenesis, respectively (Figure 3.2A-B), which were further confirmed by Western blot analysis (Figure 3.2C-D).

Next, we performed principal component analysis (PCA) for the small GTPases quantified by LC-MRM based on their relative expression to day-0 throughout the differentiation time course. The small GTPases were further categorized according to their known functions in autophagy. Strikingly, small GTPases that promote, inhibit, or exhibit no reported functions in autophagy are distinctly separated in the PCA plot (Figure S3.5), suggesting that hESCs cells may orchestrate the expression of small GTPases to modulate autophagy for successful osteogenic differentiation.

3.3.4 KRAS Regulates Osteogenesis Through Modulating the Integrity of Extracellular Matrix

We observed diminished KRAS expression upon mesenchymal differentiation to osteoblasts in both H9 hESCs and C3H10T1/2 mMPCs (Figure 3.2A and 3.3C). Thus, we hypothesized that reduced expression of KRAS might promote osteogenesis. To test this hypothesis, we overexpressed KRAS in H9 hESCs by using a lentiviral system. Interestingly, the calcification rate was significantly diminished in cells transfected with lentivirus on day-4, which led to overexpression of KRAS starting on approximately day-7 (Figure 3.4A-B). Consistently, we observed a decreased expression of osteocalcin, a protein secreted by mature osteoblasts to assist calcification of extracellular matrix, in these cells (Figure 3.4C). Cells transfected at day-10, which led to overexpression of KRAS starting on day-13, however, did not confer any apparent effect on osteogenesis rate.

During the osteogenesis of H9 hESCs, day-7 corresponds to the stage of MSCs, whereas day-13 represents the time of differentiation of osteoprogenitor cells into osteoblasts (Figure S3.1A). One possible explanation for the differential effects elicited by KRAS overexpression on different days following osteogenesis induction is that KRAS is crucial for MSCs but assumes a less significant role after they become osteoprogenitor cells. To confirm that decreased KRAS expression is required for osteogenesis from MSCs, we overexpressed KRAS in C3H10T1/2 cells (Figure 3.4D). Similar to H9 hESCs, we observed decreased calcification and diminished expression of *Ocn* transcripts in KRAS-overexpressing cells (Figures 3.4E-F and S3.6). Together, our results demonstrated that attenuated KRAS expression is required for optimal osteogenic differentiation from a mesenchymal state.

KRAS was previously shown to regulate the integrity of the extracellular matrix through modulating the activities of matrix metalloproteinases (MMPs).[36,37] Notably, extracellular matrix is involved in the differentiation and maturation of bone tissues. Collagens, especially type-I collagen, are secreted by osteoblasts and further mineralized by calcium-phosphate deposition, and subsequently crystallize with hydroxyapatite.[38,39] Given that KRAS could stimulate the activities of MMPs, diminished expression of KRAS during osteogenesis from mesenchymal cells may be a crucial step toward the formation of mineralized matrices. To examine this possibility, we collected conditioned media for gelatin zymography analysis, and our results showed that MMP9, whose identity was confirmed by LC-MS/MS analysis, displayed elevated activities upon osteogenic differentiation (Figure 3.S7A-B). Next, we analyzed secreted MMP activity from KRAS-

overexpressing H9 hESCs or C3H10T1/2 mMPCs. Interestingly, we found augmented MMP9 activity in cells overexpressing KRAS (Figure 3.4G and S3.7C). In this vein, previous studies revealed that MMP9 could digest type-I collagen, which is the predominant protein in the extracellular matrix for osteoblast mineralization.[40] Together, our LC-MRM results revealed decreased KRAS expression upon mesenchymal differentiation toward osteoblast, and we observed that diminished KRAS is crucial for ECM accumulation.

3.4 Discussion

Small GTPases are among the most important regulators in signal transduction in cells. Earlier studies documented the important roles of a few small GTPases in osteogenesis.[14,15] The expression of the entire superfamily of small GTPase proteins during osteogenic differentiation, however, remained largely unexplored owing to the lack of appropriate techniques.

Here, we employed a recently developed scheduled LC-MRM-based targeted proteomic method, coupled with synthetic stable isotope-labeled peptides,[19] to interrogate the dynamic changes in expression of the entire small GTPase proteome during osteogenesis. We were able to identify approximately 90 small GTPases and quantify over 50% of the small GTPase proteome, suggesting the excellent sensitivity of the method (Figure 3.1B and 3.2A). In addition, scheduled LC-MRM offered high throughput, where we were able to monitor approximately 90% of the small GTPase proteome in a single 60-min LC-MS/MS run. We also revealed that the method provided high accuracy, as validated by Western blot analysis (Figure 3.2C and S3.4C). It is worth noting that H9 cells

are ESCs, which, upon withdrawal from pluripotency, are differentiated sequentially into germ layers, MPCs, osteoprogenitor cells, and finally into osteoblasts (Figure S3.1A). Therefore, our systematic profiling revealed dynamic expression of the small GTPase proteome upon osteogenic differentiation from as early as the embryonic stages, and the proteomic results obtained from this study may enable future investigations about the functions of small GTPases in different stages of differentiation.

Our results validated some known small GTPases in regulating osteogenic differentiation. CDC42, for instance, was shown to promote osteogenesis of MSCs isolated from adult bone marrow.[14] Here, we observed increased CDC42 expression in H9 cells at day-6 post-induction, at which point mesenchymal cells are differentiated into osteoprogenitor cells (Table S3.3). In addition, our proteomic results showed that the expression of RAB32 was up-regulated during osteogenic differentiation of both H9 hESCs (Figure 3.2B-C) and C3H10T1/2 mMPCs (Figure 3.3B-C). Interestingly, we found previously that diminished RAB32 expression is required for successful adipogenesis, which is an alternative differentiation lineage for mesenchymal cells.[12] The fact that RAB32 protein expression displays opposite trends in osteogenesis and adipogenesis suggests that RAB32 could be a master regulator of mesenchymal cell differentiation.

Emerging lines of evidence suggests that an appropriate level of autophagy promotes the differentiation and survival of osteoblast derived from various stem cells.[31,32] An earlier genome-wide analysis of patients with defects in bone density of wrist underscored an association of autophagy with osteoporosis.[41] While previous studies documented the roles of small GTPases in the initiation and maturation of autophagy, little

is known about their involvements in mediating autophagy during osteogenic differentiation. Here, we found, from PCA analysis, that the autophagy-promoting, -suppressing, and -unrelated small GTPases are clustered together (Figure 3.4D). In addition, we observed the up- and down-regulations of RAB27B and RAB33A, respectively, during osteogenesis (Figure 3.2A and 3.2C), which is in keeping with previous findings that RAB27B and RAB33A could promote and suppress autophagy, respectively.[33,42] Thus, stem cells may coordinate the expression of those small GTPases to modulate autophagy during osteogenesis.

We also revealed an attenuated expression of KRAS during osteogenic differentiation of mesenchymal cells (Figure 3.2A and 3.3B), and diminished KRAS expression is commonly required for mesenchymal differentiation of H9 hESCs and C3H10T1/2 mMPCs to osteoblasts (Figure 3.4A, 3.4B, and 3.4E). KRAS was previously shown to regulate ECM integrity through modulating MMP activity, where elevated expression and/or activation of KRAS lead to collagen degradation, thereby promoting the migration and invasion of cancer cells. [36,37] In addition, extracellular matrix, e.g., type-I collagen, is actively secreted during osteogenesis, and deposited type-I collagen further mineralizes to allow ossification.[39] Our results showed that KRAS was down-regulated during the differentiation of mesenchymal cells to osteoblasts. Hence, stem cells may modulate the expression of KRAS for extracellular matrix accumulation.

Our zymography assay results revealed diminished MMP9 activity during osteogenesis (Figure S3.7A), and augmented MMP9 activity in cells overexpressing KRAS (Figure 3.4G and S3.7C). On the grounds that MMP9 was found to degrade type-I

collagen,[40] attenuated MMP9 activity arising from diminished KRAS expression may be crucial for osteogenesis. Interestingly, we observed decreased calcification in H9 cells overexpressing KRAS when transduced at day-4, but not day-10 (Figure 3.4A). Cells transfected at day-4 following differentiation induction would start to overexpress KRAS when cells were undergoing osteogenic differentiation from mesenchymal stage (i.e., day-6). Previous studies showed that ECMs deposited during osteogenesis could initiate a positive feedback loop to promote MSC differentiation,[43,44] and osteogenesis imperfecta is associated with mutations in genes with functions in maintaining ECM integrity.[38] Our study, hence, revealed a novel KRAS-MMP9-mediated pathway, which promotes osteogenesis via remodeling ECM.

In summary, we characterized systematically the involvements of small GTPases in osteogenesis, and our results validated the earlier findings and unveiled previously uncharacterized roles of small GTPases in osteogenesis. We showed that osteogenic differentiation is accompanied with substantial alterations in expression of a number of small GTPases with functions in autophagy. We also discovered a crucial role of KRAS in regulating ECM accumulation for optimal osteogenesis. Down-regulating KRAS expression, which attenuates MMP9 activity so as to promote ECM accumulation and mineralization, is commonly observed in the two distinct cell lines studied herein. Our findings provide important knowledge basis for future therapeutic interventions of human diseases emanating from aberrant osteogenic differentiation, e.g., osteogenesis imperfecta.

3.5 Acknowledgements

The authors would like to thank Prof. Shaochen Chen at the University of California San Diego for providing the C3H10T1/2 cells and helpful discussion with Dr. Rachel Behar from the Stem Cell Core at University of California Riverside for consulting. This work was supported by the National Institutes of Health (R01 CA 236204 to Yinsheng Wang and R01 DE 025330 to Nicole zur Nieden).

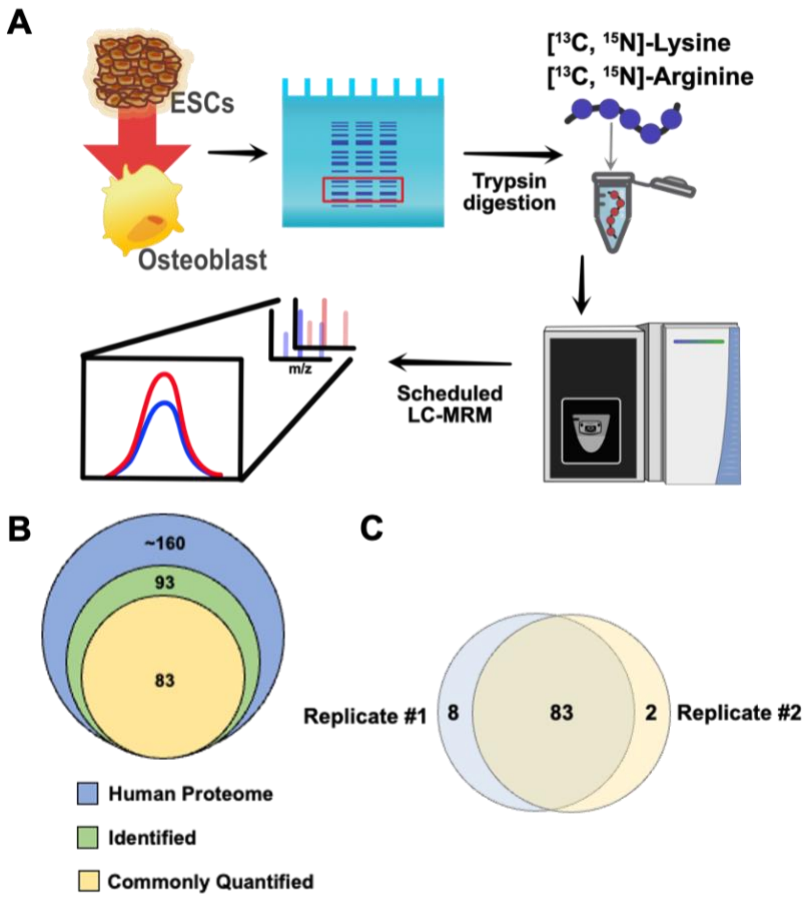


Figure 3.1 High-throughput scheduled LC-MRM for quantifying small GTPase proteins during the osteogenic differentiation of H9 hESCs.

(A) A schematic diagram illustrating the experimental workflow. Osteogenesis was induced in H9 hESCs and the cells were collected at different time points. Total proteins were extracted for in-gel tryptic digestion, and the resulting peptides were spiked-in with a mixture of synthetic stable isotope-labeled small GTPase peptides of identical amino acid sequences prior to sample injection. The MRM data were analyzed using Skyline. (B) A Venn diagram showing the numbers of small GTPases included in the MRM library,

identified, and quantified by LC-MRM analysis. (C) A Venn diagram displaying the number of commonly quantified proteins in the two biological replicates.

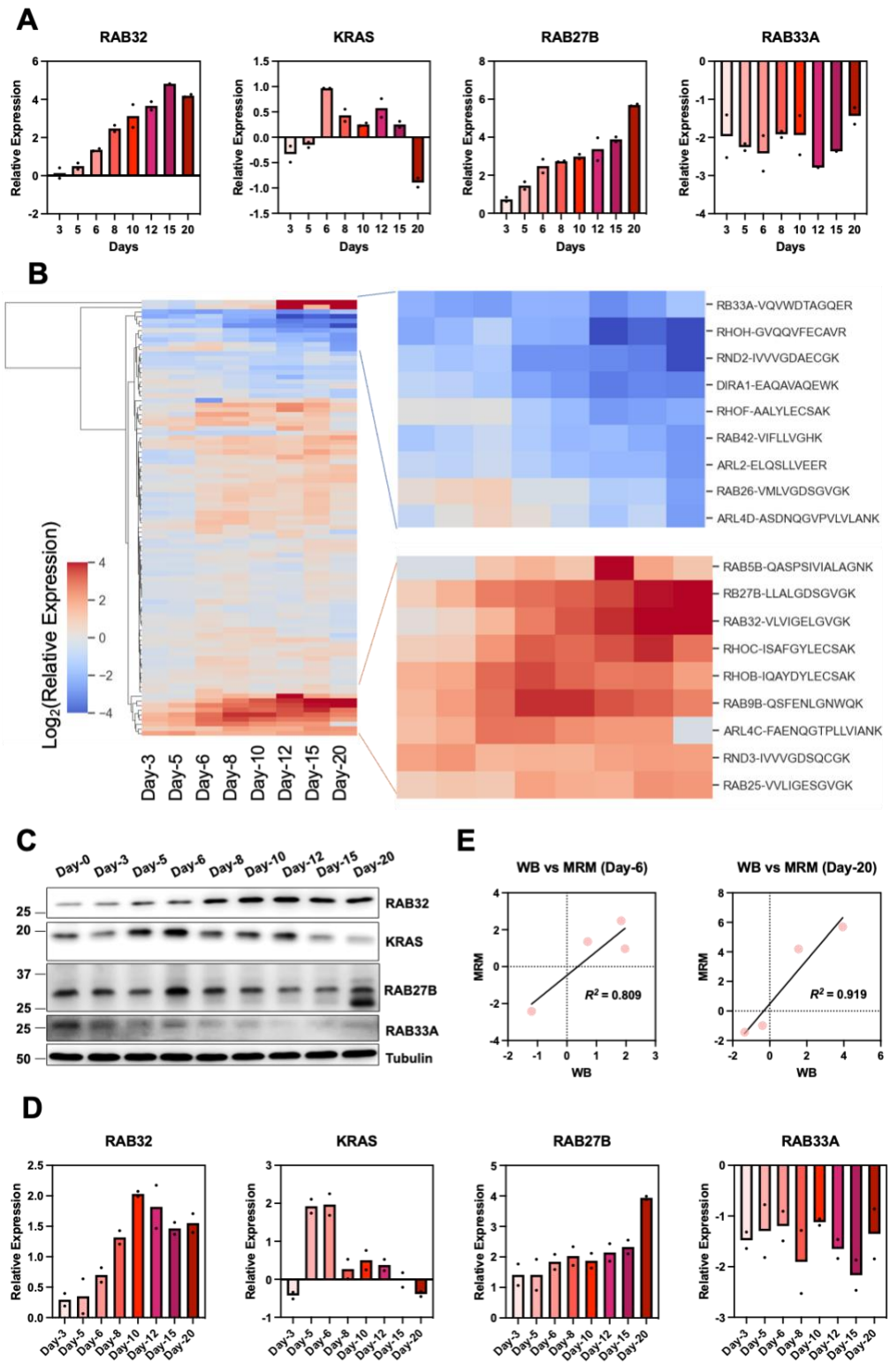

A
RAB32
KRAS
RAB27B
RAB33A
Relative Expression
Days
B
RB33A-VQVWDTAGQER
RHOH-GVQQVFECAVR
RND2-IVVVGDAECGK
DIRA1-EAQAVAQEWK
RHOF-AALYLECSAK
RAB42-VIFLLVGHK
ARL2-ELQSLLVEER
RAB26-VMLVGDSGVGK
ARL4D-ASDNQGVPVLVLANK
RAB5B-QASPSIVIALAGNK
RB27B-LLALGDSGVGK
RAB32-VLVIGELGVGK
RHOC-ISAFGYLECSAK
RHOB-IQAYDYLECSAK
RAB9B-QSFENLGNWQK
ARL4C-FAENQGTPLLVIANK
RND3-IVVVGDSQCGK
RAB25-VVLIGESGVGK
Log2(Relative Expression)
Day-3
Day-5
Day-6
Day-8
Day-10
Day-12
Day-15
Day-20
C
Day-0
RAB32
KRAS
RAB27B
RAB33A
Tubulin
E
WB vs MRM (Day-6)
WB vs MRM (Day-20)
MRM
WB
R2 = 0.809
R2 = 0.919
D

Figure 3.2. Differential expression of small GTPases upon osteogenic differentiation of H9 hESCs.

(A) Bar charts showing the relative expression of selected small GTPases, including RAB32, KRAS, RAB27B and RAB33A, during the time course of osteogenic differentiation as compared to day-0. The data represent average values, where the ratios obtained from the two individual biological replicates are marked with dots. (B) A hierarchical clustering heatmap in $\log_2$ ratio to depict the relative expression of small GTPase proteins at different time points with respect to day-0. The corresponding unique peptides used for quantification are listed after the protein name. (C) Western blot validations of the selected small GTPases during the course of osteogenic differentiation of H9 hESCs. (D) Bar charts showing the quantification results of selected small GTPases. The signal intensities of the small GTPases were first normalized against that of tubulin, and the resulting ratios were further normalized against that of control (day-0). (E) Linear regression analysis of the relative expression of selected small GTPases measured by MRM and Western Blot analysis (WB) at day-6 or day-20.

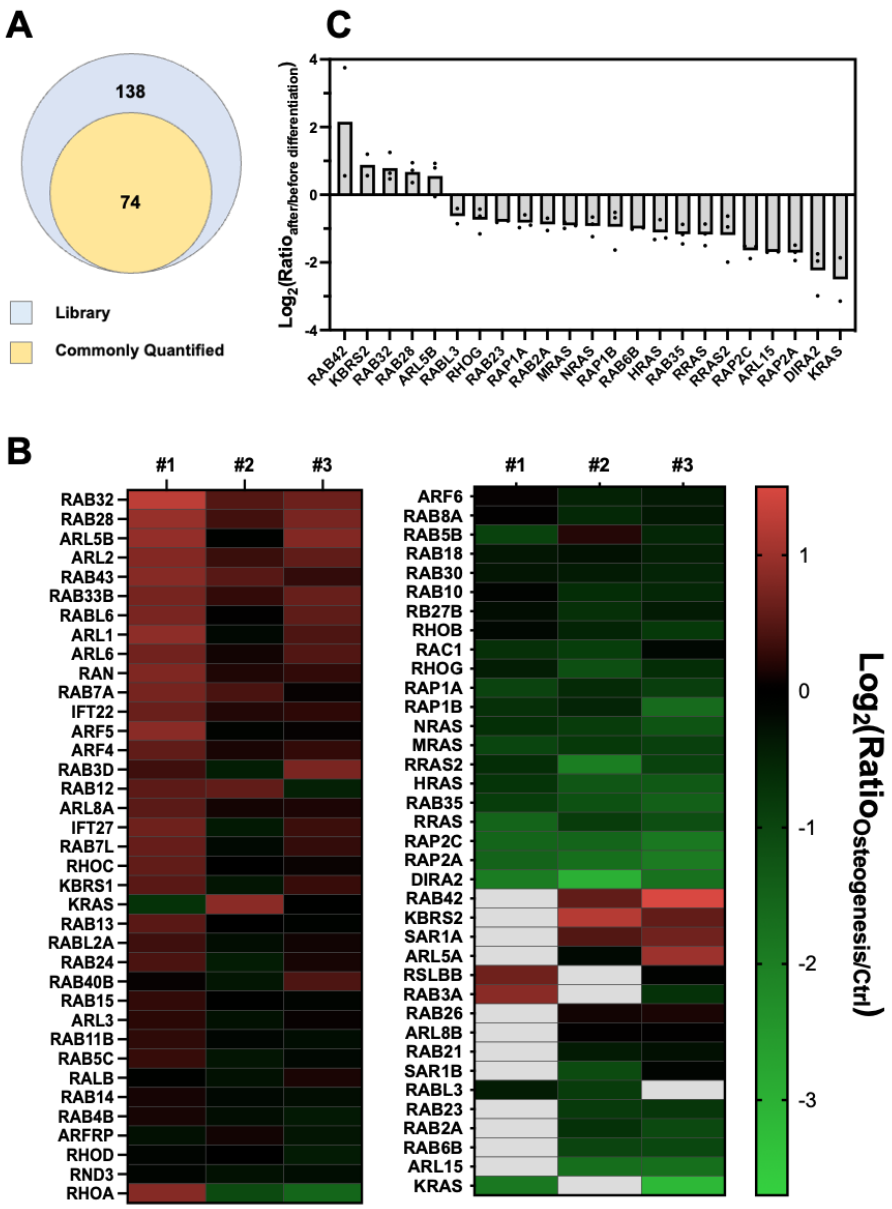


Figure 3.3 Osteogenic differentiation of murine C3H10T1/2 mesenchymal progenitor cells.

(A) A Venn diagram showing the numbers of small GTPases commonly quantified in at least two biological replicates and in the library. (B) A heatmap showing the ratio of small GTPases during osteogenesis of C3H10T1/2 cells quantified from three biological

replicates. Small GTPases with missing data are shown in gray boxes. (C) A bar chart displaying the differentially expressed small GTPases quantified in at least 2 biological replicates and with a cutoff of 1.5-fold.

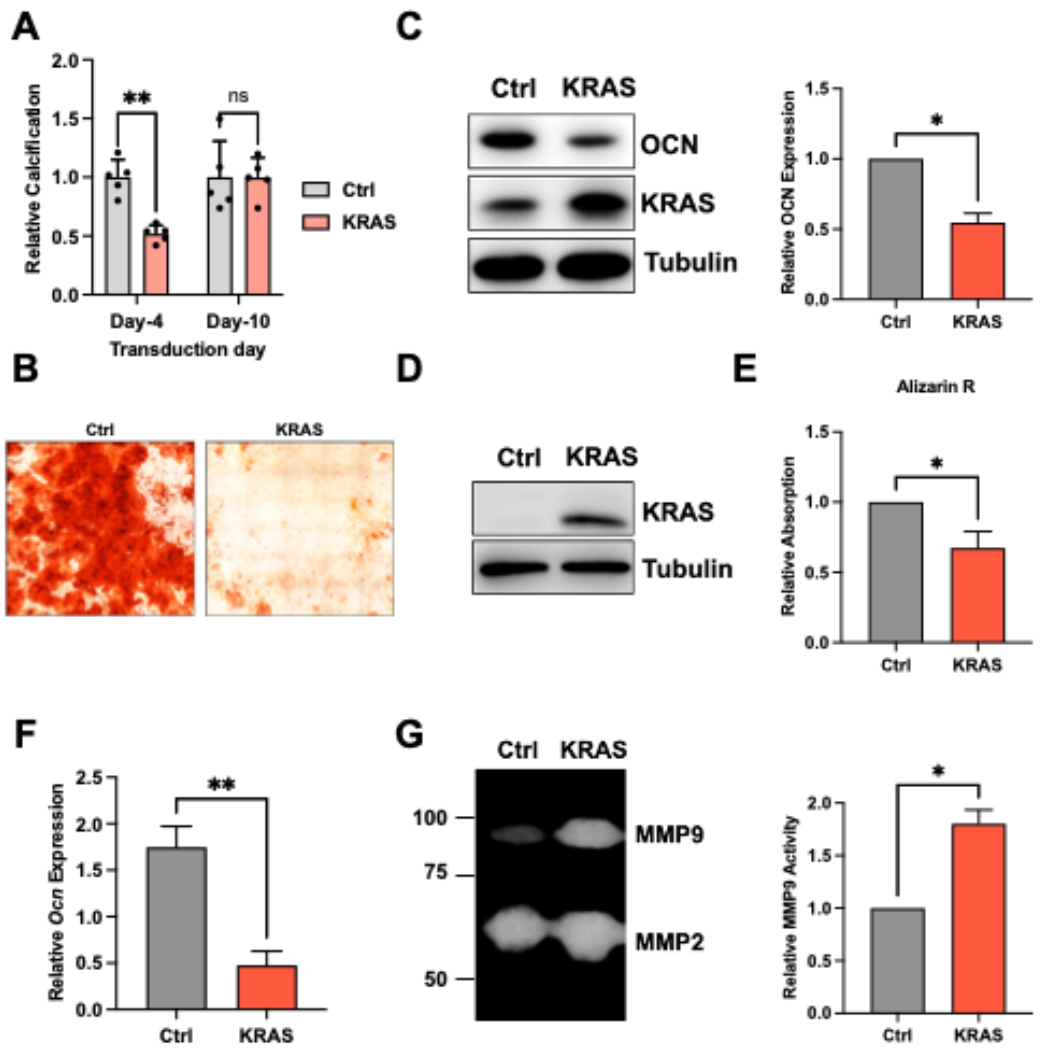


Figure 3.4 KRAS regulates osteogenesis through modulating the integrity of extracellular matrix.

(A) A bar chart displaying the results of calcium assay for H9 cells transduced with lentivirus at day-4 or day-10 for KRAS overexpression. (B) Western blot and the corresponding quantification results showing the successful overexpression of KRAS and the ensuing change in OCN expression in H9 cells at day-15 following transfection at day-4. (C) Alizarin Red S staining illustrating the osteogenesis rate of H9 cells overexpressing

KRAS. (D) Western blot analysis showing the overexpression of KRAS in C3H10T1/2 cells. (E) Quantification results of Alizarin Red S staining showing the osteogenesis rate of C3H10T1/2 cells overexpressing KRAS. (F) RT-qPCR for the quantification of *Ocn* expression in C3H10T1/2 cells overexpressing KRAS. The level of *Ocn* mRNA was normalized against that of *Gapdh* and further normalized against that of control (Ctrl) (G) Gel zymography analysis and the quantification results for the activity of secreted MMP9 from H9 cells overexpressing KRAS. The data represent mean ± S.D. of results obtained from three biological replicates. The *p*-values were calculated using two-tailed, unpaired *t*-test (*, $0.05 < p < 0.01$; **, $0.001 < p < 0.01$; ***, $p < 0.001$; ns, $p > 0.05$).

3.6 References

1. Infante, A. & Rodriguez, C.I. Osteogenesis and aging: lessons from mesenchymal stem cells. *Stem Cell Res. Ther.* **9**, 244 (2018).

2. Raggatt, L.J. & Partridge, N.C. Cellular and molecular mechanisms of bone remodeling. *J. Biol. Chem.* **285**, 25103-8 (2010).

3. Meyers, C. et al. Heterotopic Ossification: A Comprehensive Review. *JBMR Plus* **3**, e10172 (2019).

4. Grottkau, B.E. & Lin, Y. Osteogenesis of Adipose-Derived Stem Cells. *Bone Res.* **1**, 133-45 (2013).

5. Zhang, L. et al. Exosomes from bone marrow mesenchymal stem cells enhance fracture healing through the promotion of osteogenesis and angiogenesis in a rat model of nonunion. *Stem Cell Res. Ther.* **11**, 38 (2020).

6. Xu, C. et al. TGF-b receptor II plays a vital role in neural crest-derived osteogenesis but not in limb osteogenesis during postnatal development. *The FASEB J.* **35**(2021).

7. Infante, A., Cabodevilla, L., Gener, B. & Rodriguez, C.I. Circulating TGF-beta Pathway in Osteogenesis Imperfecta Pediatric Patients Subjected to MSCs-Based Cell Therapy. *Front. Cell Dev. Biol.* **10**, 830928 (2022).

8. Andrzejewska, A., Lukomska, B. & Janowski, M. Concise Review: Mesenchymal Stem Cells: From Roots to Boost. *Stem Cells* **37**, 855-864 (2019).

9. Vetter, I.R. & Wittinghofer, A. The guanine nucleotide-binding switch in three dimensions. *Science* **294**, 1299-1304 (2001).

10. Wennerberg, K., Rossman, K.L. & Der, C.J. The Ras superfamily at a glance. *J. Cell Sci.* **118**, 843-6 (2005).

11. Mishra, A.K. & Lambright, D.G. Invited review: Small GTPases and their GAPs. *Biopolymers* **105**, 431-48 (2016).

12. Yang, Y.Y., Huang, M. & Wang, Y. Targeted proteomic analysis of small GTPases in murine adipogenesis. *Anal. Chem.* (2020).

13. Villarroel-Campos, D., Bronfman, F.C. & Gonzalez-Billault, C. Rab GTPase signaling in neurite outgrowth and axon specification. *Cytoskeleton* **73**, 498-507 (2016).

14. Gao, L., Gorski, J.L. & Chen, C.S. The Cdc42 guanine nucleotide exchange factor FGD1 regulates osteogenesis in human mesenchymal stem cells. *Am. J. Pathol.* **178**, 969-74 (2011).

15. Yang, S. & Wang, C. The intraflagellar transport protein IFT80 is required for cilia formation and osteogenesis. *Bone* **51**, 407-17 (2012).

16. Shi, T. et al. Advances in targeted proteomics and applications to biomedical research. *Proteomics* **16**, 2160-82 (2016).

17. Ebhardt, H.A., Root, A., Sander, C. & Aebersold, R. Applications of targeted proteomics in systems biology and translational medicine. *Proteomics* **15**, 3193-208 (2015).

18. Huang, M., Qi, T.F., Li, L., Zhang, G. & Wang, Y. A targeted quantitative proteomic approach assesses the reprogramming of small gtpases during melanoma metastasis. *Cancer Res.* **78**, 5431-5445 (2018).

19. Huang, M., Darvas, M., Keene, C.D. & Wang, Y. Targeted Quantitative Proteomic Approach for High-Throughput Quantitative Profiling of Small GTPases in Brain Tissues of Alzheimer's Disease Patients. *Anal. Chem.* **91**, 12307-12314 (2019).

20. Mukherjee, A. & Rotwein, P. Akt promotes BMP2-mediated osteoblast differentiation and bone development. *J Cell Sci* **122**, 716-26 (2009).

21. Martinez, I.K., Bhanu, B. & Zur Nieden, N.I. Video-based calcification assay: A novel method for kinetic analysis of osteogenesis in live cultures. *MethodsX* **8**, 101265 (2021).

22. zur Nieden, N.I., Price, F.D., Davis, L.A., Everitt, R.E. & Rancourt, D.E. Gene profiling on mixed embryonic stem cell populations reveals a biphasic role for beta-catenin in osteogenic differentiation. *Mol. Endocrinol.* **21**, 674-85 (2007).

23. Davis, L., Dienelt, A. & zur Nieden, N. Absorption based assays for the identification of skeletal cell types. *Methods Mol. Biol.* **690**, 255-272 (2011).

24. Shevchenko, A., Tomas, H., Havlis, J., Olsen, J.V. & Mann, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* **1**, 2856-60 (2006).

25. Escher, C. et al. Using iRT, a normalized retention time for more targeted measurement of peptides. *Proteomics* **12**, 1111-21 (2012).

26. Perez-Riverol, Y. et al. The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.* **50**, D543-D552 (2022).

27. Pavlos, N.J. et al. Tctex-1, a novel interaction partner of Rab3D, is required for osteoclastic bone resorption. *Mol. Cell. Biol.* **31**, 1551-64 (2011).

28. Sparks, N.R.L., Martinez, I.K.C., Soto, C.H. & Zur Nieden, N.I. Low Osteogenic Yield in Human Pluripotent Stem Cells Associates with Differential Neural Crest Promoter Methylation. *Stem Cells* **36**, 349-362 (2018).

29. zur Nieden, N.I., Kempka, G. & Ahr, H.J. In vitro differentiation of embryonic stem cells into mineralized osteoblasts. *Differentiation* **71**, 18-27 (2003).

30. Yim, W.W. & Mizushima, N. Lysosome biology in autophagy. *Cell Discov.* **6**, 6 (2020).

31. Wan, Y., Zhuo, N., Li, Y., Zhao, W. & Jiang, D. Autophagy promotes osteogenic differentiation of human bone marrow mesenchymal stem cell derived from osteoporotic vertebrae. *Biochem. Biophys. Res. Commun.* **488**, 46-52 (2017).

32. Yin, X. et al. Autophagy in bone homeostasis and the onset of osteoporosis. *Bone Res.* **7**, 28 (2019).

33. Underwood, R. et al. The GTPase Rab27b regulates the release, autophagic clearance, and toxicity of alpha-synuclein. *J. Biol. Chem.* **295**, 8005-8016 (2020).

34. Sekar, M. & Thirumurugan, K. Autophagy: a molecular switch to regulate adipogenesis and lipolysis. *Mol. Cell. Biochem.* **477**, 727-742 (2022).

35. Lu, Q., Wang, P.S. & Yang, L. Golgi-associated Rab GTPases implicated in autophagy. *Cell Biosci.* **11**, 35 (2021).

36. Wang, G. et al. Insulin promotes invasion and migration of KRAS(G12D) mutant HPNE cells by upregulating MMP-2 gelatinolytic activity via ERK- and PI3K-dependent signalling. *Cell Prolif.* **52**, e12575 (2019).

37. Justilien, V. et al. Matrix metalloproteinase-10 is required for lung cancer stem cell maintenance, tumor initiation and metastatic potential. *PLoS One* **7**, e35040 (2012).

38. Lin, X., Patil, S., Gao, Y.G. & Qian, A. The Bone Extracellular Matrix in Bone Formation and Regeneration. *Front. Pharmacol.* **11**, 757 (2020).

39. Schlesinger, P.H. et al. Cellular and extracellular matrix of bone, with principles of synthesis and dependency of mineral deposition on cell membrane transport. *Am. J. Physiol. Cell Physiol.* **318**, C111-C124 (2020).

40. Bigg, H.F., Rowan, A.D., Barker, M.D. & Cawston, T.E. Activity of matrix metalloproteinase-9 against native collagen types I and III. *FEBS J.* **274**, 1246-55 (2007).

41. Zhang, L. et al. Pathway-based genome-wide association analysis identified the importance of regulation-of-autophagy pathway for ultradistal radius BMD. *J. Bone Miner. Res.* **25**, 1572-80 (2010).

42. Itoh, T. et al. Golgi-resident small GTPase Rab33B interacts with Atg16L and modulates autophagosome formation. *Mol. Biol. Cell* **19**, 2916-25 (2008).

43. Harvestine, J.N. et al. Extracellular matrix-coated composite scaffolds promote mesenchymal stem cell persistence and osteogenesis. *Biomacromolecules* **17**, 3524-3531 (2016).

44. Zhang, Z. et al. Bone marrow stromal cell-derived extracellular matrix promotes osteogenesis of adipose-derived stem cells. *Cell Biol. Int.* **39**, 291-9 (2015).

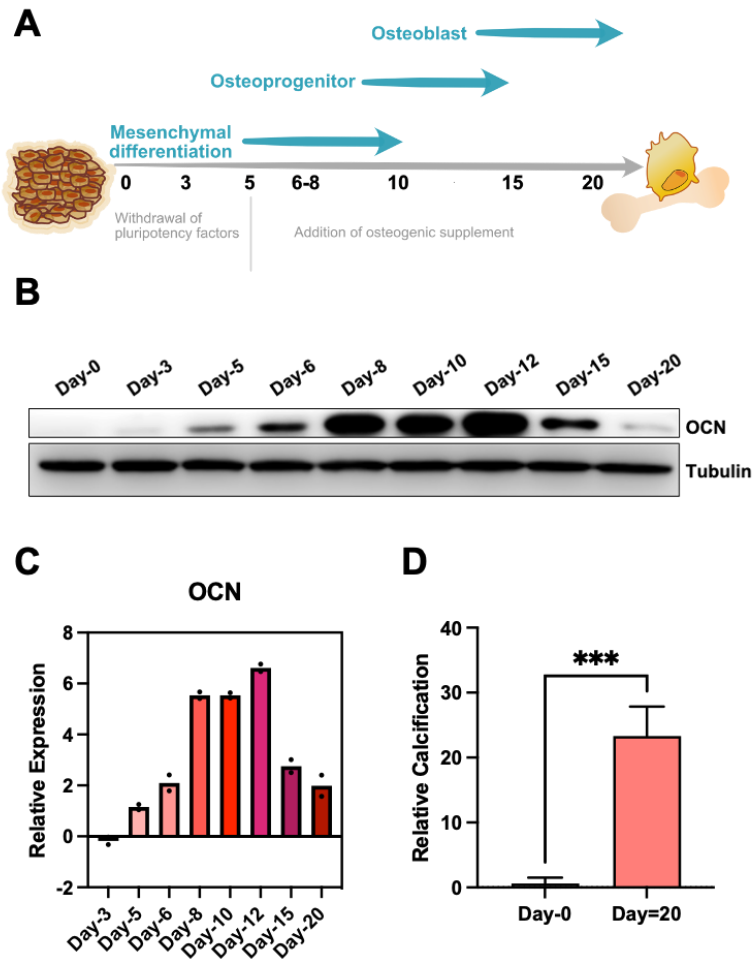


Figure S3.1. Osteogenic differentiation of H9 ESCs.

(A) A schematic diagram depicting the differentiation time course of H9 human ESCs. (B-C) Western blot analysis and the quantification results of osteocalcin (OCN) expression throughout the osteogenic differentiation of H9 cells. (D) Calcium assay to quantify calcification of cells collected during osteogenic differentiation of H9 hESCs. The average calcium quantity was further normalized to the undifferentiated cells (i.e., day-0). The data

represent mean ± S.D. of results obtained from five biological replicates. The *p*-values were calculated using two-tailed, unpaired t-test (***, $p < 0.001$; ns, $p > 0.05$).

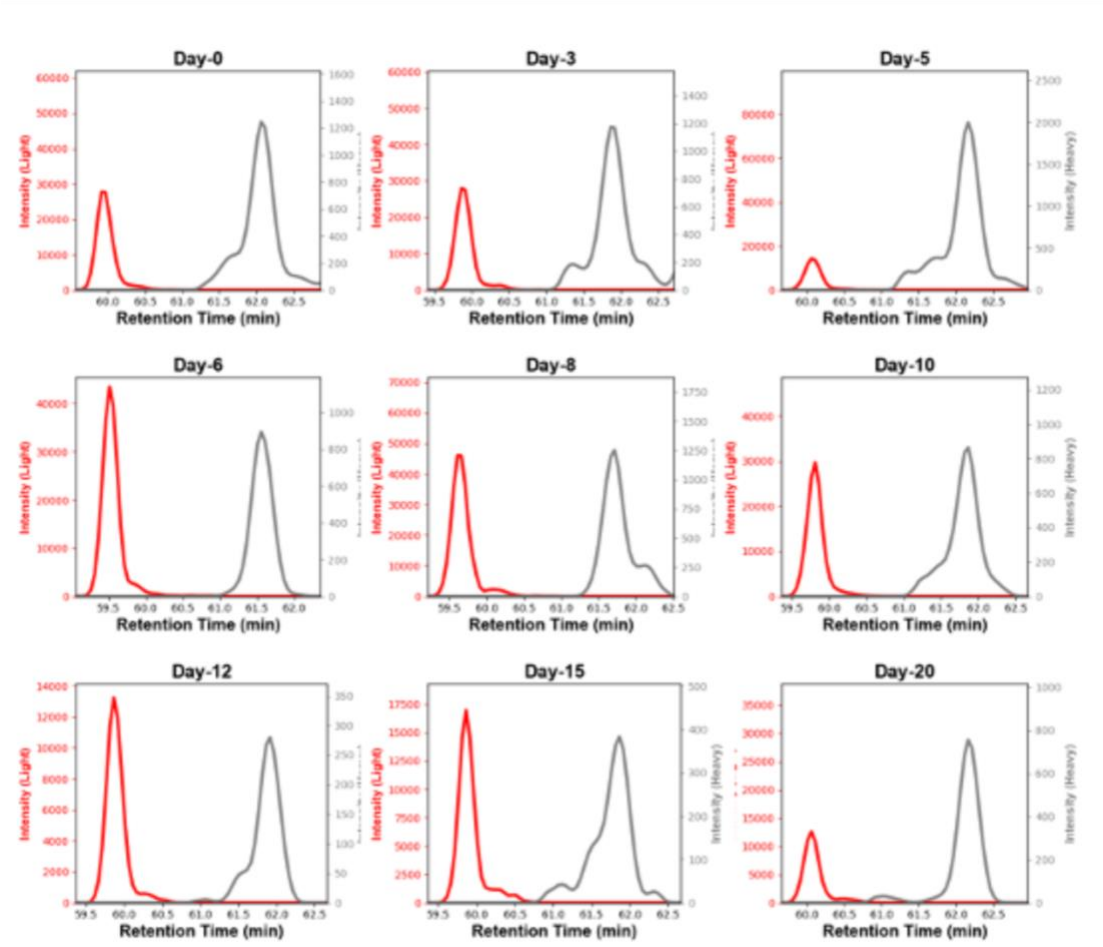


Figure S3.2 Extracted-ion chromatograms of KRAS in H9 cells.

The red traces represent the unlabeled peptides of KRAS from the samples, and the grey traces represent the spiked-in stable isotope-labeled surrogate peptides. The color-coded dual y-axis represents the peak intensities for peptides from endogenous KRAS protein (red) and spiked-in stable isotope-labeled peptides (grey).

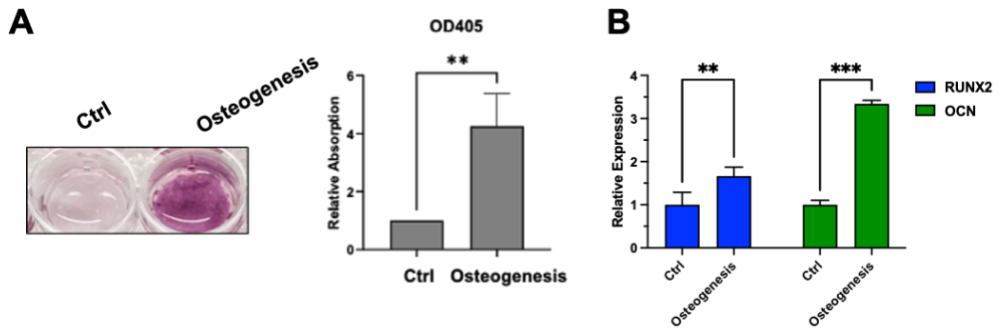


Figure S3.3 Osteogenic differentiation of murine C3H10T1/2 mesenchymal stem cells. (A) An image depicting the Alizarin Red S staining of C3H10T1/2 cells with (Osteogenesis) or without (Ctrl) osteogenesis induction. The Alizarin Red S dye was extracted from the stained cells for UV-vis spectrophotometry analysis at 405 nm. The absorbance was normalized against the control (Ctrl). (B) RT-qPCR analysis of osteogenesis markers Runx2 and Oxn. The levels of the mRNAs of selected osteogenesis markers were normalized against that of the GAPDH gene and further normalized against that of control (Ctrl). The data represent mean ± SD of results obtained from three biological replicates. The p-values were calculated using two-tailed, unpaired Student's t-test (**, $0.001 < p < 0.01$; ***, $p < 0.001$; ns, $p > 0.05$).

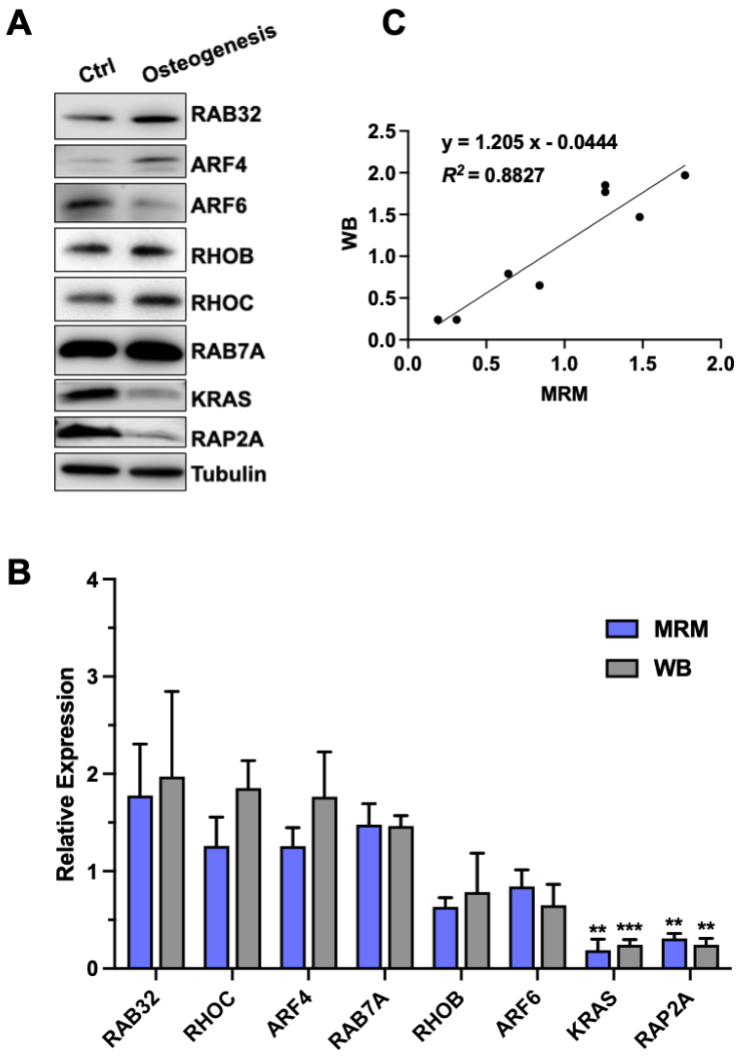


Figure S3.4 Validation of differential expression of small GTPases upon osteogenic differentiation of C3H10T1/2 mMPCs.

(A) Western blot validation of selected small GTPases during osteogenic differentiation of C3H10T1/2 cells. (B) A bar chart comparing the quantification results obtained from MRM (blue) and Western blot (grey) analyses of selected small GTPases during the osteogenic differentiation of C3H10T1/2 cells. (C) A linear regression displaying the correlation of the quantification results obtained from MRM and Western blot analyses of selected small GTPases during the osteogenesis of C3H10T1/2 cells. The data represent mean ± S.D. of

results obtained from three biological replicates. The *p*-values were calculated using two-tailed, unpaired *t*-test (**, $0.001 < p < 0.01$; ***, $p < 0.001$).

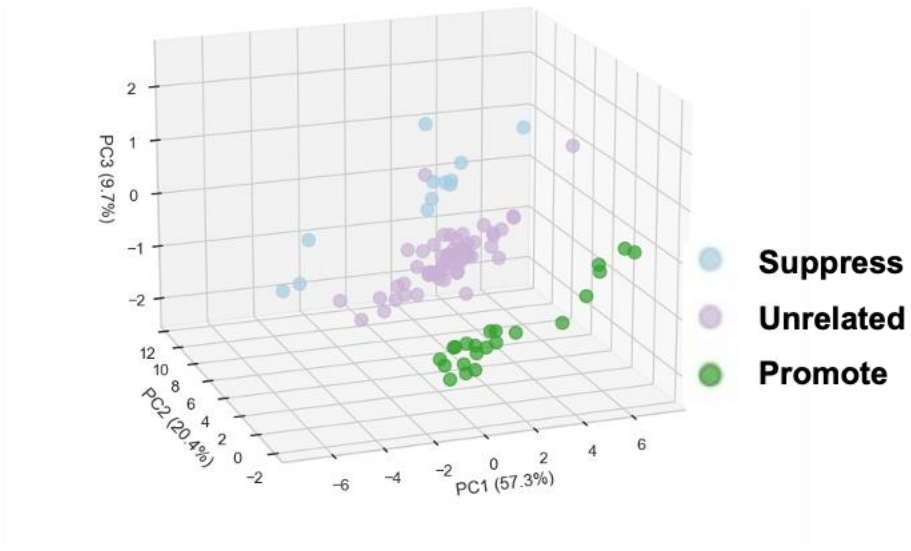


Figure S3.5 Small GTPases involved in autophagy regulation are differentially expressed during osteogenic differentiation of H9 hESCs.

A dot plot showing the PCA analysis of the quantified small GTPases. The small GTPases previously reported to promote, suppress, or unrelated to autophagy are shown in green, cyan, and pink dots, respectively.

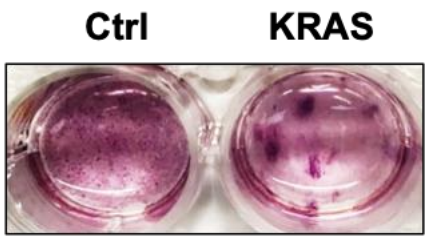


Figure S3.6 Diminished KRAS is required for osteogenesis in C3H10T1/2 mESCs. Alizarin Red S staining and the corresponding quantification results for control (Ctrl) and KRAS-overexpressing cells. The Alizarin Red S dye was extracted from the stained cells for the UV-vis analysis at 405 nm (Figure 3.4E).

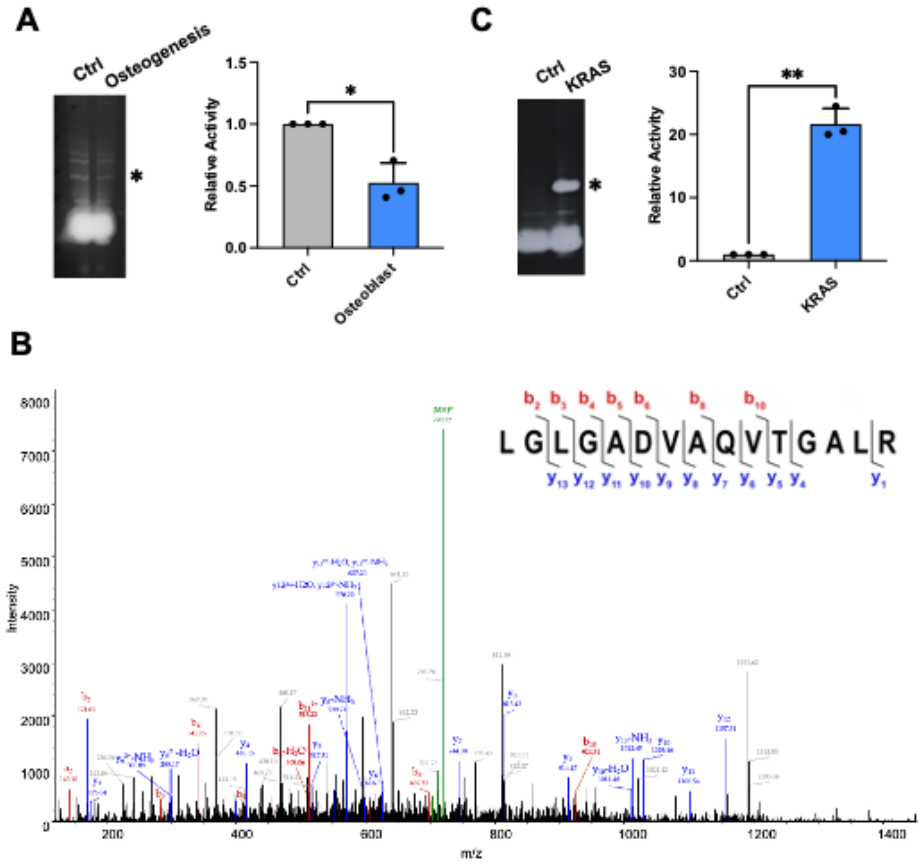


Figure S3.7 KRAS modulated the integrity of extracellular matrices.

(A) Zymography assay for analyzing the activities of secreted MMPs from C3H10T1/2 cells with (Osteogenesis) or without (Ctrl) osteogenesis induction. The band intensities of active MMP9 (arrow) were quantified and normalized against that of Ctrl. (B) LC-MS/MS analysis revealed that the band marked by asterisk corresponds to MMP9. (C) Zymography assay showed that the activities of secreted MMP9 from control cells (Ctrl) or cells overexpressing KRAS (KRAS). The band intensities of active MMP9 (asterisk) were quantified and normalized against that of Ctrl. The data represent mean ± S.D. of results

obtained from three biological replicates. The p-values were calculated using two-tailed, unpaired Student's t-test (*, $0.01 < p < 0.05$; **, $0.001 < p < 0.01$).

Table S3.1 A list of primer sequences used in RT-qPCR experiments.

Transcripts	Species	Forward	Reverse
Ocn	Mouse	CAAGTCCCACACAGCAGCTT	AAAGCCGAGCTGCCAGAGTT
Runx2	Mouse	GCCGGGAATGATGAGAACTA	GGACCGTCCACTGTCACTTT
Gapdh	Mouse	CATCACTGCCACCCAGAAGACTG	ATGCCAGTGAGCTTCCCGTTCAG

Table S3.2 Primary and secondary antibodies used in Western blot analysis.

Targets		Vendor	Cat. No.	Dilution
Primary	RAB27B	Abclonal	A10389	1:1000
	RAB33A	ProteinTech	11025-1-AP	1:1000
	KRAS	ProteinTech	1263-1-AP	1:1000
	RAB32	Thermo Fisher	PA5-48865	1:2000
	ARF4	ProteinTech	11673-1-AP	1:2000
	ARF6	Santa Cruz	sc-7971	1:1000
	RHOB	Abclonal	A2892	1:1000
	RHOC	Abclonal	A1062	1:1000
	RAB7A	Abclonal	A12784	1:1000
	RAP2A	ProteinTech	13789-1-AP	1:1000
	OCN	ProteinTech	23418-1-AP	1:1000
	Beclin-1	ProteinTech	11306-1-AP	1:1000
	p62	ProteinTech	18420-1-AP	1:1000
	LC3	Santa Cruz	sc-376404	1:1000
	TUBULIN	Santa Cruz	sc-32293	1:5000
Secondary	Mouse IgG-HRP	Santa Cruz	sc-2005	1:4000
	Rabbit IgG-HRP	Thermo Fisher	31460	1:4000

Table S3.3 A complete lists of ratios of expression levels of all small GTPases during the entire time course of osteogenic differentiation of H9 hESCs, as quantified by LC-MRM analysis (relative to Day-0).

UniProt Accession	Protein	Peptide	Family	Day-3			Day-5			Day-6			Day-8		
				Replicate #1	Replicate #2	Average	Replicate #1	Replicate #2	Average	Replicate #1	Replicate #2	Average	Replicate #1	Replicate #2	Average
P18085	**ARF4**	LGLQSLR	ARF	1.12	1.03	**1.07**	1.20	0.84	**1.02**	1.16	1.30	**1.23**	1.17	1.72	**1.44**
P84085	**ARF5**	LGLQHLR	ARF	0.99	0.97	**0.98**	1.03	0.95	**0.99**	1.30	1.48	**1.39**	1.65	1.44	**1.54**
P62330	**ARF6**	FNVWDVGGQDK	ARF	1.05	0.69	**0.87**	0.94	0.85	**0.89**	1.57	2.30	**1.93**	1.85	2.97	**2.41**
Q13795	**ARFRP**	DCLTQACSALTGK	ARF	1.04	0.98	**1.01**	0.87	0.90	**0.88**	1.14	1.32	**1.23**	1.27	0.97	**1.12**
P40616	**ARL1**	ILILGLDGAGK	ARF	0.69	0.66	**0.67**	0.76	0.74	**0.75**	0.72	1.26	**0.99**	0.86	1.63	**1.25**
Q8N8L6	**ARL10**	EVLVLGLDGAGK	ARF	1.02		**1.02**	1.18		**1.18**	1.01		**1.01**	1.33		**1.33**
Q9NXU5	**ARL15**	ELGGADNIR	ARF	0.95	0.89	**0.92**	1.26	1.11	**1.18**	1.27	1.28	**1.28**	1.66	1.47	**1.56**
P36404	**ARL2**	ELQSLLVEER	ARF	0.70	0.53	**0.62**	0.55	0.52	**0.53**	0.67	0.74	**0.70**	0.56	0.41	**0.49**
P36405	**ARL3**	ILLLGLDNAGK	ARF	1.04	0.75	**0.90**	0.90	0.86	**0.88**	1.41	2.11	**1.76**	1.53	1.42	**1.47**
P56559	**ARL4C**	FAENQGTPLLVIANK	ARF	2.45	2.05	**2.25**	2.84	3.09	**2.96**	6.22	6.41	**6.31**	6.58	6.03	**6.30**
Q9Y689	**ARL5A**	AGLLIFANK	ARF	0.76	0.76	**0.76**	0.87	0.69	**0.78**	1.06	0.97	**1.01**	0.80	0.75	**0.77**
Q96KC2	**ARL5B**	AAVLIFANK	ARF	0.77	0.75	**0.76**	0.86	0.71	**0.79**	1.09	1.00	**1.04**	0.83	0.74	**0.78**
Q96BM9	**ARL8A**	LWDIGGQPR	ARF	0.95	0.75	**0.85**	0.93	0.86	**0.89**	1.20	1.50	**1.35**	1.19	1.66	**1.43**
Q9NVJ2	**ARL8B**	IWDIGGQPR	ARF	0.89	0.74	**0.82**	0.93	0.88	**0.90**	1.16	1.41	**1.28**	1.49	1.76	**1.63**
Q9NR31	**SAR1A**	EIFGLYGQTTGK	ARF	0.93	0.83	**0.88**	0.93	0.93	**0.93**	1.12	1.77	**1.44**	1.26	1.59	**1.42**
P49703	**ARL4D**	ASDNQGVPVLVLANK	ARF	0.88	0.43	**0.66**	0.85	1.23	**1.04**	1.58	1.53	**1.55**	1.12		**1.12**
Q9H7X7	**IFT22**	ILFVGPCESGK	RAB	1.73	1.25	**1.49**	1.81	1.72	**1.77**	2.60	2.30	**2.45**	2.79	1.25	**2.02**
Q9BW83	**IFT27**	CILAGDPAVGK	RAB	1.10	0.83	**0.97**	1.01	1.02	**1.02**	1.30	1.08	**1.19**	1.19	0.82	**1.01**
P61026	**RAB10**	FHTITTSYYR	RAB	1.00	0.74	**0.87**	0.97	0.88	**0.93**	0.85	1.16	**1.01**	1.01	1.33	**1.17**
Q6IQ22	**RAB12**	LQVIIIGSR	RAB	0.63	0.47	**0.55**	0.80	0.52	**0.66**	1.11	1.18	**1.14**	1.44	1.25	**1.35**
P51153	**RAB13**	LQVWDTAGQER	RAB	1.09	1.08	**1.09**	1.00	0.89	**0.94**	0.68	0.80	**0.74**	0.54	0.71	**0.62**
P61106	**RAB14**	TGENVEDAFLEAAK	RAB	0.74	0.64	**0.69**	0.76	0.78	**0.77**	1.10	1.34	**1.22**	1.59	1.74	**1.67**
P59190	**RAB15**	IQIWDTAGQER	RAB	0.92	0.75	**0.83**	0.77	0.72	**0.74**	0.87	0.98	**0.92**	0.93	1.13	**1.03**
Q9NP72	**RAB18**	TLTPSYYR	RAB	0.97	0.78	**0.87**	0.92	0.88	**0.90**	0.96	0.82	**0.89**	1.08	1.15	**1.12**
A4D1S5	**RAB19**	IILIGDSNVGK	RAB	0.60	0.48	**0.54**	0.51	0.45	**0.48**	0.54	0.46	**0.50**	0.76	0.37	**0.56**
Q9UL25	**RAB21**	VVLLGEGCVGK	RAB	1.11	0.93	**1.02**	1.24	1.10	**1.17**	1.53	1.47	**1.50**	1.97	1.61	**1.79**
Q9ULC3	**RAB23**	TIGVDFLER	RAB	0.92	0.67	**0.79**	0.85	0.84	**0.85**	1.34	1.90	**1.62**	1.43	1.90	**1.66**
Q969Q5	**RAB24**	AQLFETSSK	RAB	0.77	0.73	**0.75**	0.77	0.75	**0.76**	0.82	0.67	**0.75**	1.07	0.71	**0.89**
P57735	**RAB25**	VVLIGESGVGK	RAB	1.87	1.44	**1.66**	1.93	2.05	**1.99**	1.94	1.79	**1.86**	3.08	4.10	**3.59**
Q9ULW5	**RAB26**	VMLVGDSGVGK	RAB	1.17		**1.17**	1.45		**1.45**	1.66		**1.66**	0.86		**0.86**
P51159	**RAB27A**	SWIPEGVVR	RAB	0.72	0.46	**0.59**	0.80	0.62	**0.71**	1.49	2.04	**1.77**	2.05	2.72	**2.38**
O00194	**RAB27B**	LLALGDSGVGK	RAB	1.79	1.54	**1.67**	2.39	3.15	**2.77**	4.35	7.19	**5.77**	6.76	6.54	**6.65**
P51157	**RAB28**	VAAEILGIK	RAB	1.06	0.95	**1.01**	1.15	1.04	**1.10**	1.34	1.25	**1.29**	1.22	0.87	**1.05**
P61019	**RAB2A**	MITIDGK	RAB	0.90	0.83	**0.86**	1.04	0.92	**0.98**	1.20	0.86	**1.03**	1.17	1.05	**1.11**
Q8WUD1	**RAB2B**	IGPQQSISTSVGPSASQR	RAB	1.39	0.99	**1.19**	0.82	0.92	**0.87**	1.16	1.53	**1.34**	1.60	0.91	**1.25**
Q15771	**RAB30**	IVLIGNAGVGK	RAB	0.72	0.51	**0.61**	0.61	0.57	**0.59**	0.84	0.76	**0.80**	0.78	0.72	**0.75**
Q13637	**RAB32**	VLVIGELGVGK	RAB	1.33	0.91	**1.12**	1.59	1.26	**1.42**	2.66	2.45	**2.56**	6.25	4.96	**5.60**
Q14088	**RAB33A**	VQVWDTAGQER	RAB	0.17	0.38	**0.28**	0.20	0.22	**0.21**	0.14	0.26	**0.20**	0.28	0.25	**0.26**
Q9H082	**RAB33B**	SAIQVPTDLAQK	RAB	1.01	0.79	**0.90**	0.94	0.86	**0.90**	0.97	0.88	**0.92**	1.14	1.11	**1.12**
Q15286	**RAB35**	LLIIGDSGVGK	RAB	0.98	0.79	**0.89**	0.92	0.85	**0.88**	0.86	0.93	**0.90**	0.97	1.47	**1.22**
Q96AX2	**RAB37**	VMLLGDTGVGK	RAB		0.65	**0.65**		1.20	**1.20**		1.49	**1.49**		0.76	**0.76**
P20336	**RAB3A**	TYSWDNAQVLLVGNK	RAB	0.72	1.39	**1.05**	0.56	0.78	**0.67**	1.02	0.70	**0.86**	0.80	0.59	**0.69**
O95716	**RAB3D**	LLLIGNSSVGK	RAB	0.64	0.53	**0.59**	0.58	0.55	**0.57**	0.61	0.65	**0.63**	0.63	0.66	**0.64**
Q5JT25-2	**RAB41**	LLFLGEQSGK	RAB	0.89		**0.89**	0.90	4.62	**2.76**	1.66	0.36	**1.01**	1.40	0.45	**0.92**
Q8N4Z0	**RAB42**	VIFLLVGHK	RAB	0.64	0.25	**0.45**	0.51	0.50	**0.51**	0.83	0.54	**0.68**	0.65	0.29	**0.47**

UniProt Accession	Protein	Peptide	Family	Day-10			Day-12			Day-15			Day-20		
				Replicate #1	Replicate #2	Average	Replicate #1	Replicate #2	Average	Replicate #1	Replicate #2	Average	Replicate #1	Replicate #2	Average
P18085	**ARF4**	LGLQSLR	ARF	1.41	1.37	**1.39**	1.19		**1.19**	1.44	1.67	**1.56**	2.00	2.17	**2.08**
P84085	**ARF5**	LGLQHLR	ARF	1.82	1.64	**1.73**	1.36	1.63	**1.50**	1.60	1.84	**1.72**	1.31	1.52	**1.41**
P62330	**ARF6**	FNVWDVGGQDK	ARF	1.84	1.36	**1.60**	1.29	1.04	**1.16**	1.50	1.69	**1.59**	1.22	1.22	**1.22**
Q13795	**ARFRP**	DCLTQACSALTGK	ARF	0.98	0.73	**0.86**	0.88	1.56	**1.22**	1.11	0.83	**0.97**	0.80	0.71	**0.75**
P40616	**ARL1**	ILILGLDGAGK	ARF	1.06	0.77	**0.91**	1.08	0.79	**0.93**	1.29	1.51	**1.40**	1.05	1.05	**1.05**
Q8N8L6	**ARL10**	EVLVLGLDGAGK	ARF	1.38		**1.38**	1.22		**1.22**	1.05		**1.05**	1.39		**1.39**
Q9NXU5	**ARL15**	ELGGADNIR	ARF	1.65	1.12	**1.39**	1.41	1.18	**1.29**	1.77	1.15	**1.46**	0.94	0.92	**0.93**
P36404	**ARL2**	ELQSLLVEER	ARF	0.42	0.28	**0.35**	0.31	0.27	**0.29**	0.33	0.27	**0.30**	0.19	0.16	**0.18**
P36405	**ARL3**	ILLLGLDNAGK	ARF	1.44	1.01	**1.22**	1.34	1.04	**1.19**	1.68	1.84	**1.76**	0.94	0.95	**0.95**
P56559	**ARL4C**	FAENQGTPLLVIANK	ARF	5.89	3.65	**4.77**	3.92	4.05	**3.99**	4.07	3.81	**3.94**	0.98	0.80	**0.89**
Q9Y689	**ARL5A**	AGLLIFANK	ARF	0.84	0.85	**0.84**	0.84	1.25	**1.05**	0.90	0.90	**0.90**	0.77	0.65	**0.71**
Q96KC2	**ARL5B**	AAVLIFANK	ARF	0.84	0.87	**0.85**	0.83	1.07	**0.95**	0.90	0.90	**0.90**	0.73	0.66	**0.69**
Q96BM9	**ARL8A**	LWDIGGQPR	ARF	1.20	1.00	**1.10**	1.11	1.02	**1.07**	1.40	1.63	**1.51**	1.41	1.44	**1.43**
Q9NVJ2	**ARL8B**	IWDIGGQPR	ARF	1.60	1.16	**1.38**	1.51	1.39	**1.45**	1.93	2.05	**1.99**	2.14	2.24	**2.19**
Q9NR31	**SAR1A**	EIFGLYGQTTGK	ARF	1.30	1.05	**1.17**	1.11	0.87	**0.99**	1.25	1.39	**1.32**	0.95	0.96	**0.95**
P49703	**ARL4D**	ASDNQGVPVLVLANK	ARF	0.70		**0.70**	0.39		**0.39**	0.60		**0.60**	0.17	0.21	**0.19**
Q9H7X7	**IFT22**	ILFVGPCESGK	RAB	2.44	1.90	**2.17**	2.13	1.25	**1.69**	2.02	1.58	**1.80**	1.19	1.00	**1.10**
Q9BW83	**IFT27**	CILAGDPAVGK	RAB	0.90	0.82	**0.86**	0.79	0.82	**0.81**	0.90	0.82	**0.86**	0.65	0.75	**0.70**
P61026	**RAB10**	FHTITTSYYR	RAB	0.90	0.84	**0.87**	0.90	0.96	**0.93**	1.10	1.35	**1.22**	0.90	0.99	**0.95**
Q6IQ22	**RAB12**	LQVIIIGSR	RAB	1.62	1.42	**1.52**	1.90	4.61	**3.25**	2.35	1.94	**2.14**	1.69	1.21	**1.45**
P51153	**RAB13**	LQVWDTAGQER	RAB	0.55	0.42	**0.48**	0.56	0.40	**0.48**	0.57	0.58	**0.57**	0.58	0.49	**0.54**
P61106	**RAB14**	TGENVEDAFLEAAK	RAB	1.89	1.05	**1.47**	1.52	1.36	**1.44**	2.31	1.79	**2.05**	1.30	1.26	**1.28**
P59190	**RAB15**	IQIWDTAGQER	RAB	0.95	0.70	**0.83**	0.88	0.59	**0.74**	1.12	1.21	**1.17**	0.85	0.87	**0.86**
Q9NP72	**RAB18**	TLTPSYYR	RAB	1.18	0.92	**1.05**	0.93	0.90	**0.91**	1.12	1.11	**1.11**	1.10	1.19	**1.14**
A4D1S5	**RAB19**	IILIGDSNVGK	RAB	0.70	0.49	**0.60**	0.56	0.39	**0.48**	0.77	0.76	**0.76**	0.38	0.45	**0.41**
Q9UL25	**RAB21**	VVLLGEGCVGK	RAB	2.03	1.29	**1.66**	1.64	1.29	**1.47**	2.06	1.72	**1.89**	1.45	1.40	**1.42**
Q9ULC3	**RAB23**	TIGVDFLER	RAB	1.28	1.01	**1.14**	0.95	0.83	**0.89**	1.11	1.46	**1.28**	0.91	0.88	**0.89**
Q969Q5	**RAB24**	AQLFETSSK	RAB	1.14	0.85	**0.99**	1.20	1.00	**1.10**	1.55	1.36	**1.46**	1.38	1.55	**1.46**
P57735	**RAB25**	VVLIGESGVGK	RAB	3.60	2.51	**3.05**	3.32	2.42	**2.87**	4.52	4.23	**4.37**	3.79	4.33	**4.06**
Q9ULW5	**RAB26**	VMLVGDSGVGK	RAB	0.86		**0.86**	0.46		**0.46**	0.49		**0.49**	0.20		**0.20**
P51159	**RAB27A**	SWIPEGVVR	RAB	2.11	1.88	**1.99**	2.13	1.51	**1.82**	2.89	3.81	**3.35**	2.37	2.02	**2.20**
O00194	**RAB27B**	LLALGDSGVGK	RAB	8.57	7.27	**7.92**	6.82	15.65	**11.24**	13.61	16.19	**14.90**	50.28	53.67	**51.98**
P51157	**RAB28**	VAAEILGIK	RAB	0.96	0.81	**0.88**	0.85	1.03	**0.94**	0.90	0.86	**0.88**	0.80	0.67	**0.74**
P61019	**RAB2A**	MITIDGK	RAB	1.34	1.58	**1.46**	1.55	2.31	**1.93**	2.06	2.11	**2.09**	2.11	2.27	**2.19**
Q8WUD1	**RAB2B**	IGPQQSISTSVGPSASQR	RAB	0.83	0.84	**0.84**	1.22		**1.22**	1.43	2.03	**1.73**	0.89	1.26	**1.08**
Q15771	**RAB30**	IVLIGNAGVGK	RAB	0.64	0.43	**0.54**	0.66	0.45	**0.55**	0.96	0.83	**0.90**	0.78	0.77	**0.77**
Q13637	**RAB32**	VLVIGELGVGK	RAB	13.30	5.79	**9.54**	14.73	10.87	**12.80**	28.27	24.20	**26.23**	19.20	17.50	**18.35**
Q14088	**RAB33A**	VQVWDTAGQER	RAB	0.37	0.18	**0.28**	0.14		**0.14**	0.19		**0.19**	0.32	0.43	**0.37**
Q9H082	**RAB33B**	SAIQVPTDLAQK	RAB	1.17	0.75	**0.96**	0.88	0.70	**0.79**	1.20	1.24	**1.22**	1.07	1.06	**1.07**
Q15286	**RAB35**	LLIIGDSGVGK	RAB	1.04	0.66	**0.85**	1.04	0.65	**0.85**	1.33	1.25	**1.29**	1.28	1.09	**1.18**
Q96AX2	**RAB37**	VMLLGDTGVGK	RAB		0.51	**0.51**					0.41	**0.41**		0.52	**0.52**
P20336	**RAB3A**	TYSWDNAQVLLVGNK	RAB	0.41	0.79	**0.60**	0.28		**0.28**	0.55	0.44	**0.50**	0.57	0.42	**0.50**
O95716	**RAB3D**	LLLIGNSSVGK	RAB	0.63	0.45	**0.54**	0.60	0.43	**0.51**	0.79	0.76	**0.77**	0.56	0.58	**0.57**
Q5JT25-2	**RAB41**	LLFLGEQSGK	RAB	1.38	0.48	**0.93**	3.35		**3.35**	7.03	0.47	**3.75**	1.41	0.55	**0.98**
Q8N4Z0	**RAB42**	VIFLLVGHK	RAB	0.36	0.24	**0.30**	0.24	0.41	**0.32**	0.33	0.18	**0.25**	0.15	0.20	**0.17**

UniProt Accession	Protein	Peptide	Family	Day-3			Day-5			Day-6			Day-8		
				Replicate #1	Replicate #2	Average	Replicate #1	Replicate #2	Average	Replicate #1	Replicate #2	Average	Replicate #1	Replicate #2	Average
P20338	**RAB4A**	IESGELDPER	RAB	1.04	0.95	**0.99**	1.19	1.23	**1.21**	1.17	1.25	**1.21**	1.43	1.17	**1.30**
P61018	**RAB4B**	FLVIGSAGTGK	RAB	0.87	0.73	**0.80**	0.81	0.83	**0.82**	0.88	0.86	**0.87**	0.88	1.10	**0.99**
P20339	**RAB5A**	QASPNIVIALSGNK	RAB		0.63	**0.63**		0.82	**0.82**		1.57	**1.57**		1.26	**1.26**
P61020	**RAB5B**	QASPSIVIALAGNK	RAB	1.17	0.68	**0.93**	0.97	0.79	**0.88**	2.12	1.92	**2.02**	2.97	2.54	**2.75**
P51148	**RAB5C**	QASPNIVIALAGNK	RAB	1.07	0.64	**0.86**	0.82	0.78	**0.80**	1.65	1.35	**1.50**	2.03	2.07	**2.05**
P20340	**RAB6A**	ELNVMFIETSAK	RAB	0.80	0.69	**0.74**	0.92	0.62	**0.77**	0.88	0.03	**0.46**	0.92	1.99	**1.45**
Q9NRW1	**RAB6B**	QITIEEGEQR	RAB	1.37	1.17	**1.27**	1.17	1.12	**1.15**	1.06	1.05	**1.05**	0.81	0.54	**0.68**
P51149	**RAB7A**	FQSLGVAFYR	RAB	0.89	0.43	**0.66**	0.73	0.74	**0.73**	1.83	2.13	**1.98**	2.34	2.01	**2.18**
O14966	**RAB7L**	VLVVGDAAVGK	RAB	0.73	0.71	**0.72**	0.65	0.64	**0.64**	0.57	0.53	**0.55**	0.52	0.60	**0.56**
P61006	**RAB8A**	LALDYGIK	RAB	1.02	0.92	**0.97**	1.03	0.79	**0.91**	0.94	0.88	**0.91**	0.99	1.01	**1.00**
P51151	**RAB9A**	DATNVAAAFEEAVR	RAB	0.75	0.54	**0.64**	0.68	(	**0.68**	1.16	1.47	**1.31**	1.89	1.98	**1.94**
Q9NP90	**RAB9B**	QSFENLGNWQK	RAB	2.46		**2.46**	3.78		**3.78**	4.48	9.92	**7.20**	8.66	17.95	**13.30**
Q9UBK7	**RABL2A**	IICLGDSAVGK	RAB	1.35	1.18	**1.26**	1.51	1.32	**1.41**	1.61	1.40	**1.50**	2.19	1.37	**1.78**
Q5HYI8	**RABL3**	YLAAGSSNAVK	RAB	1.10	1.01	**1.06**	1.14	1.09	**1.12**	1.22	1.04	**1.13**	1.61	0.97	**1.29**
Q3YEC7	**RABL6**	VEVWDVVDK	RAB	0.75	0.73	**0.74**	0.80	0.70	**0.75**	0.67	1.75	**1.21**	0.98	1.31	**1.14**
Q15907	**RB11B**	NILTEIYR	RAB	0.96	0.66	**0.81**	0.88	0.71	**0.80**	1.07	1.88	**1.48**	1.25	1.73	**1.49**
P62826	**RAN**	FNVWDTAGQEK	RAN	0.97	0.76	**0.86**	0.81	0.64	**0.72**	0.98	1.20	**1.09**	0.91	0.98	**0.95**
O95057	**DIRA1**	EAQAVAQEWK	RAS	0.62	0.54	**0.58**	0.55	0.48	**0.51**	0.40	0.39	**0.39**	0.32	0.15	**0.24**
Q96HU8	**DIRA2**	VAVFGAGGVGK	RAS	1.10	0.90	**1.00**	1.75	1.94	**1.84**	1.25	1.39	**1.32**	1.62	2.71	**2.17**
P01112	**HRAS**	SFEDIHQYR	RAS	1.38	1.28	**1.33**	1.77	1.11	**1.44**	1.79	1.78	**1.78**	2.03	2.28	**2.16**
Q9NYS0	**KBRS1**	TLIEPFTLLASK	RAS	1.19		**1.19**	0.69		**0.69**	3.43		**3.43**	3.48		**3.48**
Q9NYR9	**KBRS2**	VVVCGQASVGK	RAS	1.23	0.86	**1.04**	1.00	1.00	**1.00**	1.64	1.47	**1.55**	1.58	1.12	**1.35**
P01116-2	**KRAS**	QGVDDAFYTLVR	RAS	1.10	0.71	**0.91**	0.81	0.94	**0.87**	1.96	1.96	**1.96**	1.05	1.47	**1.26**
O14807	**MRAS**	LVVVGDGGVGK	RAS	0.87	1.17	**1.02**	1.18	1.56	**1.37**	0.97	1.12	**1.04**	1.23	1.23	**1.23**
P01111	**NRAS**	SFADINLYR	RAS	1.28	0.83	**1.05**	1.21	1.20	**1.21**	1.94	2.63	**2.28**	2.10	2.36	**2.23**
P55042	**RAD**	SMPVDER	RAS	1.17		**1.17**	1.00		**1.00**	2.12		**2.12**	1.31		**1.31**
P11234	**RALB**	AEEWGVQYVETSAK	RAS	1.86	1.43	**1.65**	1.66	1.35	**1.51**	2.51	2.49	**2.50**	2.54	2.87	**2.70**
P62834	**RAP1A**	QWCNCAFLESSAK	RAS	1.25	0.85	**1.05**	1.18	0.98	**1.08**	2.02	1.58	**1.80**	1.91	1.80	**1.85**
P10114	**RAP2A**	VPVILVGNK	RAS	0.71	0.86	**0.79**	1.05	1.11	**1.08**	0.98	1.25	**1.11**	1.09	1.30	**1.19**
P01116	**RASK**	SFEDIHHYR	RAS	1.17	1.09	**1.13**	1.17	1.09	**1.13**	1.35	1.22	**1.29**	1.40	1.07	**1.23**
Q8IYK8	**REM2**	VMLVGESGVGK	RAS	0.47	1.05	**0.76**	0.59	1.50	**1.04**	0.75	2.09	**1.42**	0.65	0.84	**0.75**
Q15382	**RHEB**	ALAESWNAAFLESSAK	RAS	0.91	0.68	**0.80**	0.65	0.47	**0.56**	1.39	1.27	**1.33**	1.08	1.43	**1.25**
P62070	**RRAS2**	QVTQEEGQQLAR	RAS	1.29	1.26	**1.27**	1.28	1.28	**1.28**	1.06	0.91	**0.99**	1.06	1.01	**1.04**
Q9BPW5	**RSLBB**	TSLIPRPK	RAS	0.68		**0.68**	0.70		**0.70**	0.56		**0.56**	0.79		**0.79**
P55040	**GEM**	VVLIGEQGVGK	RGK	0.84	0.98	**0.91**	1.00	1.17	**1.09**	1.04	1.29	**1.16**	1.66	1.42	**1.54**
P60953	**CDC42**	TCLLISYTTNK	RHO	1.09	0.87	**0.98**	1.39	1.02	**1.20**	2.17	2.62	**2.40**	3.20	3.07	**3.13**
P15153	**RAC2**	AVLCPQPTR	RHO	1.03	1.02	**1.02**	1.07	1.16	**1.11**	0.90	1.01	**0.95**	0.80	0.58	**0.69**
P60763	**RAC3**	LAPITYPQGLAMAR	RHO	0.75	0.53	**0.64**	0.53	0.69	**0.61**	1.61	1.24	**1.43**	1.39	1.04	**1.21**
P61586	**RHOA**	IGAFGYMECSAK	RHO	0.89	0.72	**0.81**	0.76	0.96	**0.86**	1.35	1.07	**1.21**	1.10	0.95	**1.03**
P62745	**RHOB**	IQAYDYLECSAK	RHO	3.34	2.36	**2.85**	3.34	3.20	**3.27**	7.08	7.22	**7.15**	9.93	10.33	**10.13**
P08134	**RHOC**	ISAFGYLECSAK	RHO	2.13	1.24	**1.69**	1.86	1.79	**1.82**	4.40	3.84	**4.12**	7.79	8.86	**8.33**
Q9HBH0	**RHOF**	AALYLECSAK	RHO	1.22	0.82	**1.02**	1.02	0.91	**0.97**	0.93	1.05	**0.99**	0.48	0.40	**0.44**
P84095	**RHOG**	YLECSALQQDGVK	RHO	1.07	0.93	**1.00**	1.02	0.97	**1.00**	1.16	1.21	**1.18**	1.33	1.30	**1.32**
Q15669	**RHOH**	GVQQVFECAVR	RHO	0.25	0.21	**0.23**	0.25	0.36	**0.30**	0.48	0.58	**0.53**	0.25	0.25	**0.25**
P17081	**RHOQ**	TVFDEAIIAILTPK	RHO	1.28	0.94	**1.11**	1.03	0.57	**0.80**	1.24	0.93	**1.08**	1.20	1.15	**1.17**
Q7L0Q8	**RHOU**	APIILVGTQSDLR	RHO	1.32		**1.32**	1.37		**1.37**	1.97		**1.97**	1.16		**1.16**
P52198	**RND2**	IVVVGDAECGK	RHO	0.49	0.43	**0.46**	0.38	0.36	**0.37**	0.44	0.39	**0.41**	0.25	0.10	**0.17**
P61587	**RND3**	IVVVGDSQCGK	RHO	3.44		**3.44**	4.18		**4.18**	3.62	1.89	**2.76**	3.63	2.18	**2.90**

UniProt Accession	Protein	Peptide	Family	Day-10 Replicate #1	Day-10 Replicate #2	Day-10 Average	Day-12 Replicate #1	Day-12 Replicate #2	Day-12 Average	Day-15 Replicate #1	Day-15 Replicate #2	Day-15 Average	Day-20 Replicate #1	Day-20 Replicate #2	Day-20 Average
P20338	**RAB4A**	IESGELDPER	RAB	1.15	0.98	**1.06**	1.20	1.03	**1.12**	1.33	1.10	**1.22**	0.88	1.05	**0.96**
P61018	**RAB4B**	FLVIGSAGTGK	RAB	0.89	0.69	**0.79**	0.75	0.64	**0.70**	1.04	1.07	**1.05**	0.91	0.92	**0.92**
P20339	**RAB5A**	QASPNIVIALSGNK	RAB		0.90	**0.90**		0.80	**0.80**		1.10	**1.10**		0.82	**0.82**
P61020	**RAB5B**	QASPSIVIALAGNK	RAB	2.58	1.50	**2.04**	2.36		**2.36**	3.13	3.63	**3.38**	2.24	1.74	**1.99**
P51148	**RAB5C**	QASPNIVIALAGNK	RAB	1.84	1.12	**1.48**	1.63	1.26	**1.45**	2.00	1.96	**1.98**	1.26	1.41	**1.33**
P20340	**RAB6A**	ELNVMFIETSAK	RAB	1.47	1.25	**1.36**	1.67	1.03	**1.35**	2.21	2.07	**2.14**	1.05	1.16	**1.10**
Q9NRW1	**RAB6B**	QITIEEGEQR	RAB	0.60	0.51	**0.56**	0.73	0.66	**0.69**	0.91	0.80	**0.86**	0.70	0.83	**0.76**
P51149	**RAB7A**	FQSLGVAFYR	RAB	2.15	1.35	**1.75**	1.74	1.56	**1.65**	2.23	1.80	**2.02**	0.96	0.92	**0.94**
O14966	**RAB7L**	VLVVGDAAVGK	RAB	0.63	0.46	**0.54**	1.03	0.78	**0.90**	1.61	1.43	**1.52**	1.61	1.78	**1.69**
P61006	**RAB8A**	LALDYGIK	RAB	1.07	1.02	**1.04**	1.12	1.27	**1.19**	1.34	1.45	**1.40**	1.48	1.34	**1.41**
P51151	**RAB9A**	DATNVAAAFEEAVR	RAB	1.80	1.73	**1.76**	1.62	1.11	**1.36**	2.11	2.23	**2.17**	1.35	1.07	**1.21**
Q9NP90	**RAB9B**	QSFENLGNWQK	RAB	9.23	17.08	**13.15**	8.94		**8.94**	7.79		**7.79**	5.34		**5.34**
Q9UBK7	**RABL2A**	IICLGDSAVGK	RAB	2.20	1.71	**1.95**	1.78	1.60	**1.69**	2.27	2.04	**2.16**	1.68	1.69	**1.69**
Q5HYI8	**RABL3**	YLAAGSSNAVK	RAB	1.33	1.04	**1.18**	1.05	0.96	**1.00**	1.16	1.04	**1.10**	0.79	0.79	**0.79**
Q3YEC7	**RABL6**	VEVWDVVDK	RAB	1.44	1.40	**1.42**	1.06	0.39	**0.73**	1.19	1.33	**1.26**	0.93	0.94	**0.93**
Q15907	**RB11B**	NILTEIYR	RAB	1.22	1.19	**1.21**	1.25	1.53	**1.39**	1.61	2.61	**2.11**	1.39	1.32	**1.35**
P62826	**RAN**	FNVWDTAGQEK	RAN	0.78	0.61	**0.69**	0.67	0.73	**0.70**	0.61	0.72	**0.67**	0.45	0.44	**0.45**
O95057	**DIRA1**	EAQAVAQEWK	RAS	0.19	0.16	**0.17**	0.12		**0.12**	0.18	0.11	**0.14**	0.08	0.20	**0.14**
Q96HU8	**DIRA2**	VAVFGAGGVGK	RAS	1.82	1.71	**1.77**	1.53		**1.53**	1.89	3.45	**2.67**	3.40	3.48	**3.44**
P01112	**HRAS**	SFEDIHQYR	RAS	1.85	1.85	**1.85**	1.70	1.75	**1.73**	2.57	2.62	**2.59**	2.27	2.46	**2.37**
Q9NYS0	**KBRS1**	TLIEPFTLLASK	RAS	1.71		**1.71**	1.25		**1.25**	2.19		**2.19**	0.73		**0.73**
Q9NYR9	**KBRS2**	VVVCGQASVGK	RAS	1.50	1.25	**1.38**	0.96	1.05	**1.00**	1.00	1.05	**1.02**	0.52	0.60	**0.56**
P01116-2	**KRAS**	QGVDDAFYTLVR	RAS	1.26	1.17	**1.22**	2.18	1.31	**1.75**	1.35	1.15	**1.25**	0.45	0.57	**0.51**
O14807	**MRAS**	LVVVGDGGVGK	RAS	1.23	1.45	**1.34**	1.04	1.34	**1.19**	1.59	1.89	**1.74**	1.63	2.53	**2.08**
P01111	**NRAS**	SFADINLYR	RAS	1.75	1.42	**1.59**	1.45	1.49	**1.47**	1.73	1.91	**1.82**	1.03	1.02	**1.03**
P55042	**RAD**	SMPVDER	RAS	1.13		**1.13**	1.31		**1.31**	1.36		**1.36**	1.10		**1.10**
P11234	**RALB**	AEEWGVQYVETSAK	RAS	2.02	1.19	**1.60**	1.81	2.21	**2.01**	2.09	1.64	**1.87**	0.95	0.76	**0.85**
P62834	**RAP1A**	QWCNCAFLESSAK	RAS	2.05	1.64	**1.85**	2.26	11.82	**7.04**	2.49	2.61	**2.55**	1.88	1.46	**1.67**
P10114	**RAP2A**	VPVILVGNK	RAS	1.08	1.01	**1.04**	1.30	1.80	**1.55**	1.53	1.37	**1.45**	1.60	1.42	**1.51**
P01116	**RASK**	SFEDIHHYR	RAS	1.28	1.71	**1.50**	0.90	3.90	**2.40**	1.30	2.13	**1.71**	0.88	0.76	**0.82**
Q8IYK8	**REM2**	VMLVGESGVGK	RAS	0.59	1.59	**1.09**	0.92		**0.92**	1.14	1.97	**1.55**	1.26	1.39	**1.32**
Q15382	**RHEB**	ALAESWNAAFLESSAK	RAS	1.46	2.38	**1.92**	2.89	182.53	**92.71**	2.70	2.44	**2.57**			
P62070	**RRAS2**	QVTQEEGQQLAR	RAS	0.57	0.46	**0.51**	0.62	0.50	**0.56**	0.82	0.60	**0.71**	0.40	0.49	**0.44**
Q9BPW5	**RSLBB**	TSLIPRPK	RAS	0.39		**0.39**	0.36		**0.36**	0.48		**0.48**	0.46		**0.46**
P55040	**GEM**	VVLIGEQGVGK	RGK	1.48	1.70	**1.59**	1.06	0.64	**0.85**	0.94	0.90	**0.92**	0.74	1.01	**0.87**
P60953	**CDC42**	TCLLISYTTNK	RHO	3.08	2.15	**2.62**	2.67	8.44	**5.56**	3.48	2.98	**3.23**	1.73	1.58	**1.65**
P15153	**RAC2**	AVLCPQPTR	RHO	0.55	0.49	**0.52**	0.49		**0.49**	0.56	0.94	**0.75**	1.07	1.37	**1.22**
P60763	**RAC3**	LAPITYPQGLAMAR	RHO	1.33	0.77	**1.05**	1.62	1.83	**1.72**	2.25	1.71	**1.98**	0.86	1.09	**0.97**
P61586	**RHOA**	IGAFGYMECSAK	RHO	1.07	0.93	**1.00**	1.15	1.08	**1.12**	1.43	1.17	**1.30**	0.82	0.90	**0.86**
P62745	**RHOB**	IQAYDYLECSAK	RHO	8.91	6.19	**7.55**	5.88	5.95	**5.92**	6.49	6.04	**6.27**	3.52	3.75	**3.64**
P08134	**RHOC**	ISAFGYLECSAK	RHO	9.90	5.00	**7.45**	10.38	8.05	**9.22**	13.98	11.89	**12.93**	6.46	6.40	**6.43**
Q9HBH0	**RHOF**	AALYLECSAK	RHO	0.35	0.28	**0.31**	0.18		**0.18**	0.21	0.19	**0.20**	0.21	0.26	**0.24**
P84095	**RHOG**	YLECSALQQDGVK	RHO	1.07	0.86	**0.96**	0.98	0.77	**0.87**	1.30	1.23	**1.27**	1.25	1.17	**1.21**
Q15669	**RHOH**	GVQQVFECAVR	RHO	0.14	0.39	**0.27**	0.05		**0.05**	0.09		**0.09**	0.05		**0.05**
P17081	**RHOQ**	TVFDEAIIAILTPK	RHO	1.03	0.44	**0.74**	1.59	1.58	**1.58**	1.74	0.45	**1.10**	1.38	0.66	**1.02**
Q7L0Q8	**RHOU**	APIILVGTQSDLR	RHO	1.59		**1.59**	1.17		**1.17**	0.95		**0.95**	1.52		**1.52**
P52198	**RND2**	IVVVGDAECGK	RHO	0.19	0.14	**0.16**	0.15	0.13	**0.14**	0.15	0.11	**0.13**	0.07	0.05	**0.06**
P61587	**RND3**	IVVVGDSQCGK	RHO	4.01	1.89	**2.95**	3.13	3.92	**3.53**	4.44	3.20	**3.82**	3.60	3.59	**3.59**

Table S3.4. A list of ratios of all small GTPases in C3H10T1/2 cells with osteogenic differentiation over those without, as measured by scheduled LC-MRM analysis.

UniProt Accession	Proteins	Peptide	Family	Biological #1			Biological #2			Biological #3			Average ratio
				Ctrl	Ost	Ratio	Ctrl	Ost	Ratio	Ctrl	Ost	Ratio	
P18085	ARF4	LGLQSLR	ARF	12.106	17.772	**1.468**	15.201	16.704	**1.099**	16.858	20.307	**1.205**	1.257
P84085	ARF5	LGLQHLR	ARF	1.621	2.927	**1.805**	2.854	2.670	**0.935**	3.532	3.646	**1.032**	1.258
P62330	ARF6	FNVWDVGGQDK	ARF	12.145	12.580	**1.036**	24.056	17.282	**0.718**	23.061	17.840	**0.774**	0.843
Q13795	ARFRP	DCLTQACSALTGK	ARF	0.796	0.659	**0.828**	1.091	1.175	**1.077**	0.755	0.606	**0.802**	0.902
P40616	ARL1	ILILGLDGAGK	ARF	1.770	3.266	**1.845**	2.897	2.563	**0.885**	2.820	3.789	**1.344**	1.358
Q9NXU5	ARL15	ELGGADNIR	ARF				0.462	0.144	**0.311**	0.408	0.126	**0.308**	0.310
P36404	ARL2	ELQSLLVEER	ARF	3.654	6.374	**1.744**	5.206	6.457	**1.240**	4.917	7.205	**1.465**	1.483
P36405	ARL3	ILLLGLDNAGK	ARF	3.189	3.720	**1.167**	4.895	4.039	**0.825**	4.442	4.629	**1.042**	1.011
Q9Y689	ARL5A	AGLLIFANK	ARF				0.102	0.089	**0.875**	0.056	0.112	**1.998**	1.437
Q96KC2	ARL5B	AAVLIFANK	ARF	0.047	0.089	**1.899**	0.095	0.091	**0.963**	0.059	0.103	**1.734**	1.532
Q9H0F7	ARL6	IPILFFANK	ARF	0.135	0.216	**1.599**	0.260	0.279	**1.074**	0.209	0.287	**1.373**	1.349
Q96BM9	ARL8A	LWDIGGQPR	ARF	3.798	5.427	**1.429**	6.026	6.535	**1.084**	6.976	7.744	**1.110**	1.208
Q9NVJ2	ARL8B	IWDIGGQPR	ARF	4.215			6.833	6.946	**1.017**	7.387	7.464	**1.010**	1.013
Q9NR31	SAR1A	EIFGLYGQTTGK	ARF	4.385			6.939	9.583	**1.381**	6.235	9.900	**1.588**	1.485
Q9Y6B6	SAR1B	EMFGLYGQTTGK	ARF				0.671	0.315	**0.469**	0.584	0.538	**0.920**	0.695
Q9H7X7	IFT22	ILFVGPCESGK	RAB	0.127	0.195	**1.533**	0.181	0.206	**1.135**	0.178	0.209	**1.178**	1.282
Q9BW83	IFT27	CILAGDPAVGK	RAB	0.104	0.164	**1.576**	0.195	0.152	**0.775**	0.175	0.218	**1.246**	1.199
P61026	RAB10	FHTITTSYYR	RAB	3.518	3.289	**0.935**	5.002	3.214	**0.642**	5.219	3.550	**0.680**	0.752
Q6IQ22	RAB12	LQVIIIGSR	RAB	1.461	2.115	**1.447**	1.740	2.569	**1.476**	3.016	2.186	**0.725**	1.216
P51153	RAB13	LQVWDTAGQER	RAB	0.325	0.462	**1.420**	0.356	0.353	**0.992**	0.343	0.326	**0.950**	1.121
P61106	RAB14	TGENVEDAFLEAAK	RAB	21.257	23.147	**1.089**	39.503	34.828	**0.882**	42.301	36.202	**0.856**	0.942
P59190	RAB15	IQIWDTAGQER	RAB	44.452	53.116	**1.195**	63.241	62.001	**0.980**	64.005	57.927	**0.905**	1.027
Q9NP72	RAB18	TLTPSYYR	RAB	2.255	1.779	**0.789**	2.494	2.055	**0.824**	2.875	2.085	**0.725**	0.780
Q9UL25	RAB21	VVLLGEGCVGK	RAB	1.234			1.890	1.452	**0.768**	1.825	1.531	**0.839**	0.804
Q9ULC3	RAB23	TIGVDFLER	RAB	2.722			4.943	2.802	**0.567**	5.016	2.933	**0.585**	0.576
Q969Q5	RAB24	AQLFETSSK	RAB	0.246	0.328	**1.333**	0.405	0.305	**0.754**	0.370	0.405	**1.095**	1.061
Q9ULW5	RAB26	VMLVGDSGVGK	RAB				1.901	2.048	**1.077**	2.418	2.666	**1.102**	1.090
P51157	RAB28	VAAEILGIK	RAB	0.077	0.148	**1.930**	0.097	0.125	**1.281**	0.104	0.172	**1.649**	1.620
P61019	RAB2A	MITIDGK	RAB		2.338		24.805	15.327	**0.618**	44.612	21.424	**0.480**	0.549
Q15771	RAB30	IVLIGNAGVGK	RAB	0.121	0.097	**0.800**	0.158	0.121	**0.765**	0.166	0.116	**0.696**	0.753
Q13637	RAB32	VLVIGELGVGK	RAB	3.834	9.134	**2.383**	10.046	13.933	**1.387**	8.084	12.573	**1.555**	1.775
Q15286	RAB35	LLIIGDSGVGK	RAB	0.658	0.359	**0.545**	0.779	0.345	**0.443**	0.955	0.349	**0.365**	0.451
P20336	RAB3A	TYSWDNAQVLLVGNK	RAB	1.352	2.444	**1.808**		7.190		5.286	3.273	**0.619**	1.213
O95716	RAB3D	LLLIGNSSVGK	RAB	0.050	0.064	**1.269**	0.071	0.053	**0.745**	0.052	0.087	**1.668**	1.227
Q8N4Z0	RAB42	VIFLLVGHK	RAB	0.882			0.768	1.132	**1.474**	0.055	0.744	**13.481**	7.478
Q86YS6	RAB43	VATELIMR	RAB	0.098	0.173	**1.775**	0.123	0.173	**1.408**	0.139	0.168	**1.206**	1.463
P61018	RAB4B	FLVIGSAGTGK	RAB	0.441	0.483	**1.095**	0.604	0.513	**0.850**	0.687	0.528	**0.769**	0.905
P61020	RAB5B	QASPSIVIALAGNK	RAB	25.744	13.324	**0.518**	13.087	14.983	**1.145**	7.244	5.013	**0.692**	0.785
P51148	RAB5C	QASPNIVIALAGNK	RAB	12.003	14.694	**1.224**	21.772	17.395	**0.799**	20.269	17.867	**0.882**	0.968
Q9NRW1	RAB6B	QITIEEGEQR	RAB	3.303			4.470	2.265	**0.507**	4.555	2.248	**0.493**	0.500
P51149	RAB7A	FQSLGVAFYR	RAB	21.630	35.326	**1.633**	61.863	82.052	**1.326**	51.490	53.556	**1.040**	1.333
O14966	RAB7L	VLVVGDAAVGK	RAB	0.066	0.099	**1.508**	0.117	0.102	**0.875**	0.107	0.130	**1.208**	1.197
P61006	RAB8A	LALDYGIK	RAB	0.435	0.439	**1.009**	0.549	0.371	**0.677**	0.556	0.435	**0.781**	0.822
Q5HYI8	RABL3	YLAAGSSNAVK	RAB	0.173	0.131	**0.755**	0.203	0.112	**0.552**		0.182		0.653
Q3YEC7	RABL6	VEVWDVVDK	RAB	0.019	0.032	**1.659**	0.034	0.034	**1.007**	0.025	0.036	**1.456**	1.374
Q15907	RAB11B	NILTEIYR	RAB	16.625	19.723	**1.186**	25.508	22.790	**0.893**	27.072	23.224	**0.858**	0.979
O00194	RAB27B	LLALGDSGVGK	RAB	0.075	0.064	**0.851**	0.094	0.059	**0.624**	0.098	0.075	**0.766**	0.747
Q9H082	RAB33B	SAIQVPTDLAQK	RAB	0.261	0.425	**1.630**	0.375	0.451	**1.203**	0.372	0.567	**1.523**	1.452
Q12829	RAB40B	EIDEHAPGVPK	RAB	0.749	0.778	**1.039**	1.753	1.386	**0.791**	0.926	1.244	**1.343**	1.058
Q9UBK7	RABL2A	IICLGDSAVGK	RAB	0.138	0.174	**1.266**	0.218	0.185	**0.850**	0.227	0.244	**1.073**	1.063
P62826	RAN	FNVWDTAGQEK	RAN	12.886	21.972	**1.705**	19.962	22.379	**1.121**	20.153	24.136	**1.198**	1.341
Q96HU8	DIRA2	VAVFGAGGVGK	RAS	0.087	0.022	**0.257**	0.113	0.014	**0.126**	0.108	0.032	**0.297**	0.227
Q9NYS0	KBRS1	TLIEPFTLLASK	RAS	0.254	0.361	**1.424**	1.095	0.874	**0.798**	0.518	0.635	**1.226**	1.149
Q9NYR9	KBRS2	VVVCGQASVGK	RAS				0.181	0.415	**2.296**	0.331	0.490	**1.482**	1.889
P11234	RALB	AEEWGVQYVETSAK	RAS	1.717	1.655	**0.964**	4.626	3.823	**0.826**	3.435	3.788	**1.103**	0.964
P62834	RAP1A	QWCNCAFLESSAK	RAS	2.530	1.292	**0.511**	2.902	1.921	**0.662**	4.023	2.160	**0.537**	0.570
P61224	RAP1B	QWNNCAFLESSAK	RAS	13.048	8.143	**0.624**	14.442	10.061	**0.697**	22.277	7.186	**0.323**	0.548
P10114	RAP2A	VPVILVGNK	RAS	0.114	0.041	**0.357**	0.153	0.047	**0.308**	0.156	0.040	**0.260**	0.308
Q9Y3L5	RAP2C	VPLILVGNK	RAS	0.164	0.057	**0.349**	0.191	0.066	**0.346**	0.227	0.061	**0.271**	0.322
P01112	HRAS	SFEDIHQYR	RAS	0.676	0.405	**0.600**	0.985	0.407	**0.413**	1.090	0.436	**0.400**	0.471
P01116-2	KRAS	QGVDDAFYTLVR	RAS	1.344	0.819	**0.610**	0.542	0.985	**1.817**	0.700	0.690	**0.985**	1.137
P01116	KRAS	SFEDIHHYR	RAS	0.867	0.238	**0.275**		0.411		1.410	0.160	**0.113**	0.194
O14807	MRAS	LVVVGDGGVGK	RAS	0.674	0.338	**0.502**	0.714	0.409	**0.573**	0.865	0.459	**0.531**	0.535
P01111	NRAS	SFADINLYR	RAS	2.731	1.726	**0.632**	4.421	2.431	**0.550**	3.984	1.688	**0.424**	0.535
P10301	RRAS	LFTQILR	RAS	3.756	1.322	**0.352**	4.955	2.724	**0.550**	4.447	1.998	**0.449**	0.450
P62070	RRAS2	QVTQEEGQQLAR	RAS	5.412	3.492	**0.645**	11.444	2.881	**0.252**	6.906	3.570	**0.517**	0.471
Q9BPW5	RSLBB	TSLIPRPK	RAS	0.053	0.084	**1.579**	0.182			0.140	0.131	**0.942**	1.260
P63000	RAC1	HHCPNTPIILVGTK	RHO	3.928	2.463	**0.627**	6.539	3.536	**0.541**	4.592	4.102	**0.893**	0.687
P61586	RHOA	IGAFGYMECSAK	RHO	1.185	2.098	**1.771**	6.629	3.120	**0.471**	7.524	2.559	**0.340**	0.861
P62745	RHOB	IQAYDYLECSAK	RHO	0.706	0.629	**0.891**	1.280	0.898	**0.702**	1.123	0.641	**0.570**	0.721
P08134	RHOC	ISAFGYLECSAK	RHO	25.405	37.434	**1.473**	70.235	70.023	**0.997**	47.303	49.681	**1.050**	1.174
O00212	RHOD	SVGAVAYLECSAR	RHO	0.812	0.741	**0.913**	1.397	1.380	**0.988**	1.291	0.987	**0.765**	0.888
P84095	RHOG	YLECSALQQDGVK	RHO	2.097	1.559	**0.744**	3.324	1.492	**0.449**	3.004	1.937	**0.645**	0.612
P61587	RND3	IVVVGDSQCGK	RHO	0.190	0.175	**0.921**	0.334	0.276	**0.825**	0.284	0.244	**0.859**	0.868

4 Chapter 4. Proteome-Wide Interrogation of Small GTPases Regulated by N^6-methyladenosine Modulators

4.1 Introduction

Small GTPases of the Ras superfamily are molecular switches that undergo conformational changes between the active guanosine triphosphate (GTP)-associated state and the inactive guanosine diphosphate (GDP)-occupied state.[1] Members of this superfamily interact with effector proteins and modulate diverse essential cellular processes including proliferation, vesicular transport, and cytoskeleton organization.[2,3] Mutations and aberrant expressions of small GTPases (e.g. KRAS) promote the progression of various types of cancer.[4,5] Indeed, clinical studies have documented that numerous small GTPases, e.g. those in the RAS and RHO subfamilies, are differentially expressed in tumor vs. normal tissues, and metastatic vs. primary tumor cells.[6] In this vein, targeting small GTPases has drawn substantial attention from academia and pharmaceutical industry for the development of cancer chemotherapeutic agents.[7] Despite intensive efforts since the initial discovery of the RAS oncogenes over three decades ago, no effective inhibitors have been approved for cancer treatment.[8-10] Therefore, alternative approaches to modulate the expression of dysregulated small GTPases may offer new venues for cancer therapy.

A growing number of studies in recent years have shown that RNA metabolism can be dynamically regulated by post-transcriptional modifications, where N^6-methyladenosine (m^6A) is the most prevalent internal modification in mRNA.[11] Previous studies have established dynamic regulation of m^6A in mRNA by the METTL3-METTL14 m^6A

methyltransferase complex and m^{6}A demethylases (i.e. FTO and ALKBH5).[12-14] When bound by its reader protein YTHDF2, the m^{6}A in mRNAs could lead to their destabilization;[15,16] the binding of m^{6}A in mRNA by other reader proteins (e.g. YTHDF1 and YTHDF3), however, enhances the stability and/or translation efficiency of the mRNA.[17-19] In addition, m^{6}A within the 5′-untranslated region (UTR) of mRNA promotes cap-independent translation in mammalian cells.[20] Interestingly, dysregulations of epitranscriptomic modulators were found to drive the abnormal expression of oncogenes at the protein level and thus promote the growth, survival, and invasion of cancer cells.[21-23] We hypothesized that the expression of small GTPase proteins may be subjected to regulation by m^{6}A modulators.

In the present study, we test the above hypothesis by analyzing comprehensively the alterations in small GTPase proteome in HEK293T cells upon CRISPR-Cas9-mediated ablation of *METTL3*, *ALKBH5*, *FTO*, *YTHDF1*, *YTHDF2*, and *YTHDF3* genes.[24] To achieve quantifications of small GTPase proteins in high sensitivity, reproducibility and throughput, we employed a previously developed scheduled multiple-reaction monitoring (MRM)-based quantitative proteomic approach, together with the use of synthetic stable isotope-labeled peptides as internal standards.[25-27] We observed altered protein expression of several small GTPases, including the known tumor suppressor RHOB and metastasis driver RHOC,[5] in cells depleted of METTL3 or ALKBH5.

4.2 Materials and Methods

4.2.1 Cell Culture and Gene Depletion

HEK293T human embryonic kidney epithelial cells (ATCC) were cultured in Dulbecco's modified Eagle's medium (DMEM, Thermo Fisher) supplemented with 10% fetal bovine serum (FBS, Thermo Fisher) and 1% penicillin-streptomycin solution (PS, GE Healthcare). The cells were maintained at 37°C in a humidified chamber supplemented with 5% CO_2. All knockout cells in the HEK293T background were generated by using the CRISPR-Cas9 genome editing method.[28] In brief, plasmids (Addgene #62988) carrying required enzymes and selected guide sequence (Table S1) were transiently transfected into HEK293T cells and the single clones were selected. Successful knockouts of target genes were confirmed by Sanger sequencing and Western blot analyses (Figure S1).

4.2.2 Cell Lysis and Proteomic Sample Preparation

To extract total proteins, approximately 8×10^6 cells were incubated on ice with 100 µL of pre-chilled CelLytic M cell lysis reagent (Sigma) containing 1% protease inhibitor cocktail (Sigma) on-ice for 30 min. The cell lysate was centrifuged at 16000 *g* for 30 min at 4°C. Supernatants containing total proteins were transferred to a chilled vial. Protein concentration was measured by using Quick Start Bradford Protein Assay (Bio-Rad). Approximately 50 µg of total proteins was mixed with 4× Laemmli SDS loading buffer and the mixtures were boiled for 10 min. Total proteins were resolved by 10% SDS-PAGE gel and then stained with Coomassie Brilliant Blue R-250.

In-gel protein digestion was conducted as described previously.[29] In brief, gel bands corresponding to a molecular weight range of 15-37 kDa were cut into 1 mm^3 cubes, and

sequentially destained with 25% and 50% CH_3CN in 50 mM ammonium bicarbonate (pH 7.8). Reduction and alkylation were performed by incubating the gel pieces in 10 mM dithiothreitol (DTT) at 37°C for 1 h and 55 mM iodoacetamide at room temperature in dark for 20 min, respectively. The proteins were digested with trypsin at 37°C for 18 h with an enzyme/protein ratio of 1:100. Peptides were extracted from the gel pieces by sonicating in a solution containing CH_3CN/H_2O/acetic acid (45/45/5, v/v) and concentrated by Speed-Vac. Prior to LC-MRM analysis, the tryptic peptides were desalted using C18 ZipTip (Agilent), concentrated by Speed-Vac, and reconstituted in 30 µL of 0.1% formic acid. Approximately 8% of the digestion mixture was spiked-in with a crude pool of synthetic small GTPase peptides (New England Peptide, Inc.) (~4 fmol each) with a C-terminal [$^{15}N_2$, $^{13}C_6$]-labeled lysine (+8.0 Da) or [$^{15}N_4$, $^{13}C_6$]-labeled arginine (+10 Da).[25]

4.2.3 Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) Analysis and Data Processing

The MRM-based LC-MS/MS experiments were performed on a TSQ Altis triple-quadrupole mass spectrometer (Thermo Fisher) coupled with an UltiMate 3000 UPLC (Thermo Fisher) and a Flex nanoelectrospray ion source (Thermo Fisher). The spiked-in sample was loaded onto a 4-cm trapping column (150 µm i.d.) packed in-house with C18 resin (5 µm in particle size and 120 Å in pore size, Dr. Maisch GmbH HPLC) with buffer A (0.1% formic acid in water) at a flow rate of 3 µL/min. The peptides were eluted from the trapping column and separated on an analytical column (75 µm i.d., ~25 cm in length) packed in-house with C18 resin (3 µm in particle size and 120 Å in pore size, Dr. Maisch GmbH HPLC), using a gradient of 12-40% buffer B (0.1% formic acid in 80% CH_3CN)

and a flow rate of 300 nL/min. The peptides were ionized with a spray voltage of 2200 V and an ion transport tube temperature of 325°C. A resolution of 0.7 F.W.H.M. was set for both Q1 and Q3. Fragmentation of precursor ions in Q2 was conducted with 1.5 mTorr argon and the collision energy was derived from default settings in Skyline (version 19.1.0.193).[30]

The mass spectrometer was scheduled to monitor the precursor to product ion transitions of 144 unique peptides of human small GTPases with a cycle time of 3 sec in a 4-min retention time (RT) window. The retention time of the monitored peptides was predicted from its normalized retention time (iRT) and the iRT-RT calibration curve determined from 10 tryptic peptides of BSA.[25,31] The three most abundant y ions found in MS/MS acquired from shotgun proteomic analyses were selected to quantify each peptide.

To enable robust peak detection, all target peptides were manually inspected and filtered with dot-product (dotp) > 0.7.[32] To calculate the relative quantity of small GTPases, the sum of peak areas in the extracted-ion chromatograms (XICs) for the three monitored transitions were normalized against those of corresponding spiked-in heavy isotope-labeled peptide. The derived ratio of the knockout cells was further normalized against that of parental HEK293T cells. A total of three and four biological replicates were conducted for *METTL3*$^{-/-}$ and other knockout cells, respectively.

4.2.4 Real-Time Quantitative PCR (RT-qPCR) and mRNA Stability Assay

RNA was extracted from approximately 1×10^6 cells with Total RNA Kit I (Omega) and purified with HiBind RNA mini columns (VWR). Total RNA (1 µg) was subjected to reverse transcription and 66 ng of the resulting cDNA was mixed with Luna qPCR Master

Mix (New England Biolabs) in a 96-well optical reaction plate (Bio-Rad). RT-qPCR was conducted on a CFX real-time PCR detection system (Bio-Rad). For mRNA stability assay, cells were treated with 5 μg/mL of actinomycin D in a reverse chronological order prior to harvesting cells for RNA extraction. The primer sequences used for RT-qPCR were: *GAPDH* forward, GGTCGGAGTCAACGGATT; *GAPDH* reverse, ATCGCCCCACTTGATTTTG; *RHOB* forward, CGGACTCGCTGGAGAACA; *RHOB* revere, GAGGTAGTCGTAGGCTTGGAT; *RHOC* forward, CAGTCTGAGCCTCCGGCAC; *RHOC* reverse, GAGGAGGCAGGTCTTCCCAC.

4.2.5 Hierarchical Clustering

The hierarchical clustering was performed using the default setting of the ClustVis online tool with minor changes.[33] Briefly, the Euclidean method was used for distance measurement, and the tightest is clustered first in the tree ordering.

4.3 Results

4.3.1 MRM-Based Targeted Proteomic Profiling of Differential Expression of Small GTPase Proteins in Cells Depleted of m^{6}A Modulators

Recent studies revealed that dynamic and reversible formations of m^{6}A in mRNA enable cells to adapt rapidly to environmental stresses.[34-36] Since small GTPases serve as upstream molecular switches for numerous signal transduction cascades in response to extracellular stimuli, we reason that the expression of small GTPase proteins may be dynamically regulated by m^{6}A modulators. To test this hypothesis, we systemically investigated how the protein expression levels of small GTPases are modulated by epitranscriptomic regulators. We utilized a previously established high-throughput MRM-

based targeted proteomic workflow to quantify, at the proteome-wide scale, the changes in expression of small GTPase proteins in cells where the m^6A writer (METTL3), eraser (ALKBH5 and FTO), or reader (YTHDF1, YTHDF2, and YTHDF3) proteins were individually knocked out by CRISPR-Cas9 (Figure S4.1, Table S4.1).[24,25]

The MRM library contained a unique peptide for each of the 144 small GTPases, which represent approximately 85% of the human small GTPase proteome.[27,37] To fulfill sensitive and reproducible quantification, we selected three most abundant fragment ions for each peptide based on MS/MS obtained from shotgun proteomic analyses. We realized that the potential differences in efficiencies of tryptic digestion and peptide recovery from gels may affect the quantification accuracies for small GTPases. Thus, we always analyzed the lysates of the knockout cells in parallel with those of parental HEK293T cells in each batch of sample preparation (Figure S4.2). A mixture of synthetic stable isotope-labeled peptides with identical amino acid sequences to our target analytes was employed as internal standards for robust quantitation and was spiked into the samples prior to LC-MS/MS analysis (Figure 4.1A). Notably, the heavy peptides and the corresponding unlabeled sample peptides exhibit nearly identical retention time and ionization efficiency.[38] MRM analyses facilitated quantification of > 80 small GTPases for each knocked-out background (Figure S4.3A-F, Table S4.2), and 78 small GTPases were commonly quantified in all genetic backgrounds (Figure 4.1B). It is of note that some peptides were only detected in the heavy form, suggesting that these small GTPases are not expressed or expressed at very low levels in HEK293T and the isogenic knockout cells. Additionally, the failure in detecting in some small GTPases may also arise from post-

translational modifications that increase pronouncedly the molecular weight of the small GTPase protein (i.e. to above 37 kDa) and/or alter the mass of the small GTPase peptides selected for MRM analysis.

Our results showed that depletion of METTL3 and ALKBH5 altered the expression of small GTPase proteins in HEK293T cells (Figure 4.1C). By imposing a cutoff ratio of 1.5-fold, we found that the expression levels of 13 and 20 small GTPase proteins were substantially altered in HEK293T cells upon genetic ablations of *METTL3* and *ALKBH5*, respectively (Figure 4.2A). By using the same threshold, only a few small GTPases were, however, up- or down-regulated at the protein level in cells depleted of FTO or any of the three m^6A reader proteins (Figure S4.3G-J). Our observation about the three m^6A reader proteins is in keeping with a recent study showing that YTHDF1, YTHDF2, and YTHDF3 assume redundant functions, and their abilities in regulating mRNA stability becomes evident only when all three paralogs are simultaneously depleted.[39]

The results from the MRM-based small GTPase quantification suggest that METTL3 and ALKBH5 may be important regulators for the expression of some small GTPases at the protein level (Figure 4.1C). In particular, 80 small GTPases were commonly quantified in *METTL3*$^{-/-}$ and *ALKBH5*$^{-/-}$ cells (Figure S4.4). A hierarchical clustering analysis for the commonly quantified small GTPases revealed the regulatory profiles of METTL3 and ALKBH5 in the expression of small GTPase proteins (Figure 4.2B).[33] Notably, the protein expressions of several small GTPases were significantly altered in cells depleted of METTL3 and/or ALKBH5 (Figure 4.2C), including the well-established regulators for cancer metastasis, e.g., RHOB and RHOC, along with the less

well-characterized RAB40A and RAC2.[5,40] Together, our results revealed the involvements of METTL3 and ALKBH5 in regulating the expression of small GTPase proteins.

4.3.2 Western Blot for the Validation of Differential Expression of Small GTPase Proteins in Cells Depleted of m^{6}A Epitranscriptomic Modulators

To further validate the findings made from the scheduled MRM analysis, we employed Western blot analysis to quantify the small GTPase expression at the protein level. The Western blot results confirmed the relative expression levels of RHOB, RHOC and several other selected small GTPase proteins in *METTL3*$^{-/-}$ and *ALKBH5*$^{-/-}$ cells vs. the parental HEK293T cells (Figure 4.3A-B, Figure S4.5). Together, the quantification results obtained from Western blot analysis are highly consistent with the proteomic data with an R^2 value of 0.991, and a slope of 0.934 in linear regression analysis (Figure 4.3C). This result demonstrates the excellent accuracy of the scheduled MRM approach in quantifying the protein expression of small GTPases in cells depleted of epitranscriptomic modulators.

To further explore the involvements of METTL3 and ALKBH5 in regulating small GTPase expression at the protein level, we compared the transcriptome datasets retrieved from the Broad Institute Cancer Cell Line Encyclopedia (CCLE) and the previously published proteomic results acquired from the paired primary/metastatic melanoma cells (WM115/WM266-4 and IGR39/IGR37) and colorectal cancer cells (SW480/SW620).[26,41,42] Interestingly, we observed a greater discrepancy between the mRNA and protein expression for the small GTPases found to be regulated by METTL3 and/or ALKBH5 in the present study than those that are not, for which a normal distribution

of mRNA/protein ratio with the median value being 0.87, 0.97, and 1.17 were observed in the WM115/WM266-4, IGR39/IGR37, and SW480/SW620 paired cell lines, respectively (Figure S4.6). Notably, although the median ratio of the small GTPases regulated by METTL3 and ALKBH5 in the IGR39/IGR37 are close to 1, the ratio distribution is distinct from the non-targets (Figure S4.6B). Interestingly, the expression of the well-characterized driver genes for melanoma metastasis, e.g., ARF6 and RHOC, were down-regulated at the mRNA level but up-regulated at the protein level in the metastatic melanoma cells relative to the paired primary melanoma cells (Table S4.3).[40,43] Together, these results support that the differential expression of small GTPases during the metastatic progression may be regulated by epitranscriptomic modulators.

4.3.3 mRNA Metabolism Regulated by Epitranscriptomic Modulators

Aberrant protein expression of RHO small GTPases, e.g., RHOB and RHOC, drives the migration and invasion of cancer cells.[40,44,45] Our MRM analysis of small GTPase proteome revealed altered expressions of RHOB and RHOC proteins in the *METTL3*$^{-/-}$ and *ALKBH5*$^{-/-}$ cells. To understand how METTL3 and ALKBH5 regulate the protein expression of RHOB and RHOC, we carried out real-time quantitative PCT (RT-qPCR) analysis for *RHOB* and *RHOC* transcripts in *METTL3*$^{-/-}$ and *ALKBH5*$^{-/-}$ cells. The results revealed a similar trend in changes of mRNA and protein expression (Figure 4.4A). Thus, we hypothesized that METTL3 and ALKBH5 may alter the transcriptional efficiencies of *RHOB* and *RHOC* genes and/or the stabilities of the resulting transcripts in cells. To distinguish these two possibilities, we tested, by employing RT-qPCR analysis, the stabilities of *RHOB* and *RHOC* mRNAs by treating cells with a transcription inhibitor,

actinomycin D, prior to harvesting cells for mRNA extraction. The results revealed an elevated stability of *RHOB* transcripts with half-life being increased by approximately 1.5-fold in cells depleted of METTL3 (Figure 4.4B). The mRNA of RHOC, on the other hand, was stable during the monitored time window with a marginal decrease in parental HEK293T cells (Figure S4.7). Together, our results demonstrated that the expression of RHOB and RHOC can be directly or indirectly regulated by epitranscriptomic modulators.

4.4 Discussion

In this study, we employed a facile MRM-based targeted proteomic analysis to study the alterations in small GTPase proteome in cells depleted of m^6A epitranscriptomic modulators. With this high-throughput method, we were able to monitor 144 small GTPases, which represent approximately 85% of the entire small GTPase proteome in human cells, and commonly quantify 78 small GTPases among the cells depleted of m^6A writer, eraser or reader proteins. Moreover, we validated the MRM data by Western blot analysis (Figure 4.3C), and the quantification results obtained from the two methods are highly consistent, underscoring the accuracy of the MRM-based target proteomic method in quantifying small GTPase protein expression. To our knowledge, this is the first comprehensive analysis of the entire small GTPase proteome regulated by m^6A writer, eraser, and reader proteins. Our results revealed the roles of epitranscriptomic modulators in regulating the expression of small GTPases at the protein level.

The expression of small GTPase genes is tightly regulated, and maintaining their expression levels within a proper range is pivotal for cell fitness.[2] Mutations and abnormal expression of small GTPase proteins are among the key drivers for tumor development and

progression.[5,40,46] The RHO small GTPases such as RHOC and RAC2 are overexpressed in cancer cells and are associated with poor patient prognosis. In contrast, RHOB was found to suppress tumorigenesis.[4,5] Despite the importance of targeting small GTPases in anti-cancer therapy and decades of intensive research efforts, no effective inhibitors for oncogenic small GTPases have reached the clinic.[9]

Here, we found that the expression of small GTPase proteins, including several known tumor drivers or suppressors, were significantly altered in the cells depleted of m^6A epitranscriptomic modulators, especially the m^6A methyltransferase METTL3 and demethylase ALKBH5 (Figure 4.1C and 4.2C). In addition, parallel RT-qPCR analyses further revealed that METTL3 and ALKBH5 directly or indirectly regulate the mRNA metabolism of small GTPases and thus modulate their expression at the protein level (Figure 4.4, Figure S4.7). These observations resemble recent findings that some epitranscriptomic regulators assume important roles in oncogenesis.[47,48] Our findings that the tumor suppressor RHOB and metastatic driver RHOC are regulated by METTL3 and ALKBH5 provide potential targets for therapeutic interventions of human cancer. Along this line, several small molecules were discovered to modulate the activities of m^6A methyltransferase (i.e., METTL3) or demethylases (i.e., ALKBH5 and FTO).[23,49-51]

Interestingly, we found that those small GTPases regulated by METTL3 and ALKBH5 in HEK293T cells also possess higher discrepancy in the expression between mRNA and protein levels in paired primary/metastatic cancer cells derived from the same patients (Figure S4.6, Table S4.3). This finding, despite with limited size of the dataset, suggests that the m^6A writer and eraser proteins may be involved in altering the protein

expression of small GTPases during cancer progression. Moreover, this observation indicates that the m^6A-mediated epitranscriptomic pathway may constitute a very important mechanism contributing to the discrepancy between the transcriptome and proteome. We realized that only a modest number of small GTPases were modulated by METTL3 and ALKBH5; hence, it will be important to examine, in the future, whether the m^6A-mediated epitranscriptomic pathway constitutes a major mechanism contributing to the differences between the transcriptome and proteome. This can be achieved by assessing how these m^6A writer and eraser proteins affect protein expression at the entire proteome level.

In summary, we demonstrated that depletion of METTL3 and ALKBH5 perturbed the expression of small GTPase proteins in cells. Interestingly, those small GTPases targeted by METTL3 and ALKBH5 also exhibited higher discrepancy between the transcriptome and proteome in paired primary/metastatic melanoma and colorectal cancer cells. To our knowledge, this is the first comprehensive analysis of the involvements of m^6A epitranscriptomic modulators in the expression of small GTPases at the protein level, and our findings lay the foundation for alternative therapeutic approach to tackle the abnormal protein expression of the currently "undruggable" small GTPases.

4.5 Acknowledgements

This work was supported by the National Institutes of Health (R35 ES031707), and M.H. was supported by an NRSA T32 Institutional Training Grant (T32 ES018827).

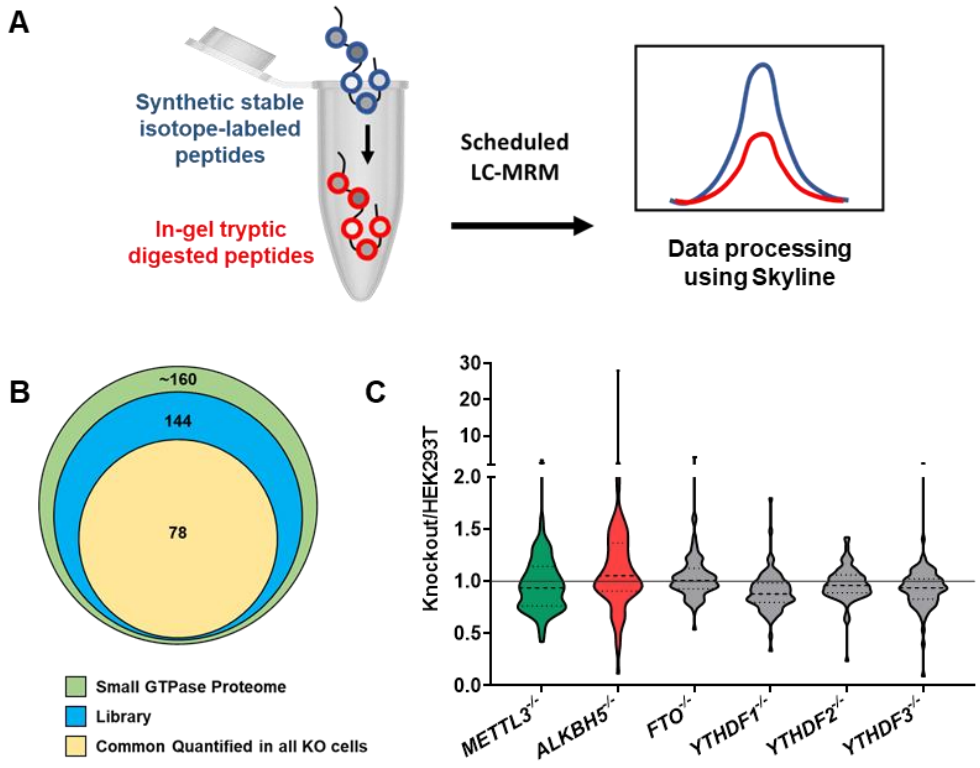


Figure 4.1 MRM-based targeted proteomic analysis for protein expression profiling of small GTPases in cells depleted of epitranscriptomic modulators.

(A) A schematic diagram illustrating the workflow of MRM-based targeted proteomic analysis of the small GTPase proteome. In-gel tryptic digested peptides were spiked-in with a mixture of synthetic stable isotope-labeled peptides of identical amino acid sequences prior to LC-MRM analysis. The LC-MS/MS data were processed using Skyline. (B) A Venn diagram showing the number of small GTPases in human proteome, the current MRM library, and the commonly quantified in all knockout cells studied herein. (C) Violin plots displaying the ratio distributions of knockout/parental cells of the quantified small GTPases in cells depleted of m^6A writer, eraser, or reader proteins.

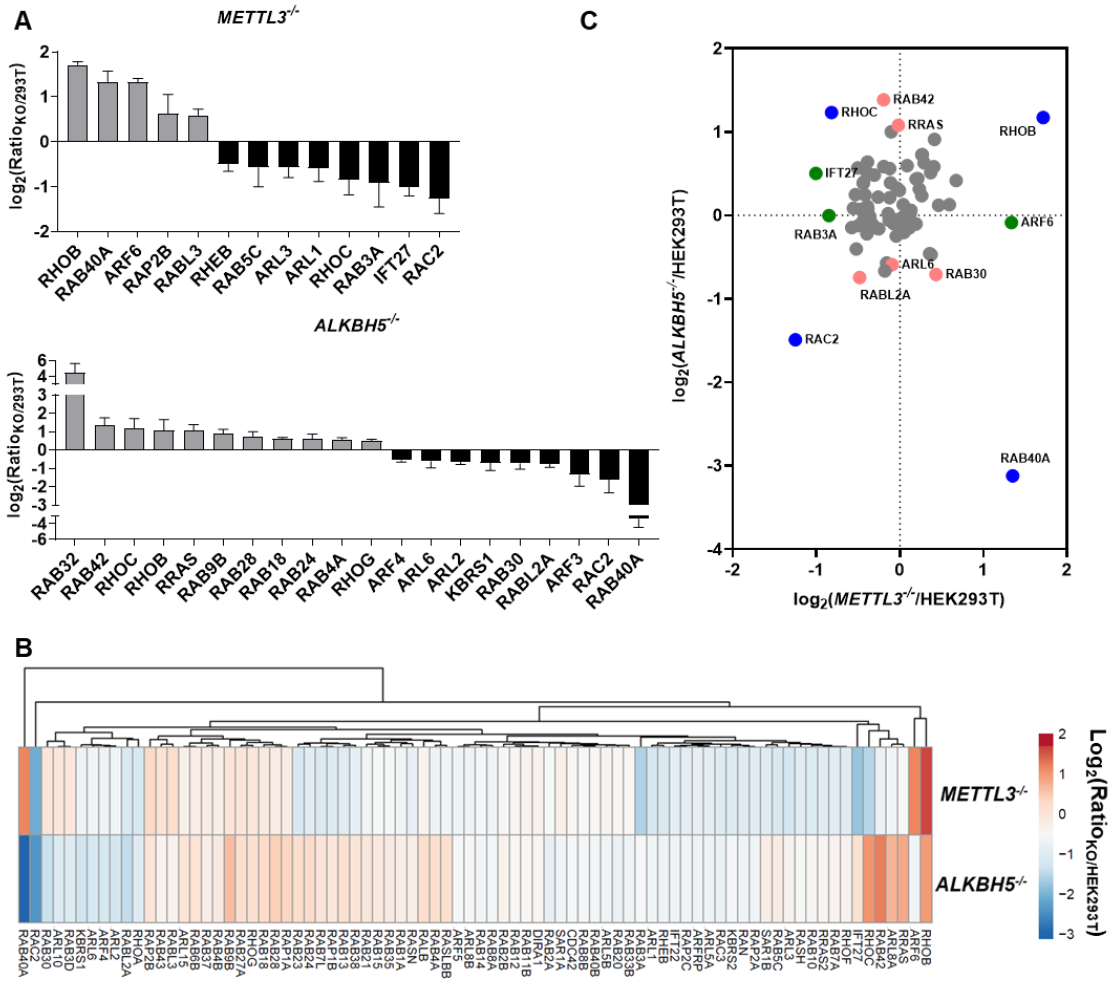


Figure 4.2 Modulations of the differential expression of the small GTPase proteome by m^{6}A writer and eraser proteins.

(A) A bar chart showing the substantially (>1.5-fold) up- (grey) or down-regulated (black) small GTPases quantified from three (for *METTL3*$^{-/-}$) or four (for *ALKBH5*$^{-/-}$) independent LC-MRM experiments. (B) Hierarchical clustering of the differential expression of small GTPase proteins between *METTL3*$^{-/-}$ vs. HEK293T and *ALKBH5*$^{-/-}$ vs. HEK293T cells. (C) A dot plot showing the significantly up- and down-regulated small GTPases in the

METTL3$^{-/-}$ (green), *ALKBH5*$^{-/-}$ (red), or both (blue) cell lines relative to the isogenic parental HEK293T cells.

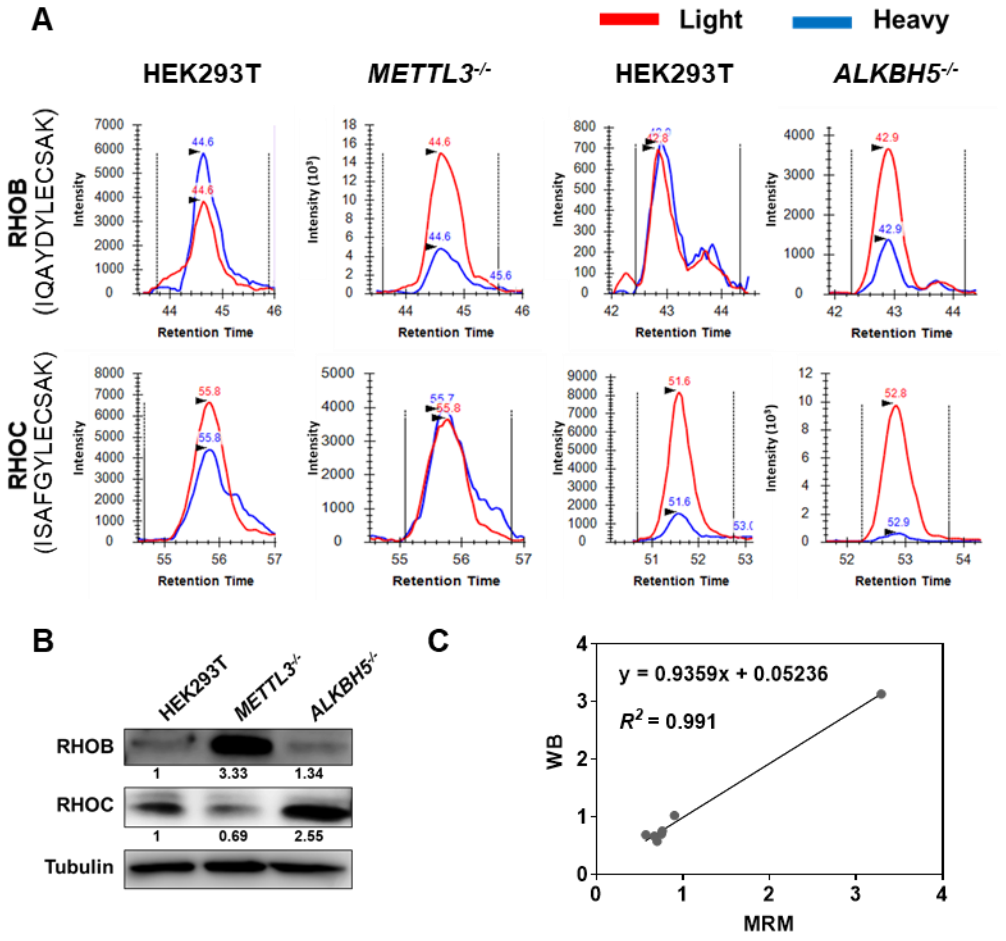


Figure 4.3 Western blot validation of the quantification results obtained from scheduled LC-MRM analysis.

(A) Ion chromatograms showing the traces of the three monitored transitions for unique tryptic peptides of RHOB (y_7, y_8, and y_9) (upper panel) and RHOC (y_6, y_8, and y_9) (lower panel) in *METTL3*$^{-/-}$ and *ALKBH5*$^{-/-}$ in comparison with parental HEK293T cells. The traces for the unlabeled peptides in parental or knockout cells are shown in red and the corresponding traces for the spiked-in heavy isotope-labeled peptides are displayed in blue. (B) Western blot for validating the MRM results of RHOB and RHOC proteins in *METTL3*$^{-/-}$ and *ALKBH5*$^{-/-}$ cells vs. parental HEK293T cells. The band intensity of small GTPases

was normalized against that of tubulin and further normalized against the corresponding ratio obtained for parental HEK293T cells. (C) Linear regression analysis of the quantification results of small GTPase expression in protein level obtained from the LC-MRM and Western blot analysis in *METTL3*$^{-/-}$ cells vs. parental HEK293T cells.

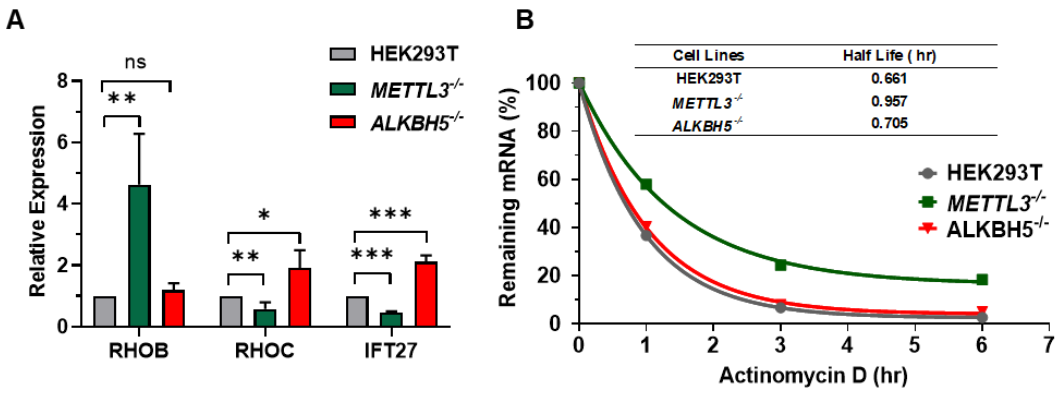


Figure 4.4 Expression and stability of selected small GTPase mRNA.

(A) RT-qPCR results for the steady-state mRNA expression levels of *RHOB*, *RHOC*, and *IFT27* genes, which encode RHOB, RHOC, and IFT27 proteins, respectively, in HEK293T cells and the isogenic *METTL3*$^{-/-}$ and *ALKBH5*$^{-/-}$ cells. The levels of transcripts of small GTPase genes were normalized against that of the *GAPDH* gene. The data represent mean ± SD of results obtained from three biological replicates. The *p* values were calculated using two-tailed, unpaired *t*-test (*, $0.01 < p < 0.05$; **, $0.001 < p < 0.01$; ***, $p < 0.001$; ns, $p > 0.05$) (B) RT-qPCR results showing the stability of *RHOB* mRNA. Cells were treated with actinomycin D to block transcription in a reverse chronological order prior to RNA extraction. The level of *RHOB* mRNA at each time point was normalized against that of *GAPDH* and further normalized against that at 0 hr. The percentage of remaining mRNA was fit with single-phase exponential decay kinetics to calculate the half-life.

4.6 References

1. van Dam, T.J., Bos, J.L. & Snel, B. Evolution of the Ras-like small GTPases and their regulators. *Small GTPases* **2**, 4-16 (2011).

2. Ridley, A.J. Rho GTPase signalling in cell migration. *Curr. Opin. Cell Biol.* **36**, 103-12 (2015).

3. Saxena, S., Bucci, C., Weis, J. & Kruttgen, A. The small GTPase Rab7 controls the endosomal trafficking and neuritogenic signaling of the nerve growth factor receptor TrkA. *J. Neurosci.* **25**, 10930-40 (2005).

4. Sahai, E. & Marshall, C.J. RHO-GTPases and cancer. *Nat. Rev. Cancer* **2**, 133-42 (2002).

5. Svensmark, J.H. & Brakebusch, C. Rho GTPases in cancer: friend or foe? *Oncogene* **38**, 7447-7456 (2019).

6. Fernandez-Medarde, A. & Santos, E. Ras in cancer and developmental diseases. *Genes Cancer* **2**, 344-58 (2011).

7. McCormick, F. KRAS as a Therapeutic Target. *Clin. Cancer Res.* **21**, 1797-1801 (2015).

8. Duan, X. et al. The small GTPase RhoA regulates the LIMK1/2-cofilin pathway to modulate cytoskeletal dynamics in oocyte meiosis. *J. Cell Physiol.* **233**, 6088-6097 (2018).

9. Cox, A.D., Fesik, S.W., Kimmelman, A.C., Luo, J. & Der, C.J. Drugging the undruggable RAS: Mission possible? *Nat. Rev. Drug Discov.* **13**, 828-51 (2014).

10. Kessler, D. et al. Drugging an undruggable pocket on KRAS. *Proc. Natl. Acad. Sci. U. S. A.* **116**, 15823-15829 (2019).

11. Perry, R.P. & Kelley, D.E. Existence of Methylated Messenger-Rna in Mouse L Cells. *Cell* **1**, 37-42 (1974).

12. Liu, J. et al. A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. *Nat. Chem. Biol.* **10**, 93-5 (2014).

13. Zhao, X. et al. FTO-dependent demethylation of N^6-methyladenosine regulates mRNA splicing and is required for adipogenesis. *Cell Res.* **24**, 1403-19 (2014).

14. Zheng, G. et al. ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol. Cell* **49**, 18-29 (2013).

15. Du, H. et al. YTHDF2 destabilizes m^{6}A-containing RNA through direct recruitment of the CCR4-NOT deadenylase complex. *Nat. Commun.* **7**, 12626 (2016).

16. Wang, X. et al. *N*^{6}-methyladenosine-dependent regulation of messenger RNA stability. *Nature* **505**, 117-20 (2014).

17. Shi, H. et al. YTHDF3 facilitates translation and decay of *N*^{6}-methyladenosine-modified RNA. *Cell Res.* **27**, 315-328 (2017).

18. Wang, X. et al. N^6-methyladenosine modulates messenger RNA translation efficiency. *Cell* **161**, 1388-99 (2015).

19. Li, A. et al. Cytoplasmic m^{6}A reader YTHDF3 promotes mRNA translation. *Cell Res.* **27**, 444-447 (2017).

20. Meyer, Kate D. et al. 5′ UTR m^{6}A Promotes cap-independent translation. *Cell* **163**, 999-1010 (2015).

21. Lin, S., Choe, J., Du, P., Triboulet, R. & Gregory, R.I. The m^{6}A Methyltransferase METTL3 Promotes Translation in Human Cancer Cells. *Mol. Cell* **62**, 335-345 (2016).

22. Zhang, S. et al. m^{6}A demethylase ALKBH5 maintains tumorigenicity of glioblastoma stem-like cells by sustaining FOXM1 expression and cell proliferation program. *Cancer Cell* **31**, 591-606 e6 (2017).

23. Huang, Y. et al. Small-molecule targeting of oncogenic FTO demethylase in acute myeloid leukemia. *Cancer Cell* **35**, 677-691 e10 (2019).

24. Miao, W. et al. HSP90 inhibitors stimulate DNAJB4 protein expression through a mechanism involving *N*^{6}-methyladenosine. *Nat. Commun.* **10**, 3613 (2019).

25. Huang, M., Darvas, M., Keene, C.D. & Wang, Y. Targeted Quantitative Proteomic Approach for High-Throughput Quantitative Profiling of Small GTPases in Brain Tissues of Alzheimer's Disease Patients. *Anal. Chem.* **91**, 12307-12314 (2019).

26. Huang, M., Qi, T.F., Li, L., Zhang, G. & Wang, Y. A targeted quantitative proteomic approach assesses the reprogramming of small gtpases during melanoma metastasis. *Cancer Res.* **78**, 5431-5445 (2018).

27. Yang, Y.Y., Huang, M. & Wang, Y. Targeted proteomic analysis of small GTPases in murine adipogenesis. *Anal. Chem.* (2020).

28. Ran, F.A. et al. Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281-2308 (2013).

29. Shevchenko, A., Tomas, H., Havlis, J., Olsen, J.V. & Mann, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* **1**, 2856-60 (2006).

30. Pino, L.K. et al. The Skyline ecosystem: Informatics for quantitative mass spectrometry proteomics. *Mass Spectrom. Rev.* (2017).

31. Escher, C. et al. Using iRT, a normalized retention time for more targeted measurement of peptides. *Proteomics* **12**, 1111-21 (2012).

32. Frewen, B. & MacCoss, M.J. Using BiblioSpec for creating and searching tandem MS peptide libraries. *Curr. Protoc. Bioinformatics* **Chapter 13**, Unit 13 7 (2007).

33. Metsalu, T. & Vilo, J. ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. *Nucleic Acids Res.* **43**, W566-70 (2015).

34. Ries, R.J. et al. m^6A enhances the phase separation potential of mRNA. *Nature* **571**, 424-428 (2019).

35. Bai, L. et al. m^6A demethylase FTO regulates dopaminergic neurotransmission deficits caused by arsenite. *Toxicol. Sci.* **165**, 431-446 (2018).

36. Engel, M. et al. The Role of m(6)A/m-RNA Methylation in Stress Response Regulation. *Neuron* **99**, 389-403 e9 (2018).

37. Wennerberg, K., Rossman, K.L. & Der, C.J. The Ras superfamily at a glance. *J. Cell Sci.* **118**, 843-6 (2005).

38. Zhang, G. & Wujcik, C.E. Overcoming ionization effects through chromatography: a case study for the ESI-LC-MS/MS quantitation of a hydrophobic therapeutic agent in human serum using a stable-label internal standard. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **877**, 2003-10 (2009).

39. Zaccara, S. & Jaffrey, S.R. A unified model for the function of YTHDF proteins in regulating m^6A-modified mRNA. *Cell*, doi: 10.1016/j.cell.2020.05.012. (2020).

40. Thomas, P., Pranatharthi, A., Ross, C. & Srivastava, S. RhoC: a fascinating journey from a cytoskeletal organizer to a Cancer stem cell therapeutic target. *J. Exp. Clin. Cancer Res.* **38**, 328 (2019).

41. Barretina, J. et al. The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603-7 (2012).

42. Huang, M. & Wang, Y. Targeted Quantitative Proteomic Approach for Probing Altered Protein Expression of Small GTPases Associated with Colorectal Cancer Metastasis. *Anal. Chem.* **91**, 6233-6241 (2019).

43. Yoo, J.H. et al. The Small GTPase ARF6 Activates PI3K in Melanoma to Induce a Prometastatic State. *Cancer Res.* **79**, 2892-2908 (2019).

44. Jiang, K. et al. Akt mediates Ras downregulation of RhoB, a suppressor of transformation, invasion, and metastasis. *Mol. Cell Biol.* **24**, 5565-76 (2004).

45. Haak, A.J. et al. Pharmacological Inhibition of Myocardin-related Transcription Factor Pathway Blocks Lung Metastases of RhoC-Overexpressing Melanoma. *Mol. Cancer Ther.* **16**, 193-204 (2017).

46. Casalou, C., Faustino, A. & Barral, D.C. Arf proteins in cancer cell migration. *Small GTPases* **7**, 270-282 (2016).

47. Choe, J. et al. mRNA circularization by METTL3-eIF3h enhances translation and promotes oncogenesis. *Nature* **561**, 556-560 (2018).

48. Barbieri, I. et al. Promoter-bound METTL3 maintains myeloid leukaemia by m(6)A-dependent translation control. *Nature* **552**, 126-131 (2017).

49. Selberg, S. et al. Discovery of Small Molecules that Activate RNA Methylation through Cooperative Binding to the METTL3-14-WTAP Complex Active Site. *Cell Rep.* **26**, 3762-3771 e5 (2019).

50. Tzelepis, K. et al. Pharmacological Inhibition of the RNA m6a Writer METTL3 As a Novel Therapeutic Strategy for Acute Myeloid Leukemia. *Blood* **134**, 403-403 (2019).

51. Malacrida, A. et al. 3D proteome-wide scale screening and activity evaluation of a new ALKBH5 inhibitor in U87 glioblastoma cell line. *Bioorg. Med. Chem.* **28**, 115300 (2020).

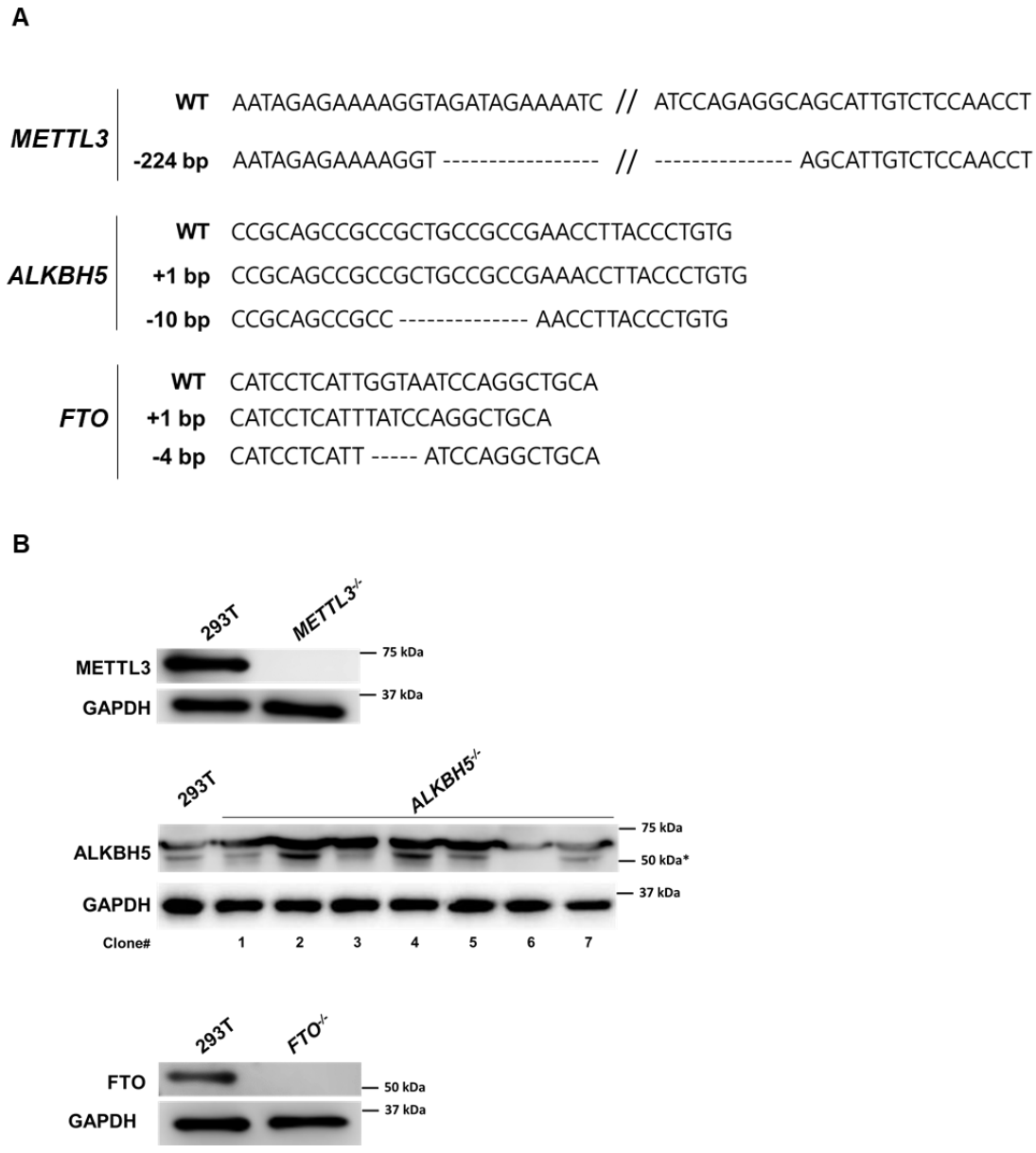


Figure S4.1 Genotyping and Western blot analyses for confirming the successful knockout of *METTL3, ALKBH5* and *FTO* genes in HEK293T cells.

(A) Genotyping of deletion mutations of *METTL3*, *ALKBH5* (clone #6), and *FTO* genes in HEK293T cells generated by using CRISPR-Cas9 (sgRNA sequences shown in Table S1).

(B) Western blot for validation of successful knockouts of target genes. Two bands were

observed in the ALKBH5 blot. The ALKBH5-specific faster migrating band (i.e. ~50 kDa) is labeled with an asterisk.

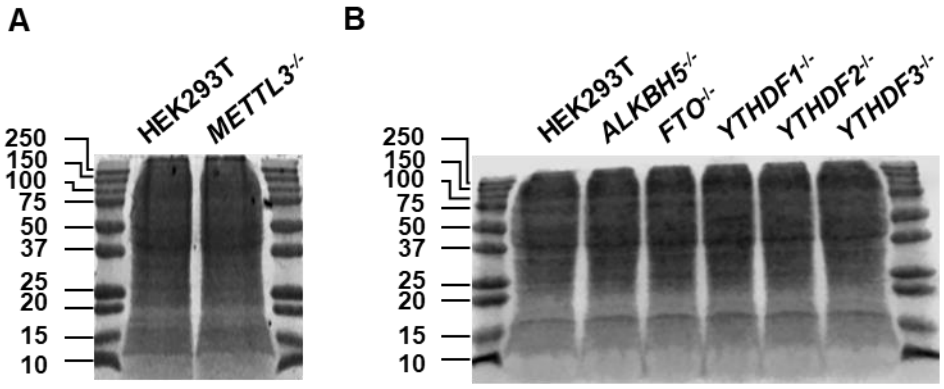
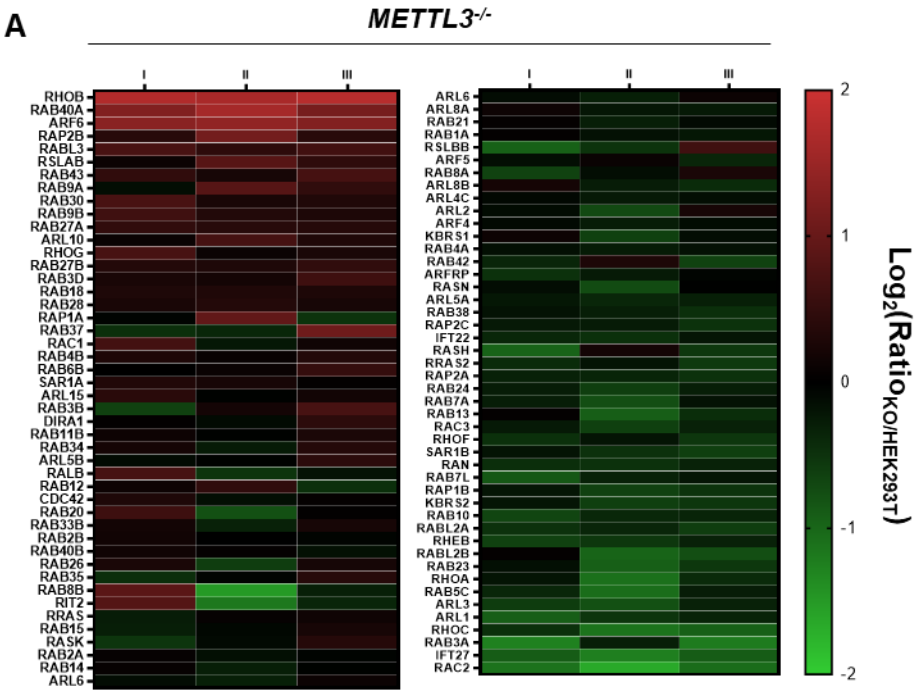
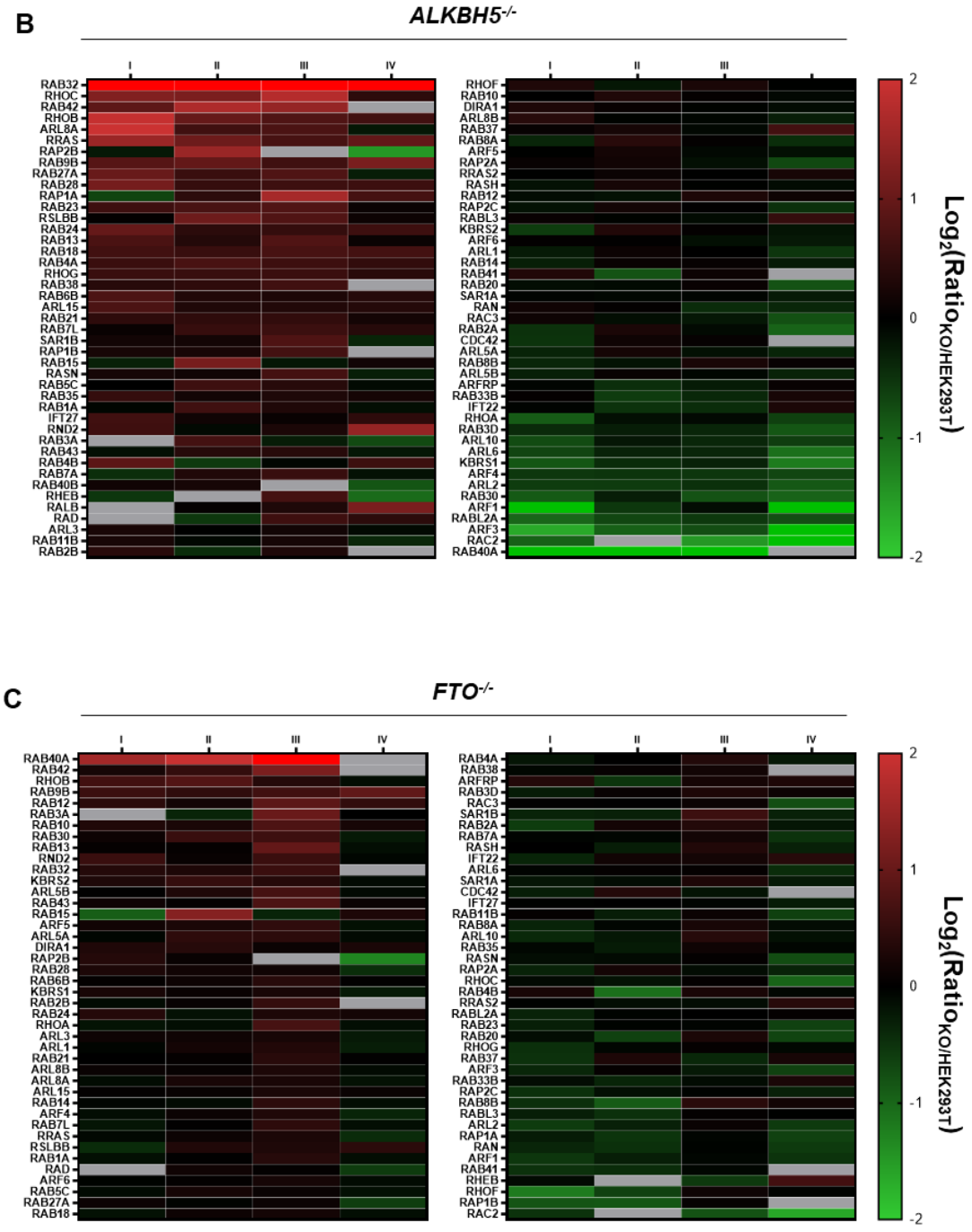
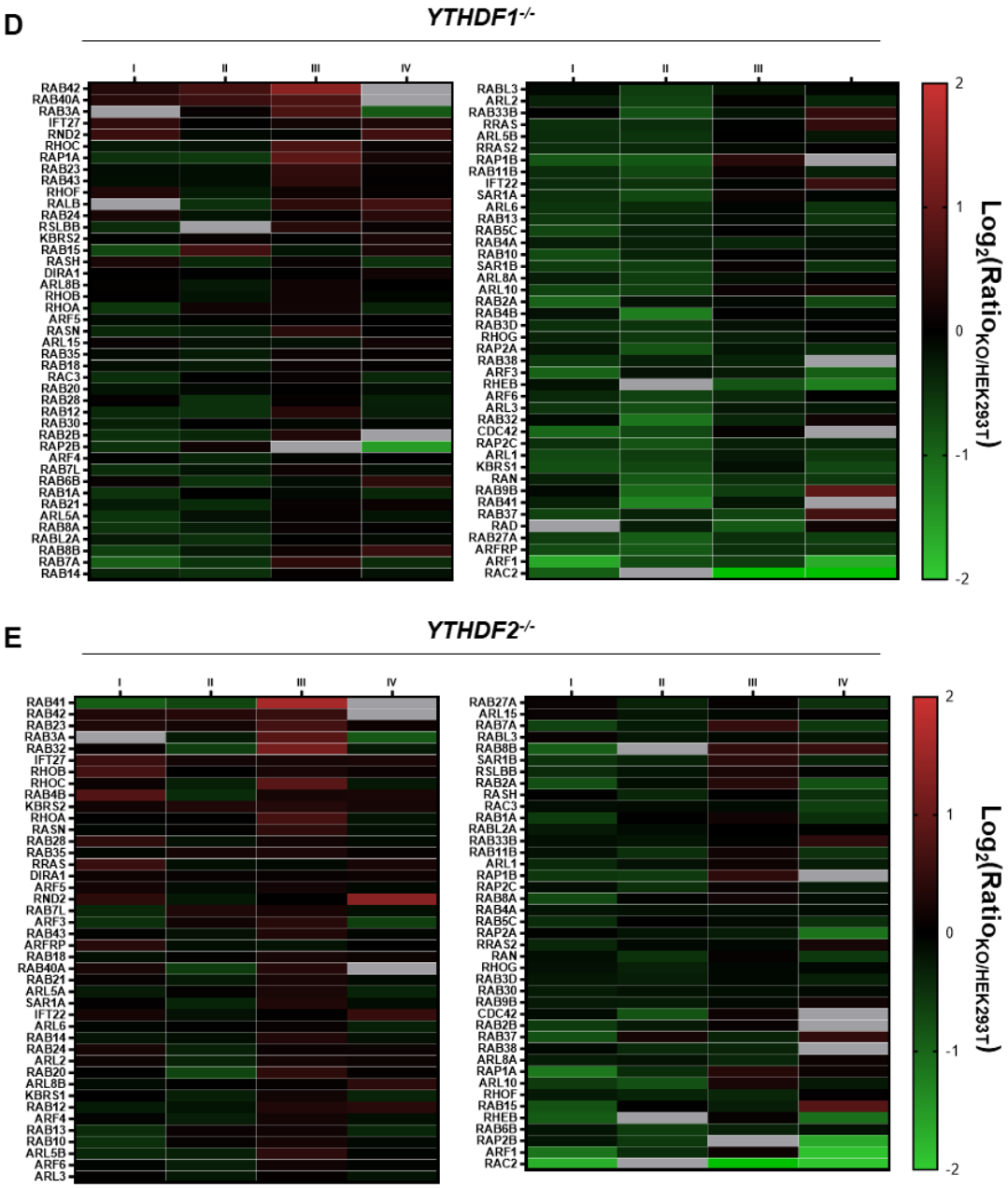


Figure S4.2 Commassie blue staining for in-gel trypsin digestion.

Cells were depleted of m^6A (A) methyltransferase (METTL3) or (B) demethylases (ALKBH5 and FTO) and reader proteins (YTHDF1, YTHDF2, and YTHDF3).


A
METTL3-/-
I
II
III
RHOB
RAB40A
ARF6
RAP2B
RABL3
RSLAB
RAB43
RAB9A
RAB30
RAB9B
RAB27A
ARL10
RHOG
RAB27B
RAB3D
RAB18
RAB28
RAP1A
RAB37
RAC1
RAB4B
RAB6B
SAR1A
ARL15
RAB3B
DIRA1
RAB11B
RAB34
ARL5B
RALB
RAB12
CDC42
RAB20
RAB33B
RAB2B
RAB40B
RAB26
RAB35
RAB8B
RIT2
RRAS
RAB15
RASK
RAB2A
RAB14
ARL6
ARL6
ARL8A
RAB21
RAB1A
RSLBB
ARF5
RAB8A
ARL8B
ARL4C
ARL2
ARF4
KBRS1
RAB4A
RAB42
ARFRP
RASN
ARL5A
RAB38
RAP2C
IFT22
RASH
RRAS2
RAP2A
RAB24
RAB7A
RAB13
RAC3
RHOF
SAR1B
RAN
RAB7L
RAP1B
KBRS2
RAB10
RABL2A
RHEB
RABL2B
RAB23
RHOA
RAB5C
ARL3
ARL1
RHOC
RAB3A
IFT27
RAC2
2
1
0
-1
-2
Log2(RatioKO/HEK293T)


B
ALKBH5-/-
I
II
III
IV
RAB32
RHOC
RAB42
RHOB
ARL8A
RRAS
RAP2B
RAB9B
RAB27A
RAB28
RAP1A
RAB23
RSLBB
RAB24
RAB13
RAB18
RAB4A
RHOG
RAB38
RAB6B
ARL15
RAB21
RAB7L
SAR1B
RAP1B
RAB15
RASN
RAB5C
RAB35
RAB1A
IFT27
RND2
RAB3A
RAB43
RAB4B
RAB7A
RAB40B
RHEB
RALB
RAD
ARL3
RAB11B
RAB2B
RHOF
RAB10
DIRA1
ARL8B
RAB37
RAB8A
ARF5
RAP2A
RRAS2
RASH
RAB12
RAP2C
RABL3
KBRS2
ARF6
ARL1
RAB14
RAB41
RAB20
SAR1A
RAN
RAC3
RAB2A
CDC42
ARL5A
RAB8B
ARL5B
ARFRP
RAB33B
IFT22
RHOA
RAB3D
ARL10
ARL6
KBRS1
ARF4
ARL2
RAB30
ARF1
RABL2A
ARF3
RAC2
RAB40A
2
1
0
-1
-2
Log2(RatioKO/HEK293T)
C
FTO-/-
I
II
III
IV
RAB40A
RAB42
RHOB
RAB9B
RAB12
RAB3A
RAB10
RAB30
RAB13
RND2
RAB32
KBRS2
ARL5B
RAB43
RAB15
ARF5
ARL5A
DIRA1
RAP2B
RAB28
RAB6B
KBRS1
RAB2B
RAB24
RHOA
ARL3
ARL1
RAB21
ARL8B
ARL8A
ARL15
RAB14
ARF4
RAB7L
RRAS
RSLBB
RAB1A
RAD
ARF6
RAB5C
RAB27A
RAB18
RAB4A
RAB38
ARFRP
RAB3D
RAC3
SAR1B
RAB2A
RAB7A
RASH
IFT22
ARL6
SAR1A
CDC42
IFT27
RAB11B
RAB8A
ARL10
RAB35
RASN
RAP2A
RHOC
RAB4B
RRAS2
RABL2A
RAB23
RAB20
RHOG
RAB37
ARF3
RAB33B
RAP2C
RAB8B
RABL3
ARL2
RAP1A
RAN
ARF1
RAB41
RHEB
RHOF
RAP1B
RAC2
2
1
0
-1
-2
Log2(RatioKO/HEK293T)


D
YTHDF1-/-
I
II
III
IV
RAB42
RAB40A
RAB3A
IFT27
RND2
RHOC
RAP1A
RAB23
RAB43
RHOF
RALB
RAB24
RSLBB
KBRS2
RAB15
RASH
DIRA1
ARL8B
RHOB
RHOA
ARF5
RASN
ARL15
RAB35
RAB18
RAC3
RAB20
RAB28
RAB12
RAB30
RAB2B
RAP2B
ARF4
RAB7L
RAB6B
RAB1A
RAB21
ARL5A
RAB8A
RABL2A
RAB8B
RAB7A
RAB14
RABL3
ARL2
RAB33B
RRAS
ARL5B
RRAS2
RAP1B
RAB11B
IFT22
SAR1A
ARL6
RAB13
RAB5C
RAB4A
RAB10
SAR1B
ARL8A
ARL10
RAB2A
RAB4B
RAB3D
RHOG
RAP2A
RAB38
ARF3
RHEB
ARF6
ARL3
RAB32
CDC42
RAP2C
ARL1
KBRS1
RAN
RAB9B
RAB41
RAB37
RAD
RAB27A
ARFRP
ARF1
RAC2
2
1
0
-1
-2
Log2(RatioKO/HEK293T)
E
YTHDF2-/-
I
II
III
IV
RAB41
RAB42
RAB23
RAB3A
RAB32
IFT27
RHOB
RHOC
RAB4B
KBRS2
RHOA
RASN
RAB28
RAB35
RRAS
DIRA1
ARF5
RND2
RAB7L
ARF3
RAB43
ARFRP
RAB18
RAB40A
RAB21
ARL5A
SAR1A
IFT22
ARL6
RAB14
RAB24
ARL2
RAB20
ARL8B
KBRS1
RAB12
ARF4
RAB13
RAB10
ARL5B
ARF6
ARL3
RAB27A
ARL15
RAB7A
RABL3
RAB8B
SAR1B
RSLBB
RAB2A
RASH
RAC3
RAB1A
RABL2A
RAB33B
RAB11B
ARL1
RAP1B
RAP2C
RAB8A
RAB4A
RAB5C
RAP2A
RRAS2
RAN
RHOG
RAB3D
RAB30
RAB9B
CDC42
RAB2B
RAB37
RAB38
ARL8A
RAP1A
ARL10
RHOF
RAB15
RHEB
RAB6B
RAP2B
ARF1
RAC2
2
1
0
-1
-2
Log2(RatioKO/HEK293T)

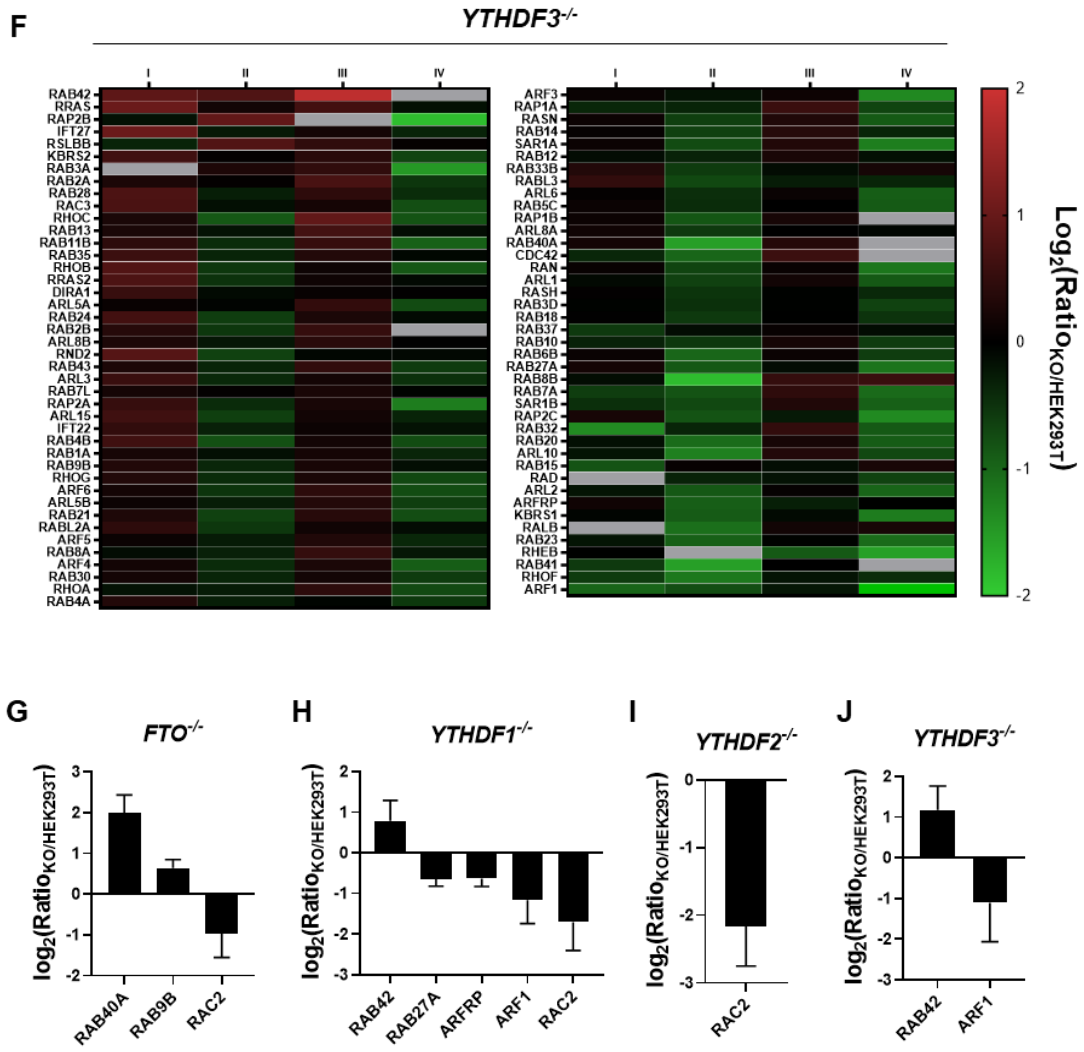


Figure S4.3 Scheduled MRM for the assessment of differential expression of small GTPase proteins in cells depleted of m^{6}A epitranscriptomic modulators.

(A-F) Heat map depicting the quantified small GTPase proteins in HEK293T cells depleted of m^{6}A methyltransferase (METTL3), demethylase (ALKBH5 and FTO), or reader proteins (YTHDF1, YTHDF2, and YTHDF3). (G-J) Quantification results of the small GTPases showing substantial up- or down-regulation (>1.5-fold) in HEK293T cells

depleted of FTO (G) or reader proteins including YTHDF1 (H), YTHDF2 (I), and YTHDF3 (J).

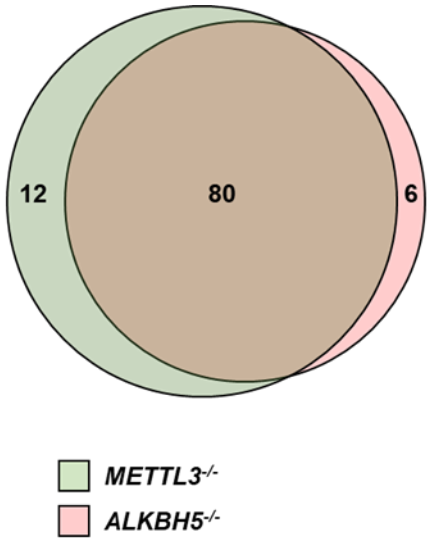


Figure S4.4 Quantified small GTPases in HEK293T cells depleted of m^{6}A methyltransferase METTL3 or demethylase ALKBH5.

A Venn diagram showing the overlap of the number of small GTPases quantified with scheduled MRM analysis in *METTL3*$^{-/-}$ cells or *ALKBH5*$^{-/-}$ cells.

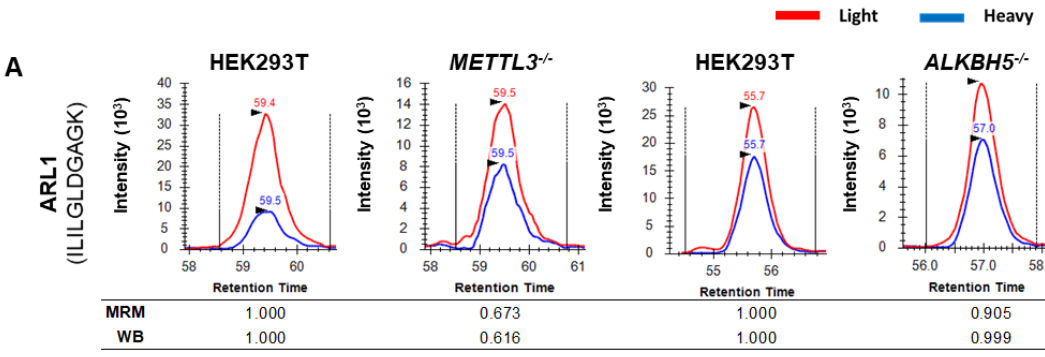

Light
Heavy
A
ARL1
(ILILGLDGAGK)
HEK293T
METTL3-/-
HEK293T
ALKBH5-/-
Intensity (10³)
Retention Time
59.4
59.5
59.5
59.5
55.7
55.7
57.0
MRM 1.000 0.673 1.000 0.905
WB 1.000 0.616 1.000 0.999
B
RAB23
(TIGVDFLER)
HEK293T
METTL3-/-
HEK293T
ALKBH5-/-
Intensity
Retention Time
58.1
58.1
58.1
58.0
55.6
55.5
55.5
55.5
MRM 1.000 0.703 1.000 1.483
WB 1.000 0.574 1.000 1.358
C
RAB7A
(FQSLGVAFYR)
HEK293T
METTL3-/-
HEK293T
ALKBH5-/-
Intensity (10³)
Retention Time
56.9
56.9
57.5
57.5
53.5
53.6
54.8
54.7
MRM 1.000 0.763 1.000 1.078
WB 1.000 0.750 1.000 1.505
D
ARF4
(LGLQSLR)
HEK293T
METTL3-/-
HEK293T
ALKBH5-/-
Intensity (10³)
Retention Time
38.2
37.2
38.2
37.5
37.5
35.0
35.0
36.2
36.2
MRM 1.000 0.897 1.000 0.675
WB 1.000 1.020 1.000 0.573

Figure S4.5 Western blot analyses for validating the quantification results for several small GTPases obtained from scheduled LC-MRM analysis.

Ion chromatograms showing the traces of (A) ARL1, (B) RAB23, (C) RAB7A, and (D) ARF4 in *METTL3*$^{-/-}$ and *ALKBH5*$^{-/-}$ cells. The tables below the ion chromatograms summarize the quantification results obtained from MRM and Western blot analyses.

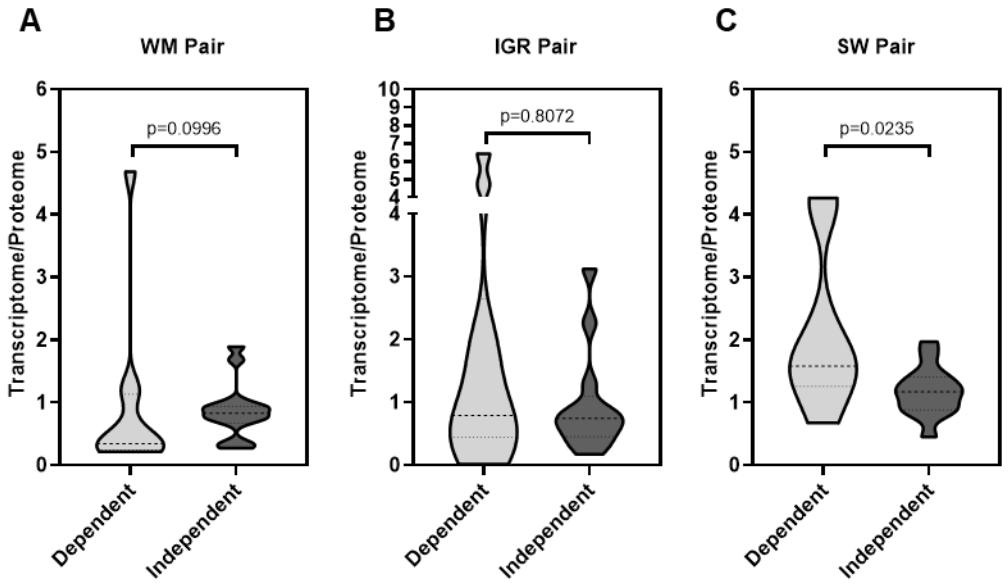


Figure S4.6 A comparison of the differential mRNA and protein expression levels of small GTPases in matched pairs of primary/metastatic melanoma or colorectal cancer cells.

Displayed are the violin plots revealing the expressions of small GTPases between the transcriptome and proteome of (A) melanoma WM266-4/WM115, (B) IGR37/IGR39 and (C) colorectal SW620/SW480 paired metastatic/primary cancer cells. The dependent group represents those small GTPases showing significant altered expression in protein level in cells depleted of METTL3 and/or ALKBH5; the independent group designates the subset of small GTPases that are not regulated by METTL3 or ALKBH5 (Figure 4.2B, Table S4.2).

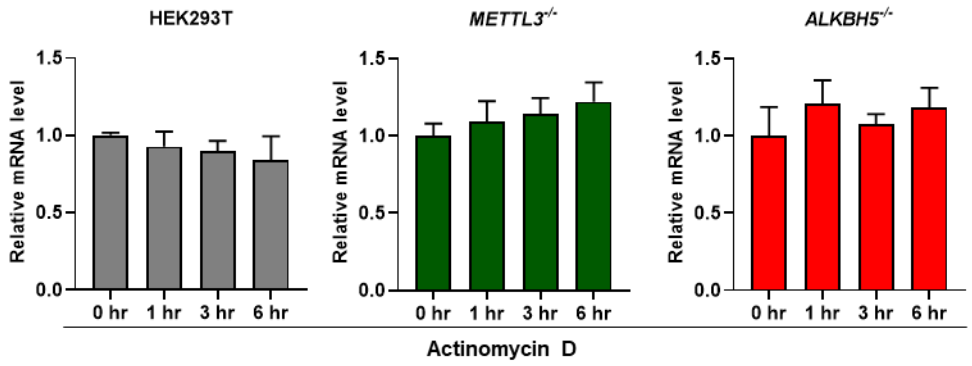


Figure S4.7 mRNA stability of selected small GTPases.

Bar chart of RT-qPCR to show the stability of *RHOC* mRNA in cells treated with actinomycin D in a reverse chronological order prior to RNA extraction. The mRNA expression of *RHOC* at each time point was normalized against that of *GAPDH* and further normalized to value at 0 hr.

Table S4.1 A list of sgRNA sequences used to generate deletion mutations of *METTL3, ALKBH5,* and *FTO* genes in HEK293T cells.

Gene	Orientation	sgRNA Sequence	Cut Position	PAM
METTL3	Sense	CTTGGATCTACGGAATCCAG	Exon 2	AGG
ALKBH5	Antisense	GGACACAGGGTAAGGTTCGG	Exon 1	AGG
FTO	Antisense	TGCAGCCTGGATTACCAATG	Exon 3	AGG

Table S4.2A A complete list of all small GTPases quantified in the scheduled LC-MRM analysis in cells depleted of eraser and reader proteins.

Uniprot Accession	Gene	Peptide	Subfamily	*ALKBH5* $^{-/-}$							
				Biological #1		Biological #2		Biological #3		Biological 4	
				L/H	dotp	L/H	dotp	L/H	dotp	L/H	dotp
P84077	ARF1	MGNIFANLFK	ARF	1.40	0.89	0.99	0.88	3.24	0.89	0.40	0.87
P61204	ARF3	MGNIFGNLLK	ARF	0.23	0.87	0.04	0.90	0.14	0.87	0.08	0.83
P18085	ARF4	LGLQSLR	ARF	15.61	0.93	7.74	0.94	7.06	0.94	6.73	0.93
P84085	ARF5	LGLQHLR	ARF	8.43	0.81	4.21	0.82	2.72	0.82	3.59	0.81
P62330	ARF6	FNVWDVGGQDK	ARF	27.17	0.89	9.48	0.90	11.44	0.91	13.75	0.89
Q13795	ARFRP	DCLTQACSALTGK	ARF	1.14	0.88	1.10	0.93	1.59	0.90	1.28	0.92
P40616	ARL1	ILILGLDGAGK	ARF	2.64	0.95	0.51	0.95	1.60	0.95	1.18	0.95
Q8N8L6	ARL10	EVLVLGLDGAGK	ARF	1.75	0.88	0.62	0.88	0.57	0.76	0.59	0.89
Q9NXU5	ARL15	ELGGADNIR	ARF	1.06	0.76	0.75	0.81	0.62	0.80	0.51	0.82
P36404	ARL2	ELQSLLVEER	ARF	11.64	0.96	6.57	0.96	7.30	0.97	7.09	0.97
P36405	ARL3	ILLLGLDNAGK	ARF	12.20	0.89	3.17	0.90	5.08	0.90	9.47	0.88
Q9Y689	ARL5A	AGLLIFANK	ARF	1.21	0.97	0.26	0.95	0.53	0.97	0.67	0.99
Q96KC2	ARL5B	AAVLIFANK	ARF	1.22	0.89	0.26	0.87	0.51	0.90	0.65	0.91
Q9H0F7	ARL6	IPILFFANK	ARF	0.63	0.91	0.13	0.90	0.27	0.90	0.46	0.88
Q96BM9	ARL8A	LWDIGGQPR	ARF	18.78	0.94	2.58	0.93	4.22	0.93	8.01	0.93
Q9NVJ2	ARL8B	IWDIGGQPR	ARF	16.05	0.88	4.77	0.89	7.53	0.89	7.37	0.89
Q9NR31	SAR1A	EIFGLYGQTTGK	ARF	13.95	0.82	8.05	0.83	5.97	0.83	5.83	0.81
Q9Y6B6	SAR1B	EMFGLYGQTTGK	ARF	4.73	0.80	2.04	0.81	1.99	0.82	1.88	0.80
Q9H7X7	IFT22	ILFVGPCESGK	RAB	0.23	0.95	0.12	0.98	0.16	0.98	0.19	0.95
Q9BW83	IFT27	CILAGDPAVGK	RAB	1.86	0.98	1.37	0.98	1.46	0.96	1.46	0.98
P61026	RAB10	FHTITTSYYR	RAB	8.00	0.95	1.64	0.96	4.29	0.97	5.05	0.96
Q6IQ22	RAB12	LQVIIIGSR	RAB	1.35	0.90	0.27	0.84	1.11	0.96	0.99	0.97
P51153	RAB13	LQVWDTAGQER	RAB	1.48	0.95	0.40	0.87	1.51	0.96	1.64	0.97
P61106	RAB14	TGENVEDAFLEAAK	RAB	8.77	0.95	7.97	0.97	11.78	0.98	5.40	0.94
P59190	RAB15	IQIWDTAGQER	RAB	1.22	0.84	0.48	0.89	0.61	0.82	0.68	0.87
Q9NP72	RAB18	TLTPSYYR	RAB	13.30	0.96	10.62	0.96	11.32	0.95	8.87	0.96
P62820	RAB1A	MGPGATAGGAEK	RAB	15.94	0.72	6.65	0.74	8.84	0.74	10.88	0.71
Q9NX57	RAB20	IVLLGDMNVGK	RAB	0.40	0.89	0.14	0.86	0.24	0.92	0.18	0.86
Q9UL25	RAB21	VVLLGEGCVGK	RAB	7.48	0.96	4.46	0.95	5.57	0.96	3.65	0.95
Q9ULC3	RAB23	TIGVDFLER	RAB	7.75	0.95	2.88	0.95	4.33	0.95	4.91	0.97
Q969Q5	RAB24	AQLFETSSK	RAB	1.61	0.96	0.90	0.95	1.00	0.96	1.06	0.96
P51157	RAB28	VAAEILGIK	RAB	0.36	0.90	0.21	0.92	0.25	0.87	0.25	0.92
P61019	RAB2A	MITIDGK	RAB	12.31	0.88	4.49	0.88	5.51	0.87	4.04	0.89
Q8WUD1	RAB2B	IGPQQSISTSVGPSASQR	RAB	3.54	0.93	1.12	0.90	2.68	0.93		
Q15771	RAB30	IVLIGNAGVGK	RAB	0.86	0.81	0.49	0.78	0.63	0.85	0.52	0.83
Q13637	RAB32	VLVIGELGVGK	RAB	3.20	0.97	0.77	0.96	4.12	0.97	1.51	0.94
Q15286	RAB35	LLIIGDSGVGK	RAB	0.72	0.92	0.27	0.93	0.87	0.92	0.91	0.92
Q96AX2	RAB37	VMLLGDTGVGK	RAB	0.24	0.78	0.10	0.72	0.10	0.77	0.62	0.96
P57729	RAB38	LLVIGDLGVGK	RAB	6.07	0.80	2.34	0.82	3.95	0.82		
P20336	RAB3A	TYSWDNAQVLLVGNK	RAB			0.91	0.89	0.48	0.89	0.21	0.99
O95716	RAB3D	LLLIGNSSVGK	RAB	1.06	0.94	0.31	0.93	0.76	0.94	0.55	0.93
Q5JT25	RAB41	LLFLGEQSVGK	RAB	0.07	0.87	0.02	0.83	0.03	0.92		
Q8N4Z0	RAB42	VIFLLVGHK	RAB	0.64	0.88	0.22	0.87	0.28	0.92		

Uniprot Accession	Gene	Peptide	Subfamily	*ALKBH5* $^{-/-}$							
				Biological #1		Biological #2		Biological #3		Biological 4	
				L/H	dotp	L/H	dotp	L/H	dotp	L/H	dotp
Q86YS6	RAB43	VATELIMR	RAB	1.40	0.89	0.55	0.91	0.65	0.90	0.76	0.89
P20338	RAB4A	IESGELDPER	RAB	5.88	0.94	3.92	0.95	3.96	0.95	2.60	0.94
P61018	RAB4B	FLVIGSAGTGK	RAB	0.19	1.00	0.11	0.95			0.10	0.96
P61020	RAB5B	QASPSIVIALAGNK	RAB	53.11	0.94	20.76	0.94				
P51148	RAB5C	QASPNIVIALAGNK	RAB	56.73	0.92	24.62	0.95	42.92	0.95	45.82	0.91
Q9NRW1	RAB6B	QITIEEGEQR	RAB	7.02	0.90	4.26	0.94	4.12	0.93	2.78	0.90
P51149	RAB7A	FQSLGVAFYR	RAB	66.81	0.95	38.21	0.97	78.70	0.97	34.57	0.94
O14966	RAB7L	VLVVGDAAVGK	RAB	0.25	0.85	0.10	0.88	0.37	0.88	0.26	0.88
P61006	RAB8A	LALDYGIK	RAB	0.73	0.92	0.35	0.92	0.93	0.91	0.75	0.92
Q92930	RAB8B	NIEEHASSDVER	RAB	1.75	0.84	1.31	0.73	1.77	0.83	1.30	0.79
Q9NP90	RAB9B	QSFENLGNWQK	RAB	1.22	0.91	0.50	0.95	0.49	0.91	0.80	0.91
Q5HYI8	RABL3	YLAAGSSNAVK	RAB	1.88	0.94	1.50	0.96	1.23	0.96	1.93	0.97
Q15907	RB11B	NILTEIYR	RAB	52.60	0.99	13.52	0.99	17.14	0.99	26.01	0.98
P51159	RB27A	SWIPEGVVR	RAB	6.79	0.94	1.67	0.93	2.35	0.93	2.24	0.94
Q9H082	RB33B	SAIQVPTDLAQK	RAB	0.70	0.90	0.33	0.82	0.59	0.87	0.53	0.88
Q8WXH6	RB40A	LPLPSTLR	RAB	0.12	0.78	0.22	0.87	0.09	0.82		
Q12829	RB40B	EIDEHAPGVPK	RAB	2.30	0.75	1.60	0.70			0.77	0.74
Q9UBK7	RBL2A	IICLGDSAVGK	RAB	0.40	0.96	0.26	0.96	0.49	0.92	0.31	0.97
P62826	RAN	FNVWDTAGQEK	RAN	148.48	0.98	44.02	0.99	54.83	0.99	65.03	0.98
O95057	DIRA1	EAQAVAQEWK	RAS	0.65	0.91	0.24	0.91	0.39	0.93	0.40	0.93
Q9NYS0	KBRS1	TLIEPFTLLASK	RAS	0.93	0.96	1.14	0.94	1.26	0.94	0.67	0.97
Q9NYR9	KBRS2	VVVCGQASVGK	RAS	1.16	0.86	0.93	0.90	1.04	0.93	0.85	0.88
P55042	RAD	SMPVDER	RAS			0.51	0.81	2.25	0.84	1.82	0.83
P11234	RALB	AEEWGVQYVETSAK	RAS			0.94	0.93	1.28	0.96	0.82	0.92
P62834	RAP1A	QWCNCAFLESSAK	RAS	17.96	0.93	17.24	0.93	22.29	0.94	20.89	0.92
P61224	RAP1B	QWNNCAFLESSAK	RAS	64.59	0.96	44.19	0.96	40.56	0.96		
P10114	RAP2A	VPVILVGNK	RAS	0.11	0.85	0.05	0.85	0.05	0.85	0.02	0.89
P61225	RAP2B	ASVDELFAEIVR	RAS	54.65	0.92	223.18	0.91			20.50	0.93
Q9Y3L5	RAP2C	VPLILVGNK	RAS	0.46	0.84	0.24	0.80	0.21	0.80	0.11	0.82
P01112	RASH	SFEDIHQYR	RAS	0.58	0.86	0.73	0.89	0.54	0.89	0.44	0.86
P01111	RASN	SFADINLYR	RAS	16.23	0.99	5.39	0.99	9.59	0.98	7.88	0.99
Q15382	RHEB	ALAESWNAAFLESSAK	RAS	11.30	0.98			25.26	0.98	6.45	0.95
P10301	RRAS	LFTQILR	RAS	0.85	0.89	0.25	0.90	0.20	0.92	0.33	0.91
P62070	RRAS2	QVTQEEGQQLAR	RAS	10.96	0.89	9.84	0.93	10.21	0.93	7.33	0.89
Q9BPW5	RSLBB	TSLIPRPK	RAS	0.55	0.74	0.16	0.74	0.23	0.75	0.43	0.73
P60953	CDC42	TCLLISYTTNK	RHO	20.33	0.98	12.60	0.97	13.23	0.98	4.93	0.97
P60763	RAC3	LAPITYPQGLAMAR	RHO	17.39	0.97	5.06	0.97	6.75	0.97	13.82	0.96
P61586	RHOA	IGAFGYMECSAK	RHO	31.63	0.97	19.00	0.97	18.64	0.96	24.20	0.97
P62745	RHOB	IQAYDYLECSAK	RHO	3.12	0.96	1.16	0.93	1.40	0.91	1.23	0.98
P08134	RHOC	ISAFGYLECSAK	RHO	15.29	0.95	15.12	0.97	15.95	0.98	6.58	0.94
Q9HBH0	RHOF	AALYLECSAK	RHO	0.23	0.92	0.07	0.98	0.27	0.85	0.40	0.81
P84095	RHOG	YLECSALQQDGVK	RHO	4.14	0.91	2.06	0.92	2.20	0.93	1.60	0.92
P52198	RND2	IVVVGDAECGK	RHO	0.11	0.91	0.08	0.96	0.12	0.88	0.08	0.89

Uniprot Accession	Gene	Peptide	Subfamily	*FTO* $^{-/-}$							
				Biological #1		Biological #2		Biological #3		Biological 4	
				L/H	dotp	L/H	dotp	L/H	dotp	L/H	dotp
P84077	ARF1	MGNIFANLFK	ARF	5.04	0.89	1.16	0.90	3.38	0.90	1.40	0.88
P61204	ARF3	MGNIFGNLLK	ARF	0.55	0.96	0.09	0.92	0.17	0.94	0.22	0.84
P18085	ARF4	LGLQSLR	ARF	20.86	0.93	11.78	0.94	12.62	0.93	8.48	0.93
P84085	ARF5	LGLQHLR	ARF	9.52	0.81	4.42	0.82	3.94	0.82	3.57	0.79
P62330	ARF6	FNVWDVGGQDK	ARF	26.30	0.89	9.58	0.90	14.78	0.90	15.04	0.88
Q13795	ARFRP	DCLTQACSALTGK	ARF	1.49	0.90	1.07	0.95	2.20	0.90	1.47	0.88
P40616	ARL1	ILILGLDGAGK	ARF	3.04	0.96	0.53	0.96	1.99	0.97	1.39	0.95
Q8N8L6	ARL10	EVLVLGLDGAGK	ARF	2.21	0.90	0.61	0.83	0.96	0.93	0.79	0.94
Q9NXU5	ARL15	ELGGADNIR	ARF	0.65	0.81	0.64	0.85	0.59	0.80	0.42	0.86
P36404	ARL2	ELQSLLVEER	ARF	11.71	0.96	7.84	0.97	10.77	0.96	8.67	0.96
P36405	ARL3	ILLLGLDNAGK	ARF	11.31	0.90	3.49	0.91	5.26	0.90	8.48	0.88
Q9Y689	ARL5A	AGLLIFANK	ARF	1.49	0.98	0.30	0.97	0.79	0.99	0.81	0.98
Q96KC2	ARL5B	AAVLIFANK	ARF	1.53	0.90	0.30	0.87	0.85	0.89	0.77	0.87
Q9H0F7	ARL6	IPILFFANK	ARF	0.97	0.89	0.17	0.90	0.36	0.89	0.73	0.91
Q96BM9	ARL8A	LWDIGGQPR	ARF	4.42	0.92	1.96	0.93	3.00	0.93	8.36	0.93
Q9NVJ2	ARL8B	IWDIGGQPR	ARF	13.17	0.89	5.05	0.90	9.47	0.88	8.35	0.88
Q9NR31	SAR1A	EIFGLYGQTTGK	ARF	12.41	0.82	7.78	0.83	7.48	0.83	5.89	0.80
Q9Y6B6	SAR1B	EMFGLYGQTTGK	ARF	3.28	0.80	1.44	0.81	1.76	0.81	1.92	0.80
Q9H7X7	IFT22	ILFVGPCESGK	RAB	0.19	0.95	0.20	0.97	0.25	0.95	0.20	0.95
Q9BW83	IFT27	CILAGDPAVGK	RAB	1.19	0.99	1.24	0.97	1.33	0.99	0.90	0.98
P61026	RAB10	FHTITTSYYR	RAB	9.88	0.95	1.59	0.96	7.00	0.96	6.18	0.94
Q6IQ22	RAB12	LQVIIIGSR	RAB	1.95	0.90	0.34	0.88	1.70	0.97	1.26	0.92
P51153	RAB13	LQVWDTAGQER	RAB	0.92	0.92	0.32	0.90	1.71	0.97	1.40	0.94
P61106	RAB14	TGENVEDAFLEAAK	RAB	10.51	0.95	6.97	0.98	15.12	0.98	5.90	0.93
P59190	RAB15	IQIWDTAGQER	RAB	0.80	0.84	0.52	0.86	0.55	0.80	0.73	0.90
Q9NP72	RAB18	TLTPSYYR	RAB	7.61	0.95	7.97	0.93	8.01	0.93	5.12	0.96
P62820	RAB1A	MGPGATAGGAEK	RAB	14.93	0.72	4.38	0.74	9.27	0.73	10.95	0.72
Q9NX57	RAB20	IVLLGDMNVGK	RAB	0.40	0.90	0.10	0.89	0.27	0.92	0.20	0.90
Q9UL25	RAB21	VVLLGEGCVGK	RAB	5.54	0.96	3.37	0.96	5.27	0.96	3.21	0.94
Q9ULC3	RAB23	TIGVDFLER	RAB	4.15	0.96	1.74	0.95	2.45	0.95	2.93	0.96
Q969Q5	RAB24	AQLFETSSK	RAB	1.04	0.96	0.61	0.92	0.86	0.93	0.81	0.97
P51157	RAB28	VAAEILGIK	RAB	0.20	0.90	0.17	0.94	0.19	0.92	0.12	0.93
P61019	RAB2A	MITIDGK	RAB	11.38	0.89	4.26	0.88	7.48	0.89	6.77	0.89
Q8WUD1	RAB2B	IGPQQSISTSVGPSASQR	RAB	2.51	0.94	1.58	0.94	3.11	0.88		
Q15771	RAB30	IVLIGNAGVGK	RAB	1.73	0.92	0.86	0.91	1.66	0.88	0.86	0.83
Q13637	RAB32	VLVIGELGVGK	RAB	0.22	0.96	0.11	0.94	0.13	0.73		
Q15286	RAB35	LLIIGDSGVGK	RAB	0.48	0.92	0.19	0.93	0.78	0.94	0.76	0.94
Q96AX2	RAB37	VMLLGDTGVGK	RAB	0.17	0.71	0.11	0.79	0.08	0.86	0.46	0.94
P57729	RAB38	LLVIGDLGVGK	RAB	4.42	0.70	1.80	0.83	2.90	0.73		
P20336	RAB3A	TYSWDNAQVLLVGNK	RAB			0.45	0.88	1.18	0.86	0.36	0.99
O95716	RAB3D	LLLIGNSSVGK	RAB	1.16	0.93	0.40	0.92	1.19	0.94	1.07	0.91
Q5JT25	RAB41	LLFLGEQSVGK	RAB	0.04	0.77	0.03	0.87	0.03	0.96		
Q8N4Z0	RAB42	VIFLLVGHK	RAB	0.40	0.88	0.10	0.85	0.24	0.91		

Uniprot Accession	Gene	Peptide	Subfamily	*FTO* $^{-/-}$							
				Biological #1		Biological #2		Biological #3		Biological 4	
				L/H	dotp	L/H	dotp	L/H	dotp	L/H	dotp
Q86YS6	RAB43	VATELIMR	RAB	1.71	0.90	0.44	0.91	0.83	0.89	0.96	0.87
P20338	RAB4A	IESGELDPER	RAB	3.45	0.96	2.42	0.97	3.31	0.95	1.58	0.95
P61018	RAB4B	FLVIGSAGTGK	RAB	0.12	0.96	0.08	0.91	0.11		0.06	0.91
P61020	RAB5B	QASPSIVIALAGNK	RAB	39.41	0.93	13.06	0.93				
P51148	RAB5C	QASPNIVIALAGNK	RAB	52.90	0.92	18.78	0.95	35.13	0.95	44.20	0.91
Q9NRW1	RAB6B	QITIEEGEQR	RAB	4.29	0.91	3.71	0.94	4.42	0.93	2.17	0.90
P51149	RAB7A	FQSLGVAFYR	RAB	88.96	0.95	29.56	0.96	61.51	0.97	27.65	0.93
O14966	RAB7L	VLVVGDAAVGK	RAB	0.20	0.92	0.07	0.93	0.33	0.86	0.18	0.89
P61006	RAB8A	LALDYGIK	RAB	0.74	0.91	0.25	0.93	1.06	0.91	0.94	0.92
Q92930	RAB8B	NIEEHASSDVER	RAB	1.65	0.81	0.80	0.70	1.99	0.85	1.38	0.90
Q9NP90	RAB9B	QSFENLGNWQK	RAB	1.02	0.88	0.36	0.93	0.50	0.90	0.68	0.83
Q5HYI8	RABL3	YLAAGSSNAVK	RAB	1.45	0.97	1.11	0.97	1.35	0.97	1.36	0.96
Q15907	RB11B	NILTEIYR	RAB	46.25	0.99	11.58	0.99	15.91	0.99	21.99	0.98
P51159	RB27A	SWIPEGVVR	RAB	3.66	0.93	1.13	0.93	1.48	0.94	1.81	0.91
Q9H082	RB33B	SAIQVPTDLAQK	RAB	0.62	0.85	0.38	0.73	0.71	0.86	0.54	0.82
Q8WXH6	RB40A	LPLPSTLR	RAB	6.05	0.99	3.77	0.99	7.71	0.99		
Q12829	RB40B	EIDEHAPGVPK	RAB								
Q9UBK7	RBL2A	IICLGDSAVGK	RAB	0.63	0.96	0.41	0.99	0.72	0.96	0.55	0.93
P62826	RAN	FNVVWDTAGQEK	RAN	103.39	0.98	31.26	0.99	71.68	0.99	57.12	0.96
O95057	DIRA1	EAQAVAQEWK	RAS	0.66	0.88	0.29	0.93	0.43	0.90	0.51	0.93
Q9NYS0	KBRS1	TLIEPFTLLASK	RAS	1.96	0.95	1.63	0.94	1.88	0.93	1.31	0.95
Q9NYR9	KBRS2	VVVCGQASVGK	RAS	2.13	0.88	1.05	0.89	1.31	0.88	0.91	0.87
P55042	RAD	SMPVDER	RAS			0.82	0.81	1.51	0.88	0.92	0.89
P11234	RALB	AEEWGVQYVETSAK	RAS			0.62	0.94	0.96	0.78		
P62834	RAP1A	QWCNCAFLESSAK	RAS	23.71	0.93	8.94	0.92	7.07	0.92	8.45	0.93
P61224	RAP1B	QWNNCAFLESSAK	RAS	31.45	0.94	21.30	0.94	26.70	0.95		
P10114	RAP2A	VPVILVGNK	RAS	0.08	0.86	0.05	0.86	0.05	0.87	0.03	0.84
P61225	RAP2B	ASVDELFAEIVR	RAS	82.23	0.92	85.00	0.93			23.25	0.92
Q9Y3L5	RAP2C	VPLILVGNK	RAS	0.38	0.84	0.19	0.82	0.20	0.79	0.12	0.79
P01112	RASH	SFEDIHQYR	RAS	0.65	0.86	0.52	0.87	0.69	0.87	0.38	0.83
P01111	RASN	SFADINLYR	RAS	12.79	0.99	4.51	0.98	6.22	0.98	5.74	0.98
Q15382	RHEB	ALAESWNAAFLESSAK	RAS	15.14	0.96			10.88	0.97	20.81	0.94
P10301	RRAS	LFTQILR	RAS	0.28	0.89	0.13	0.93	0.19	0.91	0.13	0.84
P62070	RRAS2	QVTQEEGQQLAR	RAS	10.90	0.89	8.39	0.92	9.47	0.92	8.23	0.88
Q9BPW5	RSLBB	TSLIPRPK	RAS	0.40	0.75	0.09	0.73	0.16	0.74	0.53	0.74
P60953	CDC42	TCLLISYTTNK	RHO	23.60	0.97	14.51	0.98	11.22	0.95	3.02	0.97
P60763	RAC3	LAPITYPQGLAMAR	RHO	16.19	0.96	5.67	0.97	8.47	0.97	14.02	0.96
P61586	RHOA	IGAFGYMECSAK	RHO	51.26	0.97	19.30	0.97	31.64	0.97	34.10	0.96
P62745	RHOB	IQAYDYLECSAK	RHO	1.28	0.94	1.01	0.93	1.05	0.94	0.73	0.91
P08134	RHOC	ISAFGYLECSAK	RHO	6.06	0.94	5.82	0.96	4.80	0.97	2.52	0.96
Q9HBH0	RHOF	AALYLECSAK	RHO	0.08	0.91	0.05	0.95	0.25	0.81	0.39	0.78
P84095	RHOG	YLECSALQQDGVK	RHO	2.06	0.94	1.36	0.91	1.54	0.93	1.15	0.97
P52198	RND2	IVVVGDAECGK	RHO	0.11	0.94	0.09	0.99	0.15	0.92	0.03	0.87

Uniprot Accession	Gene	Peptide	Subfamily	*YTHDF1* $^{-/-}$							
				Biological #1		Biological #2		Biological #3		Biological 4	
				L/H	dotp	L/H	dotp	L/H	dotp	L/H	dotp
P84077	ARF1	MGNIFANLFK	ARF	2.34	0.89	0.87	0.89	2.34	0.90	0.62	0.86
P61204	ARF3	MGNIFGNLLK	ARF	0.37	0.97	0.08	0.91	0.17	0.96	0.18	0.94
P18085	ARF4	LGLQSLR	ARF	22.52	0.93	8.57	0.93	9.79	0.93	10.88	0.93
P84085	ARF5	LGLQHLR	ARF	7.88	0.81	3.61	0.82	2.90	0.81	3.97	0.81
P62330	ARF6	FNVWDVGGQDK	ARF	20.47	0.89	6.02	0.90	9.64	0.91	16.19	0.89
Q13795	ARFRP	DCLTQACSALTGK	ARF	0.72	0.97	0.83	0.93	1.41	0.91	0.84	0.85
P40616	ARL1	ILILGLDGAGK	ARF	1.91	0.95	0.31	0.96	1.34	0.95	1.18	0.95
Q8N8L6	ARL10	EVLVLGLDGAGK	ARF	1.87	0.92	0.46	0.87	0.79	0.92	1.02	0.94
Q9NXU5	ARL15	ELGGADNIR	ARF	0.68	0.81	0.54	0.81	0.45	0.82	0.44	0.81
P36404	ARL2	ELQSLLVEER	ARF	14.05	0.97	6.24	0.97	10.36	0.96	10.02	0.96
P36405	ARL3	ILLLGLDNAGK	ARF	7.28	0.89	1.89	0.91	3.89	0.90	8.59	0.89
Q9Y689	ARL5A	AGLLIFANK	ARF	1.09	0.96	0.20	0.96	0.62	0.98	0.75	0.97
Q96KC2	ARL5B	AAVLIFANK	ARF	1.12	0.91	0.17	0.90	0.53	0.89	0.71	0.91
Q9H0F7	ARL6	IPILFFANK	ARF	0.70	0.89	0.12	0.90	0.33	0.89	0.70	0.91
Q96BM9	ARL8A	LWDIGGQPR	ARF	3.90	0.93	1.11	0.93	2.28	0.92	9.17	0.94
Q9NVJ2	ARL8B	IWDIGGQPR	ARF	12.59	0.89	4.09	0.89	8.74	0.89	9.00	0.89
Q9NR31	SAR1A	EIFGLYGQTTGK	ARF	10.50	0.82	5.16	0.83	6.64	0.83	6.75	0.81
Q9Y6B6	SAR1B	EMFGLYGQTTGK	ARF	2.82	0.80	1.19	0.82	1.23	0.82	1.75	0.79
Q9H7X7	IFT22	ILFVGPCESGK	RAB	0.18	0.96	0.13	0.97	0.21	0.98	0.23	0.95
Q9BW83	IFT27	CILAGDPAVGK	RAB	1.67	0.99	1.25	0.97	1.57	0.96	1.31	0.99
P61026	RAB10	FHTITTSYYR	RAB	4.79	0.95	1.06	0.96	4.16	0.97	5.02	0.94
Q6IQ22	RAB12	LQVIIIGSR	RAB	1.12	0.88	0.22	0.79	1.18	0.96	0.75	0.87
P51153	RAB13	LQVWDTAGQER	RAB	0.62	0.98	0.24	0.93	0.83	0.95	1.09	0.95
P61106	RAB14	TGENVEDAFLEAAK	RAB	8.54	0.96	5.39	0.97	11.67	0.97	5.70	0.94
P59190	RAB15	IQIWDTAGQER	RAB	0.94	0.83	0.33	0.89	0.62	0.91	0.72	0.90
Q9NP72	RAB18	TLTPSYYR	RAB	7.77	0.96	6.14	0.92	7.54	0.93	5.77	0.96
P62820	RAB1A	MGPGATAGGAEK	RAB	11.75	0.73	4.34	0.73	6.73	0.73	9.24	0.72
Q9NX57	RAB20	IVLLGDMNVGK	RAB	0.38	0.89	0.15	0.89	0.22	0.89	0.29	0.90
Q9UL25	RAB21	VVLLGEGCVGK	RAB	4.60	0.95	2.47	0.95	4.24	0.95	3.43	0.94
Q9ULC3	RAB23	TIGVDFLER	RAB	4.73	0.96	1.62	0.96	3.48	0.95	4.77	0.98
Q969Q5	RAB24	AQLFETSSK	RAB	0.96	0.94	0.59	0.97	0.73	0.96	0.92	0.99
P51157	RAB28	VAAEILGIK	RAB	0.17	0.94	0.11	0.92	0.18	0.91	0.13	0.94
P61019	RAB2A	MITIDGK	RAB	8.75	0.89	3.34	0.88	5.54	0.86	4.88	0.90
Q8WUD1	RAB2B	IGPQQSISTSVGPSASQR	RAB	1.97	0.92	1.15	0.95	2.72	0.95		
Q15771	RAB30	IVLIGNAGVGK	RAB	1.29	0.95	0.59	0.87	1.04	0.98	0.98	0.87
Q13637	RAB32	VLVIGELGVGK	RAB	0.16	0.80	0.04	0.80	0.05	0.91	0.08	0.83
Q15286	RAB35	LLIIGDSGVGK	RAB	0.48	0.90	0.19	0.91	0.75	0.93	0.81	0.93
Q96AX2	RAB37	VMLLGDTGVGK	RAB	0.15	0.91	0.07	0.84	0.07	0.84	0.63	0.90
P57729	RAB38	LLVIGDLGVGK	RAB	3.19	0.76	1.39	0.82	2.03	0.75		
P20336	RAB3A	TYSWDNAQVLLVGNK	RAB			0.62	0.87	0.97	0.86	0.19	0.85
O95716	RAB3D	LLLIGNSSVGK	RAB	1.03	0.93	0.26	0.93	0.86	0.94	1.00	0.93
Q5JT25	RAB41	LLFLGEQSVGK	RAB	0.05	0.88	0.02	0.84	0.02	0.95		
Q8N4Z0	RAB42	VIFLLVGHK	RAB	0.44	0.90	0.11	0.90	0.26	0.90		

Uniprot Accession	Gene	Peptide	Subfamily	*YTHDF1* $^{-/-}$							
				Biological #1		Biological #2		Biological #3		Biological 4	
				L/H	dotp	L/H	dotp	L/H	dotp	L/H	dotp
Q86YS6	RAB43	VATELIMR	RAB	1.40	0.89	0.38	0.91	0.67	0.90	0.90	0.89
P20338	RAB4A	IESGELDPER	RAB	3.26	0.89	1.93	0.96	2.04	0.95	1.67	0.97
P61018	RAB4B	FLVIGSAGTGK	RAB	0.09	0.92	0.07	0.85	0.09		0.06	0.87
P61020	RAB5B	QASPSIVIALAGNK	RAB	22.40	0.93	11.59	0.94				
P51148	RAB5C	QASPNIVIALAGNK	RAB	35.15	0.92	13.14	0.95	32.61	0.94	41.58	0.92
Q9NRW1	RAB6B	QITIEEGEQR	RAB	4.40	0.92	2.42	0.94	3.11	0.94	2.88	0.90
P51149	RAB7A	FQSLGVAFYR	RAB	48.01	0.95	21.51	0.97	72.74	0.96	29.41	0.95
O14966	RAB7L	VLVVGDAAVGK	RAB	0.17	0.90	0.06	0.93	0.27	0.89	0.18	0.89
P61006	RAB8A	LALDYGIK	RAB	0.67	0.92	0.22	0.90	0.98	0.89	1.04	0.91
Q92930	RAB8B	NIEEHASSDVER	RAB	1.48	0.78	1.25	0.71	1.63	0.83	1.78	0.81
Q9NP90	RAB9B	QSFENLGNWQK	RAB	0.65	0.94	0.13	0.97	0.21	0.94	0.66	0.86
Q5HYI8	RABL3	YLAAGSSNAVK	RAB	1.69	0.97	1.01	0.97	1.14	0.98	1.31	0.96
Q15907	RB11B	NILTEIYR	RAB	33.25	0.99	8.61	0.99	15.89	0.99	27.35	0.98
P51159	RB27A	SWIPEGVVR	RAB	2.24	0.92	0.62	0.94	1.00	0.94	1.81	0.93
Q9H082	RB33B	SAIQVPTDLAQK	RAB	0.68	0.85	0.28	0.81	0.67	0.89	0.63	0.91
Q8WXH6	RB40A	LPLPSTLR	RAB	2.61	0.98	1.45	0.97	2.36	0.99		
Q12829	RB40B	EIDEHAPGVPK	RAB	2.06	0.75	1.58	0.73				
Q9UBK7	RBL2A	IICLGDSAVGK	RAB	0.68	0.96	0.31	0.99	0.72	0.97	0.50	0.95
P62826	RAN	FNVVWDTAGQEK	RAN	108.89	0.98	23.74	0.99	54.28	0.99	57.88	0.98
O95057	DIRA1	EAQAVAQEWK	RAS	0.54	0.93	0.22	0.92	0.40	0.93	0.49	0.92
Q9NYS0	KBRS1	TLIEPFTLLASK	RAS	0.98	0.95	0.94	0.94	1.44	0.95	0.97	0.94
Q9NYR9	KBRS2	VVVCGQASVGK	RAS	1.71	0.87	0.78	0.89	1.02	0.91	1.15	0.88
P55042	RAD	SMPVDER	RAS			0.61	0.72	0.81	0.95	1.49	0.85
P11234	RALB	AEEWGVQYVETSAK	RAS			0.66	0.94	1.42	0.94	0.55	0.94
P62834	RAP1A	QWCNCAFLESSAK	RAS	20.40	0.92	8.69	0.91	13.45	0.94	15.25	0.89
P61224	RAP1B	QWNNCAFLESSAK	RAS	31.37	0.95	21.24	0.96	33.75	0.95		
P10114	RAP2A	VPVILVGNK	RAS	0.09	0.87	0.03	0.83	0.05	0.86	0.03	0.87
P61225	RAP2B	ASVDELFAEIVR	RAS	46.31	0.92	87.44	0.93			19.96	0.93
Q9Y3L5	RAP2C	VPLILVGNK	RAS	0.38	0.83	0.12	0.84	0.17	0.81	0.12	0.82
P01112	RASH	SFEDIHQYR	RAS	0.77	0.86	0.48	0.88	0.58	0.87	0.34	0.84
P01111	RASN	SFADINLYR	RAS	10.76	0.99	4.00	0.98	7.94	0.98	9.75	0.99
Q15382	RHEB	ALAESWNAAFLESSAK	RAS	14.40	0.96			9.07	0.96	5.70	0.95
P10301	RRAS	LFTQILR	RAS	0.22	0.86	0.09	0.94	0.16	0.90	0.24	0.91
P62070	RRAS2	QVTQEEGQQLAR	RAS	8.20	0.88	7.06	0.92	9.60	0.93	6.40	0.88
Q9BPW5	RSLBB	TSLIPRPK	RAS	0.41	0.86			0.17	0.75	0.42	0.75
P60953	CDC42	TCLLISYTTNK	RHO	14.20	0.98	6.99	0.98	13.22	0.99	4.58	0.96
P60763	RAC3	LAPITYPQGLAMAR	RHO	11.75	0.97	5.59	0.97	8.43	0.96	18.45	0.97
P61586	RHOA	IGAFGYMECSAK	RHO	40.59	0.97	23.52	0.97	22.30	0.97	29.37	0.97
P62745	RHOB	IQAYDYLECSAK	RHO	0.82	0.94	0.50	0.92	0.91	0.96	0.76	0.92
P08134	RHOC	ISAFGYLECSAK	RHO	5.40	0.95	5.88	0.97	7.71	0.97	5.25	0.94
Q9HBH0	RHOF	AALYLECSAK	RHO	0.23	0.90	0.07	0.95	0.25	0.84	0.42	0.80
P84095	RHOG	YLECSALQQDGVK	RHO	2.23	0.91	0.97	0.94	1.23	0.94	1.15	0.96
P52198	RND2	IVVVGDAECGK	RHO	0.11	0.99	0.08	0.97	0.10	0.92	0.05	0.82

Uniprot Accession	Gene	Peptide	Subfamily	*YTHDF2* $^{-/-}$							
				Biological #1		Biological #2		Biological #3		Biological 4	
				L/H	dotp	L/H	dotp	L/H	dotp	L/H	dotp
P84077	ARF1	MGNIFANLFK	ARF	3.32	0.89	1.06	0.90	3.81	0.91	0.53	0.86
P61204	ARF3	MGNIFGNLLK	ARF	0.52	0.96	0.09	0.96	0.26	0.91	0.22	0.87
P18085	ARF4	LGLQSLR	ARF	22.19	0.93	8.95	0.94	11.64	0.93	9.80	0.93
P84085	ARF5	LGLQHLR	ARF	9.65	0.81	3.40	0.82	3.24	0.81	3.78	0.80
P62330	ARF6	FNVWDVGGQDK	ARF	25.99	0.88	7.72	0.90	14.22	0.90	15.75	0.91
Q13795	ARFRP	DCLTQACSALTGK	ARF	1.54	0.87	1.38	0.92	1.70	0.91	1.23	0.88
P40616	ARL1	ILILGLDGAGK	ARF	2.40	0.95	0.42	0.96	1.73	0.96	1.40	0.96
Q8N8L6	ARL10	EVLVLGLDGAGK	ARF	1.97	0.94	0.42	0.91	0.87	0.92	0.75	0.90
Q9NXU5	ARL15	ELGGADNIR	ARF	0.68	0.80	0.53	0.83	0.47	0.82	0.40	0.80
P36404	ARL2	ELQSLLVEER	ARF	18.94	0.96	7.78	0.96	11.61	0.96	13.90	0.97
P36405	ARL3	ILLLGLDNAGK	ARF	10.55	0.90	2.70	0.91	4.64	0.90	8.82	0.90
Q9Y689	ARL5A	AGLLIFANK	ARF	1.29	0.95	0.23	0.94	0.71	1.00	0.68	0.97
Q96KC2	ARL5B	AAVLIFANK	ARF	1.17	0.90	0.20	0.88	0.71	0.89	0.79	0.90
Q9H0F7	ARL6	IPILFFANK	ARF	0.97	0.90	0.16	0.90	0.36	0.89	0.79	0.89
Q96BM9	ARL8A	LWDIGGQPR	ARF	3.95	0.92	1.48	0.92	1.97	0.94	10.19	0.94
Q9NVJ2	ARL8B	IWDIGGQPR	ARF	11.35	0.89	4.83	0.90	8.15	0.89	11.94	0.89
Q9NR31	SAR1A	EIFGLYGQTTGK	ARF	14.53	0.82	6.61	0.84	7.65	0.83	6.42	0.84
Q9Y6B6	SAR1B	EMFGLYGQTTGK	ARF	2.98	0.81	1.44	0.82	1.55	0.82	1.91	0.82
Q9H7X7	IFT22	ILFVGPCESGK	RAB	0.28	0.95	0.16	0.97	0.22	0.99	0.22	0.96
Q9BW83	IFT27	CILAGDPAVGK	RAB	1.77	0.99	1.40	0.97	1.60	0.98	1.28	0.97
P61026	RAB10	FHTITTSYYR	RAB	5.86	0.95	1.39	0.97	4.78	0.97	5.04	0.97
Q6IQ22	RAB12	LQVIIIGSR	RAB	1.22	0.82	0.26	0.88	1.14	0.93	1.20	0.91
P51153	RAB13	LQVWDTAGQER	RAB	0.67	0.92	0.34	0.97	0.98	0.96	1.20	0.97
P61106	RAB14	TGENVEDAFLEAAK	RAB	9.53	0.94	6.86	0.98	14.06	0.97	5.81	0.97
P59190	RAB15	IQIWDTAGQER	RAB	0.87	0.87	0.21	0.86			1.07	0.87
Q9NP72	RAB18	TLTPSYYR	RAB	7.79	0.96	7.47	0.93	8.19	0.95	6.11	0.94
P62820	RAB1A	MGPGATAGGAEK	RAB	11.11	0.72	4.29	0.74	7.93	0.73	8.74	0.72
Q9NX57	RAB20	IVLLGDMNVGK	RAB	0.43	0.87	0.10	0.89	0.30	0.89	0.33	0.90
Q9UL25	RAB21	VVLLGEGCVGK	RAB	5.76	0.95	2.77	0.95	4.72	0.95	2.97	0.96
Q9ULC3	RAB23	TIGVDFLER	RAB	6.35	0.96	2.02	0.97	3.97	0.95	4.98	0.96
Q969Q5	RAB24	AQLFETSSK	RAB	0.94	0.99	0.55	0.95	0.76	0.97	0.76	0.95
P51157	RAB28	VAAEILGIK	RAB	0.22	0.91	0.13	0.89	0.19	0.92	0.15	0.94
P61019	RAB2A	MITIDGK	RAB	10.06	0.89	3.41	0.89	7.61	0.87	4.57	0.89
Q8WUD1	RAB2B	IGPQQSISTSVGPSASQR	RAB	1.81	0.92	1.26	0.96	2.32	0.95		
Q15771	RAB30	IVLIGNAGVGK	RAB	1.32	0.94	0.52	0.89	1.00	0.95	0.96	0.88
Q13637	RAB32	VLVIGELGVGK	RAB	0.19	0.85	0.06	0.91	0.14	0.92	0.06	0.87
Q15286	RAB35	LLIIGDSGVGK	RAB	0.49	0.92	0.27	0.92	0.85	0.93	0.82	0.93
Q96AX2	RAB37	VMLLGDTGVGK	RAB	0.13	0.76	0.10	0.81	0.08	0.79	0.55	0.97
P57729	RAB38	LLVIGDLGVGK	RAB	4.43	0.72	1.30	0.78	2.02	0.77		
P20336	RAB3A	TYSWDNAQVLLVGNK	RAB			0.49	0.88	1.03	0.80	0.20	0.89
O95716	RAB3D	LLLIGNSSVGK	RAB	1.21	0.93	0.31	0.94	0.93	0.93	0.80	0.95
Q5JT25	RAB41	LLFLGEQSVGK	RAB	0.03	0.84	0.03	0.92	0.09	0.91		
Q8N4Z0	RAB42	VIFLLVGHK	RAB	0.43	0.81	0.09	0.83	0.15	0.81		

Uniprot Accession	Gene	Peptide	Subfamily	*YTHDF2* $^{-/-}$							
				Biological #1		Biological #2		Biological #3		Biological 4	
				L/H	dotp	L/H	dotp	L/H	dotp	L/H	dotp
Q86YS6	RAB43	VATELIMR	RAB	1.52	0.87	0.38	0.91	0.62	0.91	0.86	0.89
P20338	RAB4A	IESGELDPER	RAB	3.49	0.93	2.18	0.95	2.41	0.94	1.72	0.98
P61018	RAB4B	FLVIGSAGTGK	RAB	0.18	0.93	0.12	0.79			0.08	0.95
P61020	RAB5B	QASPSIVIALAGNK	RAB	26.90	0.92	13.90	0.95				
P51148	RAB5C	QASPNIVIALAGNK	RAB	42.28	0.92	16.41	0.95	30.82	0.94	35.85	0.95
Q9NRW1	RAB6B	QITIEEGEQR	RAB	3.66	0.91	2.28	0.96	2.71	0.92	1.90	0.92
P51149	RAB7A	FQSLGVAFYR	RAB	58.42	0.95	25.58	0.97	75.82	0.96	26.86	0.96
O14966	RAB7L	VLVVGDAAVGK	RAB	0.18	0.88	0.09	0.85	0.29	0.88	0.18	0.89
P61006	RAB8A	LALDYGIK	RAB	0.58	0.91	0.26	0.89	1.02	0.90	0.94	0.93
Q92930	RAB8B	NIEEHASSDVER	RAB	1.22	0.78			2.05	0.86	1.76	0.84
Q9NP90	RAB9B	QSFENLGNWQK	RAB	0.58	0.88	0.22	0.94	0.30	0.91	0.40	0.84
Q5HYI8	RABL3	YLAAGSSNAVK	RAB	1.87	0.97	1.32	0.97	1.26	0.97	1.18	0.97
Q15907	RB11B	NILTEIYR	RAB	40.12	0.99	10.34	0.98	16.55	0.98	24.39	0.98
P51159	RB27A	SWIPEGVVR	RAB	3.61	0.94	0.91	0.92	1.43	0.91	1.97	0.94
Q9H082	RB33B	SAIQVPTDLAQK	RAB	0.61	0.86	0.43	0.75	0.76	0.85	0.60	0.91
Q8WXH6	RB40A	LPLPSTLR	RAB	2.34	0.95	0.65	0.96	1.79	0.98		
Q12829	RB40B	EIDEHAPGVPK	RAB			1.40	0.70				
Q9UBK7	RBL2A	IICLGDSAVGK	RAB	0.68	0.97	0.39	0.99	0.72	0.96	0.52	0.98
P62826	RAN	FNVWDTAGQEK	RAN	121.66	0.98	30.76	0.99	76.55	0.99	58.11	0.99
O95057	DIRA1	EAQAVAQEWK	RAS	0.58	0.91	0.24	0.95	0.42	0.95	0.47	0.93
Q9NYS0	KBRS1	TLIEPFTLLASK	RAS	1.68	0.95	1.24	0.94	1.80	0.94	1.22	0.95
Q9NYR9	KBRS2	VVVCGQASVGK	RAS	1.95	0.88	0.92	0.89	1.29	0.90	1.13	0.91
P55042	RAD	SMPVDER	RAS			0.67	0.82	1.18	0.88		
P11234	RALB	AEEWGVQYVETSAK	RAS			0.64	0.97	1.29	0.94	0.51	0.94
P62834	RAP1A	QWCNCAFLESSAK	RAS	12.69	0.92	9.62	0.92	9.36	0.93	14.19	0.93
P61224	RAP1B	QWNNCAFLESSAK	RAS	38.35	0.94	25.62	0.94	34.65	0.95		
P10114	RAP2A	VPVILVGNK	RAS	0.10	0.86	0.04	0.86	0.05	0.87	0.02	0.87
P61225	RAP2B	ASVDELFAEIVR	RAS	56.13	0.92	55.28	0.92			18.37	0.93
Q9Y3L5	RAP2C	VPLILVGNK	RAS	0.48	0.82	0.15	0.84	0.22	0.83	0.13	0.84
P01112	RASH	SFEDIHQYR	RAS	0.66	0.85	0.49	0.89	0.55	0.86	0.35	0.86
P01111	RASN	SFADINLYR	RAS	13.43	0.99	4.96	0.99	8.32	0.98	8.72	0.99
Q15382	RHEB	ALAESWNAAFLESSAK	RAS	8.86	0.97			16.67	0.97	6.29	0.97
P10301	RRAS	LFTQILR	RAS	0.44	0.91	0.11	0.91	0.15	0.91	0.20	0.85
P62070	RRAS2	QVTQEEGQQLAR	RAS	8.47	0.89	8.66	0.92	9.83	0.91	7.29	0.91
Q9BPW5	RSLBB	TSLIPRPK	RAS	0.41	0.87	0.07	0.70	0.26	0.86	0.39	0.74
P60953	CDC42	TCLLISYTTNK	RHO	26.68	0.98	6.56	0.99	13.87	0.98	4.26	0.95
P60763	RAC3	LAPITYPQGLAMAR	RHO	14.56	0.96	5.20	0.96	7.34	0.96	15.70	0.97
P61586	RHOA	IGAFGYMECSAK	RHO	58.08	0.97	21.04	0.98	31.21	0.97	33.16	0.97
P62745	RHOB	IQAYDYLECSAK	RHO	1.30	0.95	0.58	0.90	0.89	0.93	0.85	0.90
P08134	RHOC	ISAFGYLECSAK	RHO	6.88	0.95	5.30	0.96	8.54	0.95	4.21	0.97
Q9HBH0	RHOF	AALYLECSAK	RHO	0.16	0.95	0.07	0.96	0.18	0.89	0.40	0.77
P84095	RHOG	YLECSALQQDGVK	RHO	2.54	0.92	1.14	0.89	1.43	0.93	1.15	0.92
P52198	RND2	IVVVGDAECGK	RHO	0.10	0.93	0.07	0.93	0.10	0.97	0.08	0.91

Uniprot Accession	Gene	Peptide	Subfamily	*YTHDF3* $^{-/-}$							
				Biological #1		Biological #2		Biological #3		Biological 4	
				L/H	dotp	L/H	dotp	L/H	dotp	L/H	dotp
P84077	ARF1	MGNIFANLFK	ARF	3.65	0.89	0.85	0.89	3.01	0.90	0.36	0.83
P61204	ARF3	MGNIFGNLLK	ARF	0.77	0.95	0.07	0.95	0.22	0.97	0.13	0.93
P18085	ARF4	LGLQSLR	ARF	25.79	0.93	8.22	0.94	12.21	0.93	5.79	0.93
P84085	ARF5	LGLQHLR	ARF	9.10	0.80	3.08	0.82	3.57	0.81	3.03	0.82
P62330	ARF6	FNVWDVGGQDK	ARF	30.57	0.88	6.61	0.90	17.00	0.90	9.68	0.91
Q13795	ARFRP	DCLTQACSALTGK	ARF	1.30	0.86	0.80	0.93	1.55	0.90	1.21	0.90
P40616	ARL1	ILILGLDGAGK	ARF	2.96	0.95	0.30	0.96	1.79	0.96	0.93	0.96
Q8N8L6	ARL10	EVLVLGLDGAGK	ARF	2.54	0.93	0.30	0.76	0.93	0.92	0.54	0.95
Q9NXU5	ARL15	ELGGADNIR	ARF	0.97	0.80	0.40	0.82	0.56	0.81	0.31	0.81
P36404	ARL2	ELQSLLVEER	ARF	15.31	0.97	5.35	0.97	10.47	0.96	6.52	0.96
P36405	ARL3	ILLLGLDNAGK	ARF	15.08	0.89	2.42	0.91	5.16	0.90	7.29	0.91
Q9Y689	ARL5A	AGLLIFANK	ARF	1.64	0.96	0.22	0.97	0.83	0.99	0.52	0.96
Q96KC2	ARL5B	AAVLIFANK	ARF	1.63	0.90	0.20	0.88	0.68	0.90	0.54	0.87
Q9H0F7	ARL6	IPILFFANK	ARF	1.02	0.89	0.12	0.91	0.37	0.89	0.52	0.89
Q96BM9	ARL8A	LWDIGGQPR	ARF	5.27	0.91	1.13	0.93	2.53	0.91	9.03	0.94
Q9NVJ2	ARL8B	IWDIGGQPR	ARF	14.94	0.89	4.17	0.90	10.35	0.90	9.05	0.90
Q9NR31	SAR1A	EIFGLYGQTTGK	ARF	15.35	0.81	5.07	0.84	7.68	0.82	2.99	0.84
Q9Y6B6	SAR1B	EMFGLYGQTTGK	ARF	3.05	0.79	1.10	0.82	1.45	0.81	1.26	0.82
Q9H7X7	IFT22	ILFVGPCESGK	RAB	0.33	0.95	0.14	0.97	0.24	0.97	0.14	0.99
Q9BW83	IFT27	CILAGDPAVGK	RAB	2.46	0.99	1.03	0.98	1.57	0.98	0.85	0.98
P61026	RAB10	FHTITTSYYR	RAB	6.37	0.94	0.93	0.97	4.80	0.95	3.56	0.95
Q6IQ22	RAB12	LQVIIIGSR	RAB	1.35	0.86	0.24	0.86	1.08	0.91	0.81	0.94
P51153	RAB13	LQVWDTAGQER	RAB	1.07	0.94	0.28	0.96	1.38	0.95	1.42	0.97
P61106	RAB14	TGENVEDAFLEAAK	RAB	11.30	0.94	4.83	0.97	14.58	0.97	5.06	0.97
P59190	RAB15	IQIWDTAGQER	RAB	0.87	0.70	0.22	0.87	0.65	0.86	0.71	0.86
Q9NP72	RAB18	TLTPSYYR	RAB	8.63	0.96	4.99	0.92	6.98	0.94	4.00	0.94
P62820	RAB1A	MGPGATAGGAEK	RAB	19.29	0.72	3.92	0.73	8.23	0.73	9.29	0.73
Q9NX57	RAB20	IVLLGDMNVGK	RAB	0.39	0.90	0.08	0.91	0.26	0.88	0.17	0.91
Q9UL25	RAB21	VVLLGEGCVGK	RAB	6.76	0.95	2.08	0.96	5.21	0.95	1.88	0.96
Q9ULC3	RAB23	TIGVDFLER	RAB	4.49	0.97	0.92	0.94	2.47	0.94	2.22	0.96
Q969Q5	RAB24	AQLFETSSK	RAB	1.27	0.96	0.45	0.94	0.85	0.98	0.66	0.95
P51157	RAB28	VAAEILGIK	RAB	0.27	0.92	0.12	0.94	0.23	0.89	0.12	0.93
P61019	RAB2A	MITIDGK	RAB	20.46	0.89	3.83	0.89	9.56	0.88	5.46	0.89
Q8WUD1	RAB2B	IGPQQSISTSVGPSASQR	RAB	3.50	0.89	1.03	0.90	3.12	0.94		
Q15771	RAB30	IVLIGNAGVGK	RAB	1.82	0.94	0.47	0.86	1.25	0.97	0.69	0.91
Q13637	RAB32	VLVIGELGVGK	RAB	0.07	0.89	0.07	0.82	0.09	0.87	0.04	0.78
Q15286	RAB35	LLIIGDSGVGK	RAB	0.76	0.92	0.18	0.91	0.88	0.94	0.75	0.94
Q96AX2	RAB37	VMLLGDTGVGK	RAB	0.16	0.76	0.08	0.76	0.11	0.81	0.37	0.93
P57729	RAB38	LLVIGDLGVGK	RAB			0.85	0.77	1.79	0.73		
P20336	RAB3A	TYSWDNAQVLLVGNK	RAB			0.71	0.90	0.80	0.77	0.13	0.97
O95716	RAB3D	LLLIGNSSVGK	RAB	1.37	0.94	0.27	0.95	0.96	0.93	0.64	0.94
Q5JT25	RAB41	LLFLGEQSVGK	RAB	0.04	0.88	0.01	0.96	0.03	0.89		
Q8N4Z0	RAB42	VIFLLVGHK	RAB	0.65	0.90	0.12	0.83	0.37	0.92		

Uniprot Accession	Gene	Peptide	Subfamily	*YTHDF3* $^{-/-}$							
				Biological #1		Biological #2		Biological #3		Biological 4	
				L/H	dotp	L/H	dotp	L/H	dotp	L/H	dotp
Q86YS6	RAB43	VATELIMR	RAB	1.86	0.89	0.33	0.91	0.68	0.92	0.60	0.90
P20338	RAB4A	IESGELDPER	RAB	5.03	0.93	1.95	0.93	2.54	0.95	1.28	0.99
P61018	RAB4B	FLVIGSAGTGK	RAB	0.16	0.94	0.09	0.81			0.04	0.92
P61020	RAB5B	QASPSIVIALAGNK	RAB	29.10	0.93	20.77	0.93				
P51148	RAB5C	QASPNIVIALAGNK	RAB	61.06	0.92	12.04	0.95	32.94	0.94	26.52	0.95
Q9NRW1	RAB6B	QITIEEGEQR	RAB	4.45	0.90	1.70	0.97	3.63	0.92	1.42	0.93
P51149	RAB7A	FQSLGVAFYR	RAB	60.29	0.94	17.46	0.97	73.22	0.96	18.93	0.96
O14966	RAB7L	VLVVGDAAVGK	RAB	0.27	0.87	0.07	0.82	0.29	0.85	0.19	0.86
P61006	RAB8A	LALDYGIK	RAB	0.87	0.93	0.21	0.93	1.28	0.89	0.87	0.91
Q92930	RAB8B	NIEEHASSDVER	RAB	2.08	0.80	0.41	0.50	2.12	0.86	1.81	0.83
Q9NP90	RAB9B	QSFENLGNWQK	RAB	0.86	0.86	0.20	0.84	0.38	0.89	0.32	0.77
Q5HYI8	RABL3	YLAAGSSNAVK	RAB	2.48	0.96	0.94	0.97	1.09	0.95	1.07	0.97
Q15907	RB11B	NILTEIYR	RAB	60.36	0.98	10.69	0.99	21.22	0.98	17.53	0.99
P51159	RB27A	SWIPEGVVR	RAB	3.86	0.92	0.63	0.92	1.27	0.93	1.24	0.94
Q9H082	RB33B	SAIQVPTDLAQK	RAB	0.85	0.88	0.33	0.77	0.71	0.92	0.52	0.94
Q8WXH6	RB40A	LPLPSTLR	RAB	2.31	0.99	0.33	0.90	1.81	0.99		
Q12829	RB40B	EIDEHAPGVPK	RAB			1.21	0.68				
Q9UBK7	RBL2A	IICLGDSAVGK	RAB	1.08	0.95	0.29	0.99	0.80	0.96	0.49	0.95
P62826	RAN	FNVVWDTAGQEK	RAN	140.16	0.97	27.21	0.99	77.75	0.99	37.85	1.00
O95057	DIRA1	EAQAVAQEWK	RAS	0.78	0.92	0.21	0.95	0.42	0.91	0.44	0.91
Q9NYS0	KBRS1	TLIEPFTLLASK	RAS	1.61	0.95	0.83	0.95	1.49	0.93	0.67	0.94
Q9NYR9	KBRS2	VVVCGQASVGK	RAS	2.67	0.88	0.76	0.88	1.35	0.90	0.61	0.92
P55042	RAD	SMPVDER	RAS			0.59	0.86	1.30	0.88	0.87	0.77
P11234	RALB	AEEWGVQYVETSAK	RAS			0.43	0.88	1.18	0.92	0.42	0.91
P62834	RAP1A	QWCNCAFLESSAK	RAS	21.99	0.91	10.08	0.93	10.77	0.92	8.26	0.92
P61224	RAP1B	QWNNCAFLESSAK	RAS	60.09	0.94	21.22	0.94	30.08	0.94		
P10114	RAP2A	VPVILVGNK	RAS	0.15	0.85	0.04	0.84	0.07	0.86	0.02	0.86
P61225	RAP2B	ASVDELFAEIVR	RAS	57.37	0.91	153.54	0.92			15.83	0.92
Q9Y3L5	RAP2C	VPLILVGNK	RAS	0.61	0.83	0.12	0.84	0.17	0.86	0.06	0.89
P01112	RASH	SFEDIHQYR	RAS	0.66	0.84	0.45	0.87	0.54	0.88	0.36	0.86
P01111	RASN	SFADINLYR	RAS	15.12	0.99	3.17	0.99	7.55	0.99	5.25	0.99
Q15382	RHEB	ALAESWNAAFLESSAK	RAS	16.14	0.96			8.77	0.97	4.49	0.98
P10301	RRAS	LFTQILR	RAS	0.61	0.91	0.14	0.93	0.25	0.93	0.15	0.92
P62070	RRAS2	QVTQEEGQQLAR	RAS	18.48	0.88	6.58	0.93	11.30	0.92	5.92	0.92
Q9BPW5	RSLBB	TSLIPRPK	RAS	0.42	0.77	0.13	0.76	0.18	0.72	0.41	0.76
P60953	CDC42	TCLLISYTTNK	RHO	22.14	0.97	5.73	0.98	18.97	0.98	1.16	0.94
P60763	RAC3	LAPITYPQGLAMAR	RHO	26.24	0.96	5.01	0.97	9.04	0.98	14.03	0.97
P61586	RHOA	IGAFGYMECSAK	RHO	51.60	0.97	17.58	0.97	26.53	0.97	22.76	0.97
P62745	RHOB	IQAYDYLECSAK	RHO	1.43	0.95	0.40	0.90	0.90	0.93	0.44	0.89
P08134	RHOC	ISAFGYLECSAK	RHO	7.65	0.94	3.61	0.93	9.06	0.96	2.82	0.96
Q9HBH0	RHOF	AALYLECSAK	RHO	0.12	0.93	0.04	0.87	0.20	0.81	0.30	0.77
P84095	RHOG	YLECSALQQDGVK	RHO	3.56	0.90	1.07	0.93	1.76	0.93	0.75	0.92
P52198	RND2	IVVVGDAECGK	RHO	0.13	0.97	0.06	0.92	0.10	0.90	0.03	0.89

Table S4.2B A complete list of all small GTPases quantified in the scheduled LC-MRM analysis in cells depleted of METTL3.

Uniprot Accession	Gene	Peptide	Subfamily	HEK293T-WT						*METTL3* $^{-/-}$					
				Biological #1		Biological #2		Biological #3		Biological #1		Biological #2		Biological #3	
				L/H	dotp	L/H	dotp	L/H	dotp	L/H	dotp	L/H	dotp	L/H	dotp
P61204	ARF3	MGNIFGNLLK	ARF	0.03	0.86	0.16	0.87	0.10	0.85	0.03	0.69	0.05	0.81	0.04	0.80
P18085	ARF4	LGLQSLR	ARF	13.29	0.94	14.34	0.94	16.93	0.94	12.47	0.94	11.78	0.94	13.04	0.94
P84085	ARF5	LGLQHLR	ARF	8.68	0.84	6.71	0.84	7.85	0.84	7.86	0.84	7.12	0.84	6.08	0.83
P62330	ARF6	FNVWDVGGQDK	ARF	6.70	0.93	13.53	0.93	15.16	0.92	16.89	0.92	35.92	0.93	34.31	0.93
Q13795	ARFRP	DCLTQACSALTGK	ARF	1.44	0.92	2.07	0.92	2.00	0.95	1.03	0.90	1.73	0.93	1.94	0.95
P40616	ARL1	ILILGLDGAGK	ARF	3.35	0.94	2.56	0.92	3.25	0.93	1.77	0.94	1.78	0.92	2.55	0.94
Q8N8L6	ARL10	EVLVLGLDGAGK	ARF	2.61	0.81	9.28	0.77	7.05	0.78	2.77	0.83	14.78	0.80	8.67	0.81
Q969Q4	ARL11	QEAPDALPLLK	ARF	0.12	0.50	0.14	0.32	0.08	0.35	0.13	0.50	0.13	0.47	0.11	0.35
Q9NXU5	ARL15	ELGGADNIR	ARF	0.62	0.89	0.74	0.84	0.53	0.86	0.81	0.93	0.72	0.90	0.60	0.90
P36404	ARL2	ELQSLLVEER	ARF	6.61	0.97	11.15	0.97	5.25	0.98	4.97	0.96	6.79	0.97	6.09	0.99
P36405	ARL3	ILLLGLDNAGK	ARF	2.42	0.95	4.86	0.94	3.11	0.95	1.61	0.94	2.80	0.94	2.48	0.93
P56559	ARL4C	FAENQGTPLLVIANK	ARF	0.08	0.85	0.17	0.75	0.13	0.81	0.07	0.75	0.14	0.90	0.12	0.88
Q9Y689	ARL5A	AGLLIFANK	ARF	0.42	0.98	0.26	0.95	0.43	0.99	0.36	0.96	0.20	0.91	0.35	0.93
Q96KC2	ARL5B	AAVLIFANK	ARF	0.37	0.90	0.24	0.89	0.25	0.95	0.35	0.92	0.23	0.89	0.22	0.90
Q9H0F7	ARL6	IPILFFANK	ARF	0.19	0.92	0.28	0.92	0.22	0.92	0.17	0.92	0.22	0.91	0.24	0.91
Q96BM9	ARL8A	LWDIGGQPR	ARF	1.92	0.93	6.03	0.95	6.03	0.94	2.09	0.93	4.66	0.95	5.07	0.94
Q9NVJ2	ARL8B	IWDIGGQPR	ARF	1.76	0.89	5.60	0.91	6.37	0.92	2.03	0.90	4.76	0.91	4.76	0.91
Q9NR31	SAR1A	EIFGLYGQTTGK	ARF	4.79	0.88	11.39	0.87	10.18	0.86	5.98	0.88	13.35	0.87	10.38	0.87
Q9Y6B6	SAR1B	EMFGLYGQTTGK	ARF	1.30	0.86	3.51	0.86	3.32	0.87	1.14	0.87	2.39	0.86	2.16	0.86
Q9H7X7	IFT22	ILFVGPCESGK	RAB	0.18	0.97	0.16	0.94	0.10	0.93	0.14	0.99	0.12	0.96	0.09	0.93
Q9BW83	IFT27	CILAGDPAVGK	RAB	0.79	0.90	1.14	0.88	0.81	0.89	0.42	0.86	0.48	0.86	0.44	0.89
P61026	RAB10	FHTITTSYYR	RAB	9.52	0.99	3.68	0.99	7.32	0.99	5.95	0.99	2.80	0.97	5.69	0.98
Q6IQ22	RAB12	LQVIIIGSR	RAB	0.63	0.97	0.62	0.92	1.12	0.98	0.68	0.96	0.84	0.94	0.81	0.96
P51153	RAB13	LQVWDTAGQER	RAB	3.55	0.94	1.96	0.95	2.26	0.95	1.51	0.94	1.03	0.94	1.67	0.94
P61106	RAB14	TGENVEDAFLEAAK	RAB	8.64	0.95	35.03	0.95	39.28	0.95	9.00	0.95	28.45	0.96	38.84	0.96
P59190	RAB15	IQIWDTAGQER	RAB	62.48	0.97	60.43	0.97	68.05	0.97	50.66	0.97	57.39	0.97	70.02	0.96
Q9NP72	RAB18	TLTPSYYR	RAB	0.88	0.88	0.80	0.90	0.81	0.91	1.07	0.91	0.99	0.92	0.98	0.92
P62820	RAB1A	MGPGATAGGAEK	RAB	11.06	0.76	8.16	0.76	10.69	0.77	11.25	0.77	7.25	0.77	9.09	0.76
Q9H0U4	RAB1B	MGPGAASGGER	RAB	0.04	0.86	0.11	0.76	0.12	0.78	0.05	0.74	0.04	0.66	0.06	0.76
Q9NX57	RAB20	IVLLGDMNVGK	RAB	0.06	0.93	0.40	0.92	0.27	0.99	0.09	0.91	0.23	0.85	0.27	0.95
Q9UL25	RAB21	VVLLGEGCVGK	RAB	3.08	0.95	4.04	0.96	3.92	0.96	3.16	0.96	3.29	0.97	3.66	0.96
Q9ULC3	RAB23	TIGVDFLER	RAB	1.40	0.93	4.86	0.91	3.04	0.92	1.27	0.93	2.55	0.93	2.31	0.93
Q969Q5	RAB24	AQLFETSSK	RAB	0.82	0.88	0.80	0.95	0.73	0.92	0.67	0.89	0.52	0.91	0.59	0.95
Q9ULW5	RAB26	VMLVGDSGVGK	RAB	0.24	0.86	0.58	0.71	0.28	0.79	0.28	0.82	0.38	0.79	0.32	0.85
P51157	RAB28	VAAEILGIK	RAB	0.14	0.93	0.10	0.92	0.09	0.88	0.16	0.91	0.13	0.88	0.11	0.89
P61019	RAB2A	MITIDGK	RAB	8.46	0.79	8.12	0.79	9.14	0.80	8.75	0.79	7.32	0.80	8.88	0.81
Q8WUD1	RAB2B	IGPQQSISTSVGPSASQR	RAB	2.27	0.97	2.85	0.93	2.78	0.94	2.54	0.97	2.84	0.93	2.70	0.95
Q15771	RAB30	IVLIGNAGVGK	RAB	0.62	0.91	0.59	0.84	0.61	0.85	1.01	0.89	0.70	0.90	0.75	0.93
Q9BZG1	RAB34	AEYWAVSSLTGENVR	RAB	2.13	0.90	8.96	0.81	7.10	0.91	2.46	0.88	7.49	0.87	9.01	0.91
Q15286	RAB35	LLIIGDSGVGK	RAB	1.69	0.94	0.61	0.96	0.99	0.95	1.24	0.96	0.59	0.94	1.28	0.92
Q96AX2	RAB37	VMLLGDTGVGK	RAB	0.10	0.74	0.08	0.73	0.08	0.80	0.07	0.80	0.14	0.74	0.17	0.72
P57729	RAB38	LLVIGDLGVGK	RAB	0.60	0.79	5.07	0.70	4.26	0.75	0.53	0.80	4.13	0.71	3.10	0.74
P20336	RAB3A	TYSWDNAQVLLVGNK	RAB	1.51	0.87	0.61	0.97	2.08	0.94	0.63	0.78	0.50	0.95	0.90	0.92
O95716	RAB3D	LLLIGNSSVGK	RAB	0.98	0.96	0.71	0.96	0.85	0.93	1.15	0.96	0.81	0.95	1.29	0.96
Q86YS6	RAB43	VATELIMR	RAB	0.48	0.93	0.67	0.96	0.44	0.92	0.66	0.95	0.78	0.91	0.70	0.94
P20338	RAB4A	IESGELDPER	RAB	5.14	0.86	1.30	0.91	3.25	0.88	4.63	0.87	1.08	0.90	2.89	0.86

Uniprot Accession	Gene	Peptide	Subfamily	HEK293T-WT						*METTL3* $^{-/-}$					
				Biological #1		Biological #2		Biological #3		Biological #1		Biological #2		Biological #3	
				L/H	dotp	L/H	dotp	L/H	dotp	L/H	dotp	L/H	dotp	L/H	dotp
P61020	RAB5B	QASPSIVIALAGNK	RAB	8.83	0.99	21.31	0.98	21.74	0.98	13.38	0.97	20.34	0.98	44.02	0.98
P51148	RAB5C	QASPNIVIALAGNK	RAB	9.31	0.98	25.51	0.98	18.98	0.97	7.18	0.98	12.14	0.98	15.68	0.99
Q9NRW1	RAB6B	QITIEEGEQR	RAB	3.65	0.98	3.86	0.97	2.64	0.97	3.66	0.99	4.11	0.96	3.74	0.97
P51149	RAB7A	FQSLGVAFYR	RAB	12.83	0.99	38.50	0.99	33.44	0.99	10.47	0.99	28.77	0.99	29.53	0.99
O14966	RAB7L	VLVVGDAAVGK	RAB	0.82	0.90	0.30	0.91	0.43	0.96	0.46	0.89	0.24	0.94	0.37	0.99
P61006	RAB8A	LALDYGIK	RAB	2.76	0.85	0.74	0.87	1.30	0.87	1.76	0.86	0.67	0.86	1.54	0.86
Q92930	RAB8B	NIEEHASSDVER	RAB	4.56	0.82	5.23	0.64	2.87	0.83	8.36	0.74	1.83	0.80	2.32	0.93
P51151	RAB9A	DATNVAAAFEEAVR	RAB	4.55	0.98	15.49	0.98	14.55	0.99	4.35	0.99	27.44	0.98	20.04	0.98
Q9NP90	RAB9B	QSFENLGNWQK	RAB	0.29	0.86	0.31	0.89	0.34	0.96	0.44	0.90	0.39	0.92	0.41	0.87
Q5HYI8	RABL3	YLAAGSSNAVK	RAB	1.28	0.97	1.17	0.97	0.94	0.97	2.09	0.97	1.57	0.97	1.45	0.98
Q15907	RB11B	NILTEIYR	RAB	9.64	0.97	18.12	0.97	13.39	0.98	10.44	0.96	17.89	0.97	15.99	0.98
P51159	RB27A	SWIPEGVVR	RAB	1.12	0.87	2.05	0.91	1.72	0.89	1.56	0.91	2.67	0.88	2.22	0.90
O00194	RB27B	LLALGDSGVGK	RAB	0.96	0.52	0.78	0.62	0.75	0.61	0.69	0.52	0.72	0.62	0.56	0.63
Q9H082	RB33B	SAIQVPTDLAQK	RAB	0.93	0.68	1.62	0.51	1.09	0.61	1.05	0.75	1.30	0.60	1.27	0.64
Q9UBK7	RBL2A	IICLGDSAVGK	RAB	0.77	0.85	0.75	0.83	0.75	0.84	0.55	0.82	0.39	0.79	0.50	0.82
Q9UNT1	RBL2B	IDDINVTQK	RAB	0.24	0.87	0.18	0.82	0.21	0.77	0.14	0.86	0.09	0.88	0.12	0.79
P62826	RAN	FNVVWDTAGQEK	RAN	74.77	0.93	#######	0.93	92.88	0.93	54.77	0.94	73.65	0.94	73.27	0.94
O95057	DIRA1	EAQAVAQEWK	RAS	0.38	0.96	0.27	0.93	0.18	0.94	0.39	0.94	0.25	0.94	0.24	0.89
P55040	GEM	VVLIGEQGVGK	RAS	0.07	0.92	0.05	0.80	0.06	0.76	0.09	0.91	0.05	0.74	0.06	0.78
Q9NYS0	KBRS1	TLIEPFTLLASK	RAS	0.32	0.92	2.68	0.91	1.82	0.90	0.34	0.91	1.73	0.91	1.66	0.91
Q9NYR9	KBRS2	VVVCGQASVGK	RAS	0.51	0.88	0.86	0.92	0.57	0.92	0.44	0.89	0.56	0.92	0.39	0.93
P11234	RALB	AEEWGVQYVETSAK	RAS	0.45	0.84	1.36	0.93	1.33	0.94	0.71	0.89	0.95	0.92	0.52	0.52
P62834	RAP1A	QWCNCAFLESSAK	RAS	3.48	0.96	24.82	0.96	9.84	0.95	3.39	0.91	5.10	0.96	6.97	0.93
P61224	RAP1B	QWNNCAFLESSAK	RAS	12.15	0.99	28.27	0.98	20.99	0.98	10.51	0.98	18.14	0.97	14.86	0.98
P10114	RAP2A	VPVILVGNK	RAS	0.10	0.89	0.12	0.89	0.13	0.89	0.08	0.90	0.09	0.89	0.09	0.90
Q9Y3L5	RAP2C	VPLILVGNK	RAS	0.42	0.90	0.52	0.83	0.60	0.87	0.35	0.89	0.46	0.87	0.43	0.85
P01112	RASH	SFEDIHQYR	RAS	0.55	0.93	0.24	0.91	0.41	0.91	0.28	0.90	0.27	0.93	0.28	0.92
P01116	RASK	SFEDIHHYR	RAS	0.28	0.96	0.39	0.87	0.27	0.94	0.20	0.94	0.37	0.90	0.35	0.89
O14807	RASM	LVVVGDGGVGK	RAS	0.43	0.53	0.87	0.51	0.62	0.46	0.34	0.62	0.38	0.52	0.42	0.56
P01111	RASN	SFADINLYR	RAS	6.58	0.80	16.10	0.85	9.32	0.79	6.02	0.82	9.66	0.87	9.22	0.82
Q8IYK8	REM2	VMLVGESGVGK	RAS	0.13	0.67	0.16	0.63	0.11	0.64	0.08	0.66	0.10	0.53	0.12	0.66
Q15382	RHEB	ALAESWNAAFLESSAK	RAS	7.69	0.97	15.19	0.98	22.95	0.97	4.92	0.97	10.48	1.00	18.20	0.97
Q99578	RIT2	FCIDDAFHGLVR	RAS	0.10	0.42	1.52	0.47	0.86	0.49	0.17	0.55	0.67	0.50	0.67	0.50
P62070	RRAS2	QVTQEEGQQLAR	RAS	7.06	0.98	7.20	0.98	7.94	0.98	5.27	0.97	6.32	0.97	5.35	0.97
Q96S79	RSLAB	VIGTSETPIIIVGNK	RAS	0.08	0.72	0.09	0.76	0.11	0.69	0.09	0.71	0.16	0.83	0.15	0.83
Q9BPW5	RSLBB	TSLIPRPK	RAS	0.48	0.74	0.41	0.74	0.68	0.68	0.42	0.70	0.29	0.73	1.04	0.72
P60953	CDC42	TCLLISYTTNK	RHO	11.18	1.00	49.05	0.99	35.64	0.99	13.68	0.99	44.41	0.98	36.70	0.99
P63000	RAC1	HHCPNTPIILVGTK	RHO	5.57	0.92	12.83	0.95	8.05	0.93	8.72	0.92	10.92	0.94	8.92	0.93
P15153	RAC2	AVLCPQPTR	RHO	0.25	0.90	0.38	0.89	0.19	0.94	0.11	0.88	0.29	0.60	0.09	0.84
P60763	RAC3	LAPITYPQGLAMAR	RHO	2.40	0.99	9.45	0.99	7.51	0.99	2.02	0.99	6.09	0.98	5.97	0.97
P61586	RHOA	IGAFGYMECSAK	RHO	12.38	0.91	40.52	0.91	25.54	0.92	10.74	0.91	18.90	0.91	19.24	0.92
P62745	RHOB	IQAYDYLECSAK	RHO	0.72	0.83	1.16	0.82	0.74	0.83	3.11	0.77	4.76	0.81	2.54	0.81
P08134	RHOC	ISAFGYLECSAK	RHO	1.27	0.99	5.16	0.98	3.30	0.98	0.92	0.99	2.37	0.99	1.72	0.98
Q9HBH0	RHOF	AALYLECSAK	RHO	0.15	0.96	0.16	0.97	0.17	0.90	0.15	0.89	0.13	0.87	0.12	0.99
P84095	RHOG	YLECSALQQDGVK	RHO	2.15	0.90	5.05	0.92	4.69	0.92	3.46	0.91	5.40	0.91	5.58	0.91
P61587	RND3	IVVVGDSQCGK	RHO	1.02	0.52	0.94	0.58	0.97	0.53	1.26	0.55	1.21	0.53	1.06	0.52

Table S4.2C A complete list of ratio all small GTPases quantified in the scheduled LC-MRM analysis in cells depleted of writer, eraser, and reader proteins.

Uniprot Accession	Gene	Peptide	Subfamily	*METTL3* $^{-/-}$			*ALKBH5* $^{-/-}$			*FTO* $^{-/-}$		
				Mean	S.D.	R.S.D.	Mean	S.D.	R.S.D.	Mean	S.D.	R.S.D.
P84077	ARF1	MGNIFANLFK	ARF				0.50	0.36	72.68	0.79	0.12	15.56
P61204	ARF3	MGNIFGNLLK	ARF				0.42	0.17	39.77	0.82	0.16	19.79
P18085	ARF4	LGLQSLR	ARF	0.90	0.06	7.23	0.67	0.04	5.53	1.00	0.20	19.58
P84085	ARF5	LGLQHLR	ARF	0.91	0.14	15.68	0.99	0.10	10.43	1.15	0.20	17.18
P62330	ARF6	FNVWDVGGQDK	ARF	2.52	0.13	5.15	0.94	0.08	8.77	1.02	0.10	9.83
Q13795	ARFRP	DCLTQACSALTGK	ARF	0.84	0.13	15.02	0.89	0.14	15.85	1.08	0.26	23.87
P40616	ARL1	ILILGLDGAGK	ARF	0.67	0.13	19.53	0.90	0.17	18.98	1.04	0.19	18.18
Q8N8L6	ARL10	EVLVLGLDGAGK	ARF	1.30	0.27	21.01	0.72	0.12	16.03	0.95	0.24	25.12
Q9NXU5	ARL15	ELGGADNIR	ARF	1.14	0.17	14.89	1.35	0.20	14.94	1.08	0.07	6.70
P36404	ARL2	ELQSLLVEER	ARF	0.90	0.28	30.85	0.65	0.07	11.05	0.80	0.18	23.08
P36405	ARL3	ILLLGLDNAGK	ARF	0.68	0.11	16.33	1.06	0.11	10.59	1.05	0.14	13.66
P56559	ARL4C	FAENQGTPLLVIANK	ARF	0.90	0.05	5.97						
Q9Y689	ARL5A	AGLLIFANK	ARF	0.81	0.04	5.32	0.90	0.17	19.34	1.14	0.23	20.05
Q96KC2	ARL5B	AAVLIFANK	ARF	1.08	0.21	19.74	0.90	0.12	13.78	1.19	0.29	24.52
Q9H0F7	ARL6	IPILFFANK	ARF	0.94	0.14	14.85	0.66	0.15	22.92	0.94	0.14	14.56
Q96BM9	ARL8A	LWDIGGQPR	ARF	0.93	0.14	14.97	2.00	1.35	67.32	1.04	0.15	14.84
Q9NVJ2	ARL8B	IWDIGGQPR	ARF	0.91	0.21	23.66	1.01	0.20	19.97	1.06	0.11	10.66
Q9NR31	SAR1A	EIFGLYGQTTGK	ARF	1.14	0.11	10.04	0.93	0.07	7.09	0.96	0.17	17.88
Q9Y6B6	SAR1B	EMFGLYGQTTGK	ARF	0.75	0.11	15.34	1.18	0.38	32.02	0.97	0.35	36.30
Q9H7X7	IFT22	ILFVGPCESGK	RAB	0.78	0.07	9.13	0.91	0.23	25.13	1.08	0.22	19.98
Q9BW83	IFT27	CILAGDPAVGK	RAB	0.50	0.06	12.92	1.42	0.19	13.52	0.95	0.08	8.25
P61026	RAB10	FHTITTSYYR	RAB	0.72	0.08	11.57	1.05	0.12	11.69	1.32	0.24	18.45
Q6IQ22	RAB12	LQVIIIGSR	RAB	1.06	0.32	30.41	1.03	0.15	14.82	1.42	0.31	21.59
P51153	RAB13	LQVWDTAGQER	RAB	0.76	0.25	32.36	1.42	0.31	21.89	1.23	0.48	39.26
P61106	RAB14	TGENVEDAFLEAAK	RAB	0.95	0.12	12.70	0.94	0.13	14.37	1.04	0.20	19.49
P59190	RAB15	IQIWDTAGQER	RAB	0.98	0.18	18.40	1.25	0.67	53.08	1.24	0.85	68.21
Q9NP72	RAB18	TLTPSYYR	RAB	1.22	0.01	0.79	1.55	0.07	4.42	1.01	0.13	12.76
P62820	RAB1A	MGPGATAGGAEK	RAB	0.92	0.09	9.56	1.16	0.29	24.75	1.03	0.18	17.81
Q9NX57	RAB20	IVLLGDMNVGK	RAB	1.03	0.47	45.01	0.87	0.21	23.69	0.85	0.27	31.95
Q9UL25	RAB21	VVLLGEGCVGK	RAB	0.93	0.11	11.57	1.31	0.11	8.68	1.09	0.15	13.96
Q9ULC3	RAB23	TIGVDFLER	RAB	0.70	0.19	26.94	1.48	0.29	19.26	0.85	0.16	18.92
Q969Q5	RAB24	AQLFETSSK	RAB	0.76	0.09	12.44	1.56	0.29	18.65	1.13	0.17	14.99
Q9ULW5	RAB26	VMLVGDSGVGK	RAB	1.00	0.30	29.63						
P51157	RAB28	VAAEILGIK	RAB	1.20	0.05	4.48	1.66	0.37	22.31	1.04	0.21	20.06
P61019	RAB2A	MITIDGK	RAB	0.97	0.07	6.91	0.84	0.30	35.07	0.98	0.27	27.81
Q8WUD1	RAB2B	IGPQQSISTSVGPSASQR	RAB	1.03	0.08	7.59	1.09	0.31	28.24	1.13	0.25	22.47
Q15771	RAB30	IVLIGNAGVGK	RAB	1.35	0.24	18.02	0.61	0.14	23.05	1.22	0.31	25.65
Q13637	RAB32	VLVIGELGVGK	RAB				28.00	24.66	88.08	1.32	0.14	10.62
Q9BZG1	RAB34	AEYWAVSSLTGENVR	RAB	1.09	0.22	20.64						
Q15286	RAB35	LLIIGDSGVGK	RAB	1.00	0.28	28.14	1.23	0.13	10.58	0.96	0.11	11.70
Q96AX2	RAB37	VMLLGDTGVGK	RAB	1.20	0.78	64.90	1.18	0.27	22.80	0.96	0.26	27.00
P57729	RAB38	LLVIGDLGVGK	RAB	0.81	0.08	9.63	1.40	0.12	8.60	1.04	0.09	8.50
P20336	RAB3A	TYSWDNAQVLLVGNK	RAB	0.56	0.22	40.01	1.00	0.51	50.75	1.27	0.66	52.39
P20337	RAB3B	DASDQNFDYMFK	RAB	1.13	0.48	42.92						
O95716	RAB3D	LLLIGNSSVGK	RAB	1.28	0.21	16.28	0.73	0.12	15.97	1.05	0.16	15.38
Q5JT25	RAB41	LLFLGEQSVGK	RAB				0.97	0.35	36.54	0.81	0.14	16.84
Q8N4Z0	RAB42	VIFLLVGHK	RAB	0.87	0.30	34.39	2.61	0.69	26.46	1.62	0.60	37.30
Q86YS6	RAB43	VATELIMR	RAB	1.38	0.21	15.48	1.09	0.24	22.22	1.22	0.30	24.82
P20338	RAB4A	IESGELDPER	RAB	0.87	0.04	4.23	1.50	0.10	6.57	0.99	0.19	19.10

Uniprot Accession	Gene	Peptide	Subfamily	*METTL3* $^{-/-}$			*ALKBH5* $^{-/-}$			*FTO* $^{-/-}$		
				Mean	S.D.	R.S.D.	Mean	S.D.	R.S.D.	Mean	S.D.	R.S.D.
P61018	RAB4B	FLVIGSAGTGK	RAB	1.17	0.11	9.58	1.24	0.52	42.02	0.94	0.33	35.24
P51148	RAB5C	QASPNIVIALAGNK	RAB	0.69	0.19	27.23	1.19	0.27	22.67	1.01	0.12	11.51
Q9NRW1	RAB6B	QITIEEGEQR	RAB	1.16	0.22	19.25	1.36	0.21	15.74	1.11	0.14	12.30
P51149	RAB7A	FQSLGVAFYR	RAB	0.76	0.15	20.29	1.08	0.33	30.42	0.95	0.18	18.63
O14966	RAB7L	VLVVGDAAVGK	RAB	0.74	0.16	21.87	1.32	0.19	14.64	1.03	0.22	20.87
P61006	RAB8A	LALDYGIK	RAB	0.91	0.28	30.25	0.96	0.27	27.71	0.96	0.16	16.57
Q92930	RAB8B	NIEEHASSDVER	RAB	1.00	0.76	76.14	0.97	0.18	18.56	0.93	0.36	38.57
P51151	RAB9A	DATNVAAAFEEAVR	RAB	1.36	0.43	31.63						
Q9NP90	RAB9B	QSFENLGNWQK	RAB	1.33	0.17	12.56	1.88	0.31	16.43	1.58	0.23	14.80
Q5HYI8	RABL3	YLAAGSSNAVK	RAB	1.51	0.14	9.60	1.09	0.22	19.93	0.89	0.14	16.23
Q15907	RB11B	NILTEIYR	RAB	1.09	0.10	9.53	1.01	0.19	18.98	0.89	0.20	22.13
P51159	RB27A	SWIPEGVVR	RAB	1.33	0.05	4.14	1.50	0.51	33.82	0.95	0.20	20.86
O00194	RB27B	LLALGDSGVGK	RAB	1.29	0.08	6.11						
Q9H082	RB33B	SAIQVPTDLAQK	RAB	1.03	0.20	19.73	0.91	0.23	25.24	0.95	0.17	17.82
Q8WXH6	RB40A	LPLPSTLR	RAB	2.55	0.44	17.08	0.11	0.09	79.47	4.10	1.25	30.54
Q12829	RB40B	EIDEHAPGVPK	RAB	1.01	0.11	10.68	0.95	0.33	35.14			
Q9UBK7	RBL2A	IICLGDSAVGK	RAB	0.72	0.06	8.02	0.60	0.08	12.62	0.95	0.11	11.50
Q9UNT1	RBL2B	IDDINVTQK	RAB	0.71	0.28	39.12						
P62826	RAN	FNVWDTAGQEK	RAN	0.75	0.04	5.17	0.91	0.18	19.97	0.79	0.13	16.95
O95057	DIRA1	EAQAVAQEWK	RAS	1.10	0.22	19.72	1.04	0.12	11.76	1.19	0.09	7.68
Q9NYS0	KBRS1	TLIEPFTLLASK	RAS	0.88	0.22	24.99	0.63	0.16	26.18	1.06	0.16	14.81
Q9NYR9	KBRS2	VVVCGQASVGK	RAS	0.73	0.11	15.52	0.95	0.25	25.83	1.21	0.20	16.67
P55042	RAD	SMPVDER	RAS				1.18	0.43	36.87	0.93	0.23	24.94
P11234	RALB	AEEWGVQYVETSAK	RAS	1.06	0.48	44.93	1.51	0.68	44.74			
P62834	RAP1A	QWCNCAFLESSAK	RAS	1.20	0.63	52.90	1.65	1.02	61.53	0.79	0.15	18.47
P61224	RAP1B	QWNNCAFLESSAK	RAS	0.74	0.11	15.55	1.30	0.24	18.09	0.72	0.27	37.90
P10114	RAP2A	VPVILVGNK	RAS	0.77	0.04	5.49	0.92	0.23	24.56	0.91	0.16	18.06
P61225	RAP2B	ASVDELFAEIVR	RAS	1.59	0.50	31.31	1.34	1.29	96.65	0.91	0.45	49.51
Q9Y3L5	RAP2C	VPLILVGNK	RAS	0.79	0.07	9.10	0.93	0.18	18.73	0.85	0.11	13.46
P01112	RASH	SFEDIHQYR	RAS	0.78	0.32	41.48	0.99	0.12	12.06	0.97	0.21	21.56
P01116	RASK	SFEDIHHYR	RAS	0.97	0.30	30.30						
P01116-2	RASK	QGVDDAFYTLVR	RAS	0.97	0.30	30.30						
P01111	RASN	SFADINLYR	RAS	0.84	0.21	24.82	1.16	0.31	26.93	0.86	0.19	21.66
Q15382	RHEB	ALAESWNAAFLESSAK	RAS	0.71	0.08	11.04	0.92	0.58	63.17	1.06	0.45	42.95
Q99578	RIT2	FCIDDAFHGLVR	RAS	0.99	0.68	69.08						
P10301	RRAS	LFTQILR	RAS	0.99	0.14	14.66	2.11	0.52	24.46	0.99	0.19	19.36
P62070	RRAS2	QVTQEEGQQLAR	RAS	0.77	0.10	13.49	1.05	0.08	7.65	1.03	0.19	18.06
Q96S79	RSLAB	VIGTSETPIIIVGNK	RAS	1.41	0.36	25.25						
Q9BPW5	RSLBB	TSLIPRPK	RAS	0.92	0.54	58.69	1.48	0.51	34.65	1.13	0.26	23.22
P60953	CDC42	TCLLISYTTNK	RHO	1.05	0.16	15.23	0.95	0.21	22.25	0.99	0.24	24.76
P63000	RAC1	HHCPNTPIILVGTK	RHO	1.17	0.36	30.77						
P15153	RAC2	AVLCPQPTR	RHO	0.42	0.09	20.89	0.36	0.16	45.58	0.54	0.20	36.45
P60763	RAC3	LAPITYPQGLAMAR	RHO	0.76	0.10	13.44	0.86	0.21	24.81	0.92	0.23	24.51
P61586	RHOA	IGAFGYMECSAK	RHO	0.70	0.21	29.70	0.76	0.20	25.80	1.07	0.35	32.64
P62745	RHOB	IQAYDYLECSAK	RHO	3.28	0.16	5.00	2.25	1.03	45.76	1.37	0.36	26.02
P08134	RHOC	ISAFGYLECSAK	RHO	0.57	0.14	24.24	2.35	0.84	35.65	0.84	0.23	26.85
Q9HBH0	RHOF	AALYLECSAK	RHO	0.76	0.10	12.59	1.07	0.18	16.63	0.78	0.30	38.49
P84095	RHOG	YLECSALQQDGVK	RHO	1.29	0.28	21.99	1.43	0.07	4.84	0.92	0.13	14.35
P52198	RND2	IVVVGDAECGK	RHO				1.59	0.78	49.32	1.24	0.27	22.03
				1.02	Mean RSD	20.90	1.45	Mean RSD	26.70	1.06	Mean RSD	22.09

Uniprot Accession	Gene	Peptide	Subfamily	*YTHDF1* $^{-/-}$			*YTHDF2* $^{-/-}$			*YTHDF3* $^{-/-}$		
				Mean	S.D.	R.S.D.	Mean	S.D.	R.S.D.	Mean	S.D.	R.S.D.
P84077	ARF1	MGNIFANLFK	ARF	0.47	0.18	38.82	0.64	0.36	55.72	0.53	0.28	52.21
P61204	ARF3	MGNIFGNLLK	ARF	0.69	0.20	28.43	0.95	0.31	32.69	0.86	0.32	37.44
P18085	ARF4	LGLQSLR	ARF	0.93	0.10	11.04	0.96	0.14	14.16	0.90	0.32	34.97
P84085	ARF5	LGLQHLR	ARF	0.98	0.04	3.68	1.03	0.11	11.03	0.98	0.22	22.61
P62330	ARF6	FNVWDVGGQDK	ARF	0.79	0.15	18.93	0.97	0.12	12.50	0.94	0.35	37.33
Q13795	ARFRP	DCLTQACSALTGK	ARF	0.65	0.08	13.04	1.03	0.19	18.76	0.86	0.25	29.26
P40616	ARL1	ILILGLDGAGK	ARF	0.70	0.10	14.38	0.89	0.14	15.69	0.81	0.27	32.83
Q8N8L6	ARL10	EVLVLGLDGAGK	ARF	0.88	0.27	30.78	0.82	0.27	32.39	0.79	0.36	45.93
Q9NXU5	ARL15	ELGGADNIR	ARF	0.99	0.12	11.63	0.97	0.09	9.14	1.02	0.39	38.21
P36404	ARL2	ELQSLLVEER	ARF	0.81	0.16	19.52	1.02	0.16	15.50	0.74	0.25	34.37
P36405	ARL3	ILLLGLDNAGK	ARF	0.75	0.13	17.15	0.94	0.09	10.05	1.02	0.35	34.22
P56559	ARL4C	FAENQGTPLLVIANK	ARF									
Q9Y689	ARL5A	AGLLIFANK	ARF	0.88	0.14	15.66	0.96	0.19	19.48	1.00	0.33	32.58
Q96KC2	ARL5B	AAVLIFANK	ARF	0.82	0.13	16.28	0.97	0.25	26.28	0.95	0.27	28.31
Q9H0F7	ARL6	IPILFFANK	ARF	0.78	0.13	16.11	0.96	0.11	11.21	0.84	0.25	30.12
Q96BM9	ARL8A	LWDIGGQPR	ARF	0.84	0.13	16.06	0.89	0.13	15.05	0.94	0.18	19.44
Q9NVJ2	ARL8B	IWDIGGQPR	ARF	0.99	0.11	11.40	1.07	0.18	16.88	1.10	0.20	18.60
Q9NR31	SAR1A	EIFGLYGQTTGK	ARF	0.85	0.21	25.22	0.99	0.19	19.67	0.84	0.39	46.15
Q9Y6B6	SAR1B	EMFGLYGQTTGK	ARF	0.78	0.18	23.66	0.90	0.28	30.73	0.77	0.32	41.30
Q9H7X7	IFT22	ILFVGPCESGK	RAB	0.97	0.34	35.13	1.12	0.24	21.17	1.04	0.26	25.01
Q9BW83	IFT27	CILAGDPAVGK	RAB	1.18	0.15	12.99	1.23	0.16	12.66	1.20	0.58	48.28
P61026	RAB10	FHTITTSYYR	RAB	0.83	0.18	21.42	0.97	0.17	18.09	0.83	0.22	26.95
Q6IQ22	RAB12	LQVIIIGSR	RAB	0.90	0.26	29.18	1.06	0.25	23.61	0.95	0.17	17.56
P51153	RAB13	LQVWDTAGQER	RAB	0.78	0.12	15.22	0.93	0.19	20.45	1.14	0.32	27.93
P61106	RAB14	TGENVEDAFLEAAK	RAB	0.85	0.14	15.87	0.98	0.18	17.83	0.94	0.29	30.61
P59190	RAB15	IQIWDTAGQER	RAB	1.05	0.40	37.79	1.04	0.52	49.93	0.92	0.26	27.85
Q9NP72	RAB18	TLTPSYYR	RAB	0.96	0.11	11.03	1.04	0.11	10.33	0.85	0.18	20.93
P62820	RAB1A	MGPGATAGGAEK	RAB	0.85	0.14	16.11	0.87	0.21	23.89	1.00	0.19	19.27
Q9NX57	RAB20	IVLLGDMNVGK	RAB	0.93	0.05	5.60	1.00	0.29	29.10	0.77	0.32	41.98
Q9UL25	RAB21	VVLLGEGCVGK	RAB	0.93	0.16	17.46	1.00	0.14	14.21	0.94	0.37	39.91
Q9ULC3	RAB23	TIGVDFLER	RAB	1.06	0.22	20.57	1.26	0.22	17.34	0.71	0.25	34.91
Q969Q5	RAB24	AQLFETSSK	RAB	1.10	0.19	17.37	1.03	0.15	15.08	1.09	0.38	35.15
Q9ULW5	RAB26	VMLVGDSGVGK	RAB									
P51157	RAB28	VAAEILGIK	RAB	0.90	0.16	17.70	1.06	0.20	19.09	1.14	0.44	38.92
P61019	RAB2A	MITIDGK	RAB	0.74	0.21	28.13	0.84	0.34	40.03	1.13	0.39	34.25
Q8WUD1	RAB2B	IGPQQSISTSVGPSASQR	RAB	0.91	0.28	30.97	0.86	0.19	22.66	1.14	0.39	34.49
Q15771	RAB30	IVLIGNAGVGK	RAB	0.93	0.08	8.43	0.89	0.05	5.27	0.93	0.24	26.12
Q13637	RAB32	VLVIGELGVGK	RAB	0.81	0.28	34.37	1.19	0.69	57.98	0.78	0.44	57.06
Q9BZG1	RAB34	AEYWAVSSLTGENVR	RAB									
Q15286	RAB35	LLIIGDSGVGK	RAB	0.97	0.11	11.10	1.09	0.11	9.85	1.12	0.31	28.08
Q96AX2	RAB37	VMLLGDTGVGK	RAB	0.91	0.45	49.78	0.98	0.36	36.96	0.89	0.16	17.42
P57729	RAB38	LLVIGDLGVGK	RAB	0.76	0.07	8.84	0.83	0.10	12.41	0.40	0.36	89.94
P20336	RAB3A	TYSWDNAQVLLVGNK	RAB	1.09	0.56	51.32	1.06	0.64	60.32	0.98	0.55	55.97
P20337	RAB3B	DASDQNFDYMFK	RAB									
O95716	RAB3D	LLLIGNSSVGK	RAB	0.83	0.14	17.30	0.86	0.07	7.69	0.83	0.18	21.28
Q5JT25	RAB41	LLFLGEQSVGK	RAB	0.69	0.25	35.71	1.40	1.42	101.71	0.67	0.33	49.06
Q8N4Z0	RAB42	VIFLLVGHK	RAB	1.80	0.66	36.51	1.33	0.10	7.82	2.39	1.05	44.05
Q86YS6	RAB43	VATELIMR	RAB	1.05	0.21	20.24	1.02	0.16	15.59	1.01	0.34	33.65
P20338	RAB4A	IESGELDPER	RAB	0.82	0.05	6.30	0.90	0.02	2.52	0.93	0.25	26.63

Uniprot Accession	Gene	Peptide	Subfamily	*YTHDF1* $^{-/-}$			*YTHDF2* $^{-/-}$			*YTHDF3* $^{-/-}$		
				Mean	S.D.	R.S.D.	Mean	S.D.	R.S.D.	Mean	S.D.	R.S.D.
P61018	RAB4B	FLVIGSAGTGK	RAB	0.82	0.26	32.27	1.21	0.40	33.54	0.96	0.46	47.66
P51148	RAB5C	QASPNIVIALAGNK	RAB	0.81	0.15	18.87	0.85	0.14	15.89	0.84	0.24	29.14
Q9NRW1	RAB6B	QITIEEGEQR	RAB	1.00	0.26	26.18	0.81	0.10	12.17	0.82	0.29	35.08
P51149	RAB7A	FQSLGVAFYR	RAB	0.83	0.36	42.72	0.89	0.35	39.26	0.77	0.40	51.67
O14966	RAB7L	VLVVGDAAVGK	RAB	0.90	0.14	15.59	1.02	0.22	21.71	1.08	0.10	9.08
P61006	RAB8A	LALDYGIK	RAB	0.91	0.17	18.85	0.91	0.21	23.44	1.00	0.28	28.09
Q92930	RAB8B	NIEEHASSDVER	RAB	1.01	0.34	34.00	1.11	0.49	44.63	1.02	0.55	54.06
P51151	RAB9A	DATNVAAAFEEAVR	RAB									
Q9NP90	RAB9B	QSFENLGNWQK	RAB	0.99	0.61	61.74	0.93	0.13	14.19	1.03	0.22	21.91
Q5HYI8	RABL3	YLAAGSSNAVK	RAB	0.86	0.14	16.05	0.93	0.09	9.81	0.90	0.34	37.56
Q15907	RB11B	NILTEIYR	RAB	0.81	0.20	24.18	0.87	0.19	21.57	1.01	0.45	44.31
P51159	RB27A	SWIPEGVVR	RAB	0.65	0.08	11.97	0.90	0.18	19.47	0.77	0.32	41.92
O00194	RB27B	LLALGDSGVGK	RAB									
Q9H082	RB33B	SAIQVPTDLAQK	RAB	0.96	0.34	35.29	1.02	0.21	20.75	1.00	0.26	26.19
Q8WXH6	RB40A	LPLPSTLR	RAB	1.48	0.19	12.91	1.03	0.32	30.70	0.92	0.51	55.14
Q12829	RB40B	EIDEHAPGVPK	RAB									
Q9UBK7	RBL2A	IICLGDSAVGK	RAB	0.88	0.12	13.55	0.94	0.07	7.49	1.02	0.28	27.56
Q9UNT1	RBL2B	IDDINVTQK	RAB									
P62826	RAN	FNVWDTAGQEK	RAN	0.70	0.11	16.02	0.84	0.17	20.29	0.80	0.31	38.54
O95057	DIRA1	EAQAVAQEWK	RAS	1.03	0.06	5.85	1.07	0.02	1.65	1.11	0.23	20.68
Q9NYS0	KBRS1	TLIEPFTLLASK	RAS	0.68	0.15	21.55	0.93	0.16	17.48	0.72	0.27	37.86
Q9NYR9	KBRS2	VVVCGQASVGK	RAS	1.05	0.09	8.24	1.20	0.07	5.63	1.13	0.39	35.03
P55042	RAD	SMPVDER	RAS	0.82	0.27	32.99				0.77	0.12	15.97
P11234	RALB	AEEWGVQYVETSAK	RAS	1.21	0.43	35.87	1.12	0.38	33.72	0.92	0.39	42.45
P62834	RAP1A	QWCNCAFLESSAK	RAS	1.10	0.54	49.20	0.89	0.37	41.40	0.91	0.38	41.66
P61224	RAP1B	QWNNCAFLESSAK	RAS	0.81	0.43	53.18	0.91	0.38	42.08	0.94	0.33	35.12
P10114	RAP2A	VPVILVGNK	RAS	0.76	0.14	18.14	0.78	0.23	28.96	0.94	0.44	47.00
P61225	RAP2B	ASVDELFAEIVR	RAS	0.72	0.38	51.90	0.63	0.28	44.64	1.03	0.83	80.63
Q9Y3L5	RAP2C	VPLILVGNK	RAS	0.72	0.11	14.89	0.89	0.14	15.75	0.74	0.34	45.26
P01112	RASH	SFEDIHQYR	RAS	0.93	0.23	24.47	0.88	0.15	16.69	0.86	0.16	18.32
P01116	RASK	SFEDIHHYR	RAS									
P01116-2	RASK	QGVDDAFYTLVR	RAS									
P01111	RASN	SFADINLYR	RAS	0.97	0.24	24.98	1.06	0.21	19.79	0.88	0.33	38.15
Q15382	RHEB	ALAESWNAAFLESSAK	RAS	0.63	0.23	37.49	0.69	0.31	45.18	0.63	0.33	53.45
Q99578	RIT2	FCIDDAFHGLVR	RAS									
P10301	RRAS	LFTQILR	RAS	0.96	0.31	32.45	1.12	0.25	22.60	1.40	0.50	35.85
P62070	RRAS2	QVTQEEGQQLAR	RAS	0.86	0.13	15.39	0.95	0.16	16.63	1.11	0.41	37.21
Q96S79	RSLAB	VIGTSETPIIIVGNK	RAS									
Q9BPW5	RSLBB	TSLIPRPK	RAS	1.04	0.27	26.20	0.94	0.18	18.67	1.23	0.41	33.59
P60953	CDC42	TCLLISYTTNK	RHO	0.71	0.28	39.84	0.86	0.26	30.61	0.92	0.51	55.32
P63000	RAC1	HHCPNTPIILVGTK	RHO									
P15153	RAC2	AVLCPQPTR	RHO	0.33	0.17	52.34	0.24	0.09	36.12	0.09	0.08	87.69
P60763	RAC3	LAPITYPQGLAMAR	RHO	0.89	0.16	18.41	0.86	0.13	15.71	1.07	0.45	41.85
P61586	RHOA	IGAFGYMECSAK	RHO	0.93	0.22	24.03	1.11	0.31	28.31	0.91	0.31	33.63
P62745	RHOB	IQAYDYLECSAK	RHO	0.98	0.10	9.98	1.20	0.26	21.38	1.02	0.53	51.69
P08134	RHOC	ISAFGYLECSAK	RHO	1.11	0.37	32.98	1.14	0.46	40.87	1.06	0.65	61.27
Q9HBH0	RHOF	AALYLECSAK	RHO	1.05	0.17	16.42	0.85	0.11	13.28	0.68	0.18	26.71
P84095	RHOG	YLECSALQQDGVK	RHO	0.81	0.11	13.35	0.90	0.07	8.00	0.95	0.31	33.08
P52198	RND2	IVVVGDAECGK	RHO	1.23	0.32	25.92	1.42	0.75	52.94	1.08	0.48	44.89
				0.90	Mean RSD	23.65	0.97	Mean RSD	23.73	0.94	Mean RSD	37.31

Table S4.3 The expression of selected small GTPases between transcriptomes.

The transcriptome results were obtained from the CCLE database, and the proteome, which was obtained from the previously published data, of melanoma WM266/WM115, IGR37/IGR39 and colorectal SW620/SW480 paired cancer cell lines.

Epitranscriptomic Modulator Target	Gene	WM266/WM115				SW620/SW480				IGR39/IGR37			
		Transcriptome	Proteome	T/P	P/T	Transcriptome	Proteome	T/P	P/T	Transcriptome	Proteome	T/P	P/T
METTL3	IFT27	1.14	5.00	**0.23**	**4.39**	1.00				1.27	1.27	**1.00**	**1.00**
	RAB3A	0.65	2.27	**0.29**	**3.50**	1.89				1.80	0.28	**6.44**	**0.16**
	ARF6	0.69	1.11	**0.62**	**1.62**	1.65	1.13	1.46	0.69	0.70	1.47	**0.48**	**2.09**
	RRAS	0.50	1.49	**0.34**	**2.96**	0.15	0.14	1.05	0.95	0.05	5.00	**0.01**	**97.56**
ALKBH5	RAB30	1.58	0.34	**4.69**	**0.21**	10.08	2.43	4.14	0.24	0.42	1.30	**0.32**	**3.09**
	RBL2A	0.51	1.18	**0.43**	**2.30**	2.85	1.80	1.58	0.63	0.65	1.14	**0.58**	**1.74**
	RAB43	0.52	1.61	**0.32**	**3.08**	2.54	**0.60**	4.27	0.23	0.76	0.45	**1.70**	**0.59**
	ARF1		0.99			1.34	0.97	1.38	0.72		0.77	**0.00**	
METTL3 & ALKBH5	RHOC	0.29	1.19	**0.24**	**4.16**	2.11	**0.92**	2.28	0.44	0.48	0.92	**0.53**	**1.90**
	RHOB	0.53	0.47	**1.13**	**0.88**	0.77	0.50	1.55	0.65	1.21	0.61	**1.96**	**0.51**
	RAC2	0.28	1.33	**0.21**	**4.84**	0.56	0.83	0.67	1.48	0.03			
	ARL1	1.09	0.85	**1.27**	**0.79**	1.34	0.75	1.79	0.56	1.78	0.38	**4.71**	**0.21**
	RAB9B	2.41	0.98	**2.46**	**0.41**	0.49	1.06	0.46	2.16	2.41	3.34	**0.72**	**1.38**
Non-targets	RAB23	0.77	**1.04**	**0.74**	**1.35**	0.69	0.53	1.30	0.77	0.77	2.13	**0.36**	**2.78**
	RABL3	0.92	**1.22**	**0.75**	**1.33**	0.78	0.95	0.82	1.22	0.77	0.86	**0.89**	**1.12**
	ARL3	0.90	**0.96**	**0.94**	**1.07**	0.83	0.97	0.85	1.18	0.25	1.52	**0.17**	**5.97**
	RB33B	0.99	**1.08**	**0.92**	**1.08**	0.87	0.63	1.39	0.72	0.80	1.06	**0.75**	**1.33**
	CDC42	0.94	**1.20**	**0.78**	**1.28**	1.13	1.16	0.98	1.03	1.72	0.76	**2.26**	**0.44**
	SAR1A	1.16	**0.61**	**1.89**	**0.53**	1.60	1.80	0.89	1.13	0.66	1.18	**0.56**	**1.80**
	RAB2A	0.49	**1.30**	**0.37**	**2.67**	0.93	0.80	1.17	0.86	1.78	0.57	**3.12**	**0.32**
	DIRA1	0.85	**1.03**	**0.83**	**1.21**								
	RB11B	0.70	**0.82**	**0.86**	**1.17**	2.20	1.26	1.75	0.57	0.59	0.68	**0.87**	**1.15**
	RAB12	0.58	**1.85**	**0.31**	**3.22**	0.54	1.21	0.45	2.23	1.48	1.15	**1.29**	**0.78**
	RAB2B	0.73	**1.09**	**0.67**	**1.50**	2.62	1.33	1.97	0.51	0.78	1.45	**0.54**	**1.85**
	RAB8A	0.95	**1.02**	**0.93**	**1.08**	1.75	1.20	1.46	0.68	0.85	1.15	**0.74**	**1.35**
	RAB14	0.49	**1.82**	**0.27**	**3.68**	1.06	1.01	1.06	0.95	0.98	1.47	**0.66**	**1.50**
	ARL8B	1.29	**0.76**	**1.68**	**0.59**	1.10	0.94	1.17	0.86				
	ARF5	0.91	**0.88**	**1.03**	**0.97**	1.27	1.04	1.22	0.82	0.57	1.67	**0.34**	**2.92**

5 Chapter 5. METTL3 Promotes Histone H3K9 Acetylation Through Recruitment of Histone Acetyltransferase 1 to Chromatin

5.1 Introduction

The methyltransferase-like proteins 3 and 14 (METTL3/METTL14), along with several other proteins, form a complex that deposits a methyl group at the N^6 position of adenosine in mRNA[1]. The resulting N^6-methyladenosine (m^6A) is the most prevalent internal modification in mRNA[2-5]. In addition, the methyl group on m^6A can be removed by demethylases, including ALKBH5 and FTO[6,7]. This reversible methylation enables cells to achieve dynamic gene regulation through modulating the stability, splicing and translational efficiencies of mRNA[8].

Recently, m^6A was also found to be present in chromatin-associated RNAs (caRNAs) and regulate gene expression through reshaping the genomic landscape of histone post-translational modifications (PTMs)[9,10]. In this vein, Xu et al.[10] demonstrated that, at loci with IAPE-z retrotransposons, METTL3 catalyzes the formation of m^6A in caRNAs, which recruit histone methyltransferases to deposit H3K27me3 heterochromatin mark in mouse embryonic stem cells (mESCs). In addition, Wei et al.[11] showed that FTO could demethylate m^6A in long-interspersed element-1 (LINE1) RNA in mESCs to regulate the expression of LINE1-containing genes. Furthermore, depletion of FTO in mESCs gives rise to closed chromatin in mESCs, as manifested by augmented H3K9me3 and attenuated H3K4me3/H3K27Ac for those m^6A-marked LINE1 RNA loci[11]. These findings unveiled an epitranscriptomic mechanism in reversible chromatin regulation.

Recent studies also revealed that m^6A in RNA is involved in modulating gene expression in differentiated cells[12]. For instance, Xu et al.[13] showed that METTL3-installed m^6A in nascent RNAs prevents their association with the integrator complex, which facilitates premature transcription termination of protein-coding genes and attenuates gene expression. Although these studies established a crosstalk between METTL3-modulated epitranscriptome and epigenome, the protein interaction networks of METTL3 involved with such regulations remain largely unexplored.

HAT1 was initially characterized as a type-B histone acetyltransferase that acetylates newly translated histones H3 and H4 in the cytosol, but exhibited no detec christina.kary@us.nature.comactivity toward histones in nucleosomes[14,15]. Recent studies, however, also revealed a nuclear localization of HAT1 in mammalian cells. In this vein, alternative translation of HAT1 mRNA yields two protein isoforms; while the short isoform (HAT1β) exists in both the cytosol and nucleus, the long isoform (HAT1α) is present exclusively in the nucleus, suggesting potentially unrecognized functions of HAT1 in the nucleus[16].

Proximity biotin labeling using engineered ascorbate peroxidase (APEX) followed by enrichment with streptavidin beads allows for harsh washing conditions for removing non-specific binders[17]. Furthermore, the short APEX labeling time enables discovery of transient and/or dynamic protein-protein interactions[18]. In this study, we employed APEX-based proximity labeling to interrogate the proximity proteome of METTL3-containing methyltransferase complex. Our results led to the discovery of an interaction between METTL3 and HAT1, and a crosstalk between METTL3-mediated m^6A formation in

caRNAs and HAT1α-catalyzed formation of H3K9Ac in chromatin, especially at loci of LTR and LINE retrotransposons.

5.2 Materials and Methods

5.2.1 Cell Culture and Plasmid Transfection

HEK293T and HeLa cells (ATCC) were cultured in Dulbecco's modified Eagle's medium (DMEM, Thermo Fisher) supplemented with 10% fetal bovine serum (FBS, Thermo Fisher) and 1% penicillin-streptomycin solution (GE Healthcare). The cells were maintained at 37°C in a humidified incubator supplemented with 5% CO_2. Transient transfection was conducted by using transIT-2020 (Mirus) for 24 h according to the manufacturer's instructions. To express APEX-fused proteins at similar molar quantities, approximately 6.5×10^6 cells in a T-75 flask were transiently transfected with 8.0 and 4.8 μg of plasmid for ectopic expression of human METTL3-V5-APEX and EGFP-V5-APEX, respectively, or 8.0 and 3.2 μg of plasmids for expressing human METTL14-V5-APEX and EGFP-V5-APEX, respectively.

5.2.2 Lentivirus Packaging and Transduction

HEK293T cells were transfected with shRNA cloned into pLKO.1/puro plasmid (Addgene #8453) together with pLTR-G (Addgene #17532) envelope plasmid and pCMV-dR8.2 dvpr (Addgene #8455) package plasmid using TransIT-X2 dynamic transfection system (Mirus). Viral particles were collected from the culture medium at 48 h after transfection and filtered through a 0.45-μm sterile filter. Cells were transduced with lentivirus for 48 h and subsequently screened with 1.0 μg/mL puromycin.

5.2.3 siRNA Knock-Down

For transient knock-down of gene expression, approximately 3×10^5 cells were transfected with 50 pmol of siRNA using Lipofectamine RNAiMax transfection system (Invitrogen). The siRNAs used in this study included siMETTL3 (CUGCAAGUAUGUUCACUAUGA)[19], siYTHDC1 (Dharmacon J-015332-17-0010), and siHAT1 (Santa Cruz #sc-37948 for a pool of three target-specific siRNAs, and GUGAUGAUAUGAUUCUUAAAUAC for targeting 3'UTR of HAT1 mRNA, Origene).

5.2.4 Proximity Labeling

The APEX-based proximity labeling was conducted as described elsewhere[17]. In brief, cells cultured in a T-75 flask were first incubated with 500 μM biotin phenol (Sigma) in complete DMEM for 30 min. Labeling was initiated by adding H_2O_2 to a final concentration of 1 mM and agitating gently at room temperature for 1 min. The cells were subsequently washed with a quencher solution (10 mM sodium azide, 10 mM sodium ascorbate, 5 mM Trolox in 1× PBS) and 1× PBS twice each. The cells were pipetted off from the bottom of the flask with the quencher solution and subsequently centrifuged at 500 g for 5 min. Cell pellets were lysed with CelLytic M containing a protease inhibitor cocktail (Sigma) and quenchers (10 mM sodium azide, 10 mM sodium ascorbate, 5 mM Trolox) on ice for 30 min. The cell lysates were sonicated with a 2-sec pulse for 3 times and centrifuged at 16,000 g for 30 min at 4°C to remove cell debris.

To enrich biotinylated proteins, approximately 600 μg of cell lysates were adjusted to equal volume with lysis buffer and incubated with streptavidin-conjugated agarose beads (Thermo) on a rotator (15 rpm) at 4°C overnight. Proteins bound nonspecifically to the

agarose beads were removed by washing sequentially with CelLytic M twice, once each with 1 M KCl and 0.1 M Na_2CO_3, and twice with CelLytic M. For each washing step, agarose bead suspension was rotated (15 rpm) at 4°C for 5 min. Biotinylated proteins were subsequently eluted from the beads by boiling for 5 min in 3× loading buffer (Bio-Rad) containing 2 mM biotin.

5.2.5 Proteomic Sample Preparation

Eluted proteins were resolved on a 14% SDS-PAGE gel for approximately 1 cm and the gel was subsequently stained with Coomassie blue. The gel band in the molecular weight range of 20-250 kDa was excised for in-gel trypsin digestion, as described previously[20]. In brief, the gel band was cut into 1 mm^3 cubes and de-stained by vortexing for 15 min in 25% and 50% acetonitrile in 50 mM ammonium bicarbonate twice each. Cysteine reduction and alkylation were conducted by incubating the gel pieces in 10 mM DTT at 37°C for 1 h and 55 mM iodoacetamide in the dark at room temperature for 20 min, respectively. Proteins were digested with 300 ng trypsin (VWR) at 37°C for 18 hr. The tryptic peptides were extracted by vortexing the gel pieces in 50% acetonitrile/0.1% formic acid for 20 min twice. The resulting peptide mixtures were desalted with C_{18} ziptip and reconstituted in 0.1% formic acid for LC-MS/MS analysis.

For limited trypsin digestion, histone proteins were resolved with a 14% SDS-PAGE and the gel pieces corresponding to molecular weight range of 10-20 kDa were excised and cut into 1 mm^3 cubes. After destaining (*vide supra*), the proteins were digested with trypsin at an enzyme-to-substrate ratio of 1:100 at 37°C for 30 min.

5.2.6 TMT Labeling

Approximately 5 µg of histone tryptic peptides were reconstituted in 8.5 µL of 50 mM HEPES (pH 8.5). TMTsixplex reagents (Thermo Fisher) were resuspended in 40 µL of anhydrous acetonitrile (ACN) and 1.5 µL of each reagent was added to each sample. The resulting mixture was vortexed at room temperature for 2 h. The reactions were quenched by adding 0.9 µL of 5% hydroxylamine and by vortexing the resulting mixture at room temperature for an additional 15 min. To each sample was subsequently added 8.9 µL solution of water/ACN/formic acid (8/1/1, v/v), and the mixture was desalted with C_{18} ziptip.

5.2.7 Liquid Chromatography-Tandem Mass Spectroscopy (LC-MS/MS) Analysis of APEX Samples

The LC-MS/MS experiments were conducted on a Q Exactive Plus Quadrupole-Orbitrap Mass Spectrometer (Thermo Fisher) coupled with an Ultimate 3000 UPLC (Thermo Fisher). Desalted peptides were resuspended in 10 µL of 0.1% formic acid. Peptides were loaded onto a 3-cm capillary column (150 µm i.d.) packed in-house with C_{18} resin (5 µm in particle size and 120 Å in pore size, Dr. Maisch GmbH HPLC) with buffer A (0.1% formic acid in water) at a flow rate of 3 µL/min. The peptides were eluted from the trapping column and separated on a ~25-cm analytical column (5 µm in particle size and 120 Å in pore size, Dr. Maisch GmbH HPLC) packed in-house with C_{18} resin (3 µm in particle size and 120 Å in pore size, Dr. Maisch GmbH HPLC) at a flow rate of 300 nL/min using a linear gradient of 5-37% buffer B (80% acetonitrile and 0.1% formic acid in water) over 135 min.

Peptides eluted from the analytical column were ionized with a Flex nanoelectrospray ion source (Thermo Fisher). The spray voltage and the capillary inlet temperature were set at 2 kV and 325°C, respectively. Full-scan MS in the *m/z* range of 350-1200 were acquired at a resolution of 35 k. Maximal injection time for full-scan MS was set to 100 ms with an AGC of $1e^6$. For MS/MS acquisition, top 25 precursor ions were isolated at a width of 1.6 Th and subsequently fragmented by higher-energy collisional dissociation (HCD) with a normalized collision energy (NCE) of 28. MS/MS were recorded at a resolution of 17.5 k. Maximal ion injection time for MS/MS was 75 ms with an AGC of $1e^5$. The dwell time was 3 sec, the minimal AGC to trigger MS/MS acquisition was set at $1e^3$, and the duration for dynamic exclusion was set at 35 sec.

The LC-PRM experiments were performed on a Q Exactive Plus Hybrid Quadrupole-Orbitrap Mass Spectrometer (Thermo) coupled with a Dionex UltiMate 3000 UPLC system (Thermo) or an Orbitrap Fusion Lumos Tribrid Mass Spectrometer (Thermo) coupled with an Easy-nLC 1200 UPLC system (Thermo). A Skyline PRM library containing 29 peptides representing 25 histone regulatory proteins enriched in METTL3 proximity proteome was developed from LC-MS/MS data acquired in the DDA mode (*vide supra*). The precursor ions were isolated in the quadrupole with an isolation window of 1.6 Th, followed by HCD with NCE values of 28 and 35 on the Q Exactive Plus and Fusion Lumos Mass Spectrometers, respectively. MS/MS were acquired in the Orbitrap with a resolution of 50 k.

5.2.8 LC-MS/MS Analysis of TMT Samples

LC-MS/MS analysis of TMT-labeled samples was conducted on an Orbitrap Fusion Lumos tribrid mass spectrometer (Thermo Fisher) coupled with an Easy-nLC 1200 UPLC system (Thermo Fisher). Sample loading and elution were conducted as described above. Eluted peptides were ionized with a Flex nanoelectrospray ion source (Thermo Fisher). The spray voltage was set to 2 kV and a capillary inlet temperature was set at 305°C. Full-scan MS were acquired in the range of *m/z* 375-1400 and at a resolution of 50 k. Maximal injection time and the AGC were set as default for full-scan MS. For MS/MS acquisition, precursor ions were isolated at a window of 0.5 Th and subsequently fragmented by HCD at an NCE of 35. Fragment ions were scanned at a resolution of 50 k.

5.2.9 LC-MS/MS Data Processing

Raw LC-MS/MS files for proximity proteome were processed in MaxQuant (version 2.0.1.0)[21]. Methionine oxidation and N-terminal acetylation were set as variable modifications; cysteine carbamidomethylation was specified as a fixed modification. Label-free quantification (LFQ) was enabled with the normalization set as none[22]. Mass tolerance for full-scan MS was set as 20 and 4.5 ppm for the first and main search, respectively. The mass tolerance for MS/MS was set at 20 ppm. Two missed cleavages were allowed for trypsin. Peptide spectra were searched against target-decoy Uniprot human proteome database (UP000005640) and the proteins were subsequently filtered at 1% FDR. Unnormalized LFQ with a minimum ratio determined from 2 peptides was used for protein quantification. The potential contamination and the decoys were removed from the output files. Gene ontology and Kyoto Encyclopedia of Genes and Genome (KEGG)

pathway analysis were conducted on the functional annotation tool embedded in Database for Annotation, Visualization and Integrated Discovery (DAVID) Resources v6.8 (https://david.ncifcrf.gov/)[23].

For the LC-MS/MS data of TMT-labeled histone samples, the type of LC-MS run was set to 'Reporter ion MS2" with "6plexTMT" as isobaric labels. Reporter ion mass tolerance and the mass tolerance for MS/MS were set at 0.01 Th and 20 ppm, respectively. Cysteine carbamidomethylation was specified as a fixed modification. Methionine oxidation, N-terminal and lysine acetylation, lysine mono-, di- and tri-methylation, lysine di-glycine modification, arginine mono-methylation and serine/threonine phosphorylation were set as variable modifications. Up to 6 missed trypsin cleavages were allowed. The PSM and peptides were both filtered at 1% FDR.

The acquired LC-PRM data were processed in Skyline. The extracted-ion chromatograms of MS/MS spectra for all precursor ions were manually inspected. A dot product (dotp) value of > 0.7 was imposed to ensure a similar fragmentation pattern as the corresponding spectrum in the library for high-confidence peptide identification. Approximately 4-6 fragment ions with nearly identical retention time were selected, and potential interfering fragment ions that did not overlay with other fragment ions were excluded. The sum of the peak areas of all 4-6 fragment ions were reported as the total area for the quantification of the corresponding protein.

5.2.10 Western Blot Analysis

Proteins were resolved on a 14% SDS-PAGE gel and transferred onto a nitrocellulose membrane, which was subsequently blocked with 5% blocker (Bio-Rad) in

0.05% PBST at room temperature for 40 min. The resulting membrane was hybridized sequentially with primary antibodies at 4°C overnight and horseradish peroxidase (HRP)-conjugated secondary antibodies at room temperature for 1 h. HRP signal was developed by using Pierce ECL Western Blotting Substrate (Thermo) and detected on an Odessey Imaging System (LI-COR Biosciences). The following antibodies were used: histone H3 (Proteintech, 17168-1-AP, 1:1000), H3K9Ac (Abcam, Ab4441, 1:1000), H3K9me2 (Cell Signaling Technology, 9753S, 1:1000), H3K9me3 (Abcam, AB8898, 1:1000), H3K14Ac (Abclonal, A7255, 1:1000), HAT1 (Proteintech, 11432-1-AP, 1:1000), METTL3 (Proteintech, 15073-I-AP, 1:2000), METTL14 (Abclonal, A8530, 1:2000), Lamin B (Proteintech, 66095-1-IG, 1:3000), V5 (Proteintech, 14440, 1:2000), GAPDH (Santa Cruz, SC-32233, 1:5000), hnRNPC (Abclonal, A19137, 1:1000), RBM17 (Abclonal, A3240, 1:1000), SF3B1 (Abclonal, A15801, 1:1000), SRSF2 (Abclonal, A3635, 1:1000).

5.2.11 Subcellular Fractionation

Subcellular fractionation was performed as described previously[24]. In brief, approximately 5 x 10^6 cells were lysed in pre-chilled cytoplasmic lysis buffer (10 mM Tris-HCl, pH 8.0, 0.34 M sucrose, 3 mM $CaCl_2$, 2 mM $MgCl_2$, 0.1 mM EDTA, 1 mM DTT, 0.5% Nonidet P-40, and protease inhibitor cocktail) on ice for 30 min. The intact nuclei were pelleted by centrifugation at 2,100 g for 2 min at 4°C. The supernatant was collected as cytoplasmic fraction, the nuclei pellet was incubated in a nuclear lysis buffer (20 mM HEPES, pH 7.9, 1.5 mM $MgCl_2$, 1 mM EDTA, 150 mM KCl, 0.1% Nonidet P-40, 1 mM DTT, 10% glycerol, and protease inhibitor cocktail) with vigorous pipetting followed by a 5-min incubation on ice. The resulting mixture was centrifuged at 16,000 g for 30 min at

4°C. The supernatant was collected as nucleoplasmic fraction, and the chromatin-enriched pellet was incubated in chromatin isolation buffer [20 mM HEPES, pH 7.9, 1.5 mM $MgCl_2$, 150 mM KCl, 10% glycerol, protease inhibitor cocktail, and 0.15 unit/μL benzonase (Sigma)] at 4°C for 2 h with a 20 sec vortexing in every 10 min. Debris was removed by centrifugation at 2,100 g for 2 min at 4°C, and the supernatant was subsequently collected as chromatin fraction.

5.2.12 Histone Extraction

Histones were extracted from cells as described previously[25]. In brief, intact nuclei were isolated from approximately 6.5×10^6 cells with NE-PER Nuclear and Cytoplasmic Extraction Reagents (Pierce). The isolated nuclei were resuspended in 400 μL of ice-chilled 0.4 N HCl. The resulting mixture was vortexed at 4°C for 4 h and subsequently centrifuged at 16,000 g for 30 min at 4°C. The histone-containing supernatant was transferred to a pre-chilled tube, and 200 μL of 100% trichloroacetic acid (TCA) was added dropwise. The mixture was incubated on ice for 2 h and centrifuged at 16,000 g for 30 min at 4°C. The histone-containing pellet was washed once with ice-cold acetone, and denatured by dissolving in a buffer containing 8 M urea and 50 mM Tris-HCl (pH 7.5). The histone proteins were quantified by using Bradford assay.

5.2.13 Co-Immunoprecipitation

Cells expressing Flag-tagged METTL3 or a comparable level of untagged METTL3 were washed twice with 1× PBS and centrifuged at 300 g for 5 min at room temperature. The cell pellets were lysed in chilled CelLytic M containing protease inhibitor cocktails. Protein lysates were quantified using Bradford assay. Approximately 600 μg of

proteins were adjusted to equal volume with a lysis buffer, mixed with 0.15 unit/μL of benzonase (Sigma) and 1 mM of $MgCl_2$, and to the mixture was subsequently added anti-Flag affinity beads. The resulting mixture was incubated on a rotator (15 rpm) at 4°C for 18 hr. The beads were subsequently washed with chilled CelLytic M for 5 times and the bound proteins were eluted by boiling in 2× loading buffer for 5 min.

To prevent non-specific interactions, beads were pre-blocked with BSA (Bio-Rad, 0.1 mg/mL), single-stranded salmon sperm DNA (Life technologies, 75 μg/mL), and glycogen (Invitrogen, 0.2 mg/mL) at room temperature for 2 h and subsequently washed twice with CelLytic M prior to the co-immunoprecipitation. The eluted proteins were resolved on an SDS-PAGE for Western blot analysis or in-gel trypsin digestion followed by LC-PRM analysis.

5.2.14 Purification of Flag-Tagged HAT1

Recombinant Flag-tagged HAT1 protein was purified from HEK293T cells. Briefly, HEK293T cells were plated in T75 flasks at ~35% confluency a day before transfection. On the next day, the cells were transfected with 15 μg Flag-tagged HAT1 plasmid. After 24 h, the cells were harvested and lysed in chilled CelLytic M containing protease inhibitor cocktail (1 mL lysis buffer per flask). Cell lysates were incubated with pre-washed Anti-Flag M2 Affinity Gel (40 μL gel per mL cell lysate, Sigma) at 4°C for 2 h on a rotator (12 rpm). After washing for three times with 0.1% PBST, the Flag-tagged proteins were eluted with PBS containing 250 ng/mL Flag peptide. Protein purity was verified by SDS-PAGE analysis, quantified by LI-COR Image Studio with BSA as the standard, and used immediately or stored at -80°C until use.

5.2.15 *In vitro* Acetylation Assay

In vitro acetylation assay was performed as described elsewhere[26]. Briefly, 1 μg of recombinant mononucleosome (H3.1) (Active Motif) or 154 ng of recombinant histone H3.1 (Active Motif) was incubated for 2 h at 30 °C in a 20-μL reaction mixture containing 50 mM Tris-HCl (pH 8.0), 10% glycerol, 1 mM DTT, 0.1 mM EDTA, 5 μM acetyl coenzyme A (acetyl-CoA, Cayman), 1× protease inhibitor cocktails (Sigma), and 0.5 μM of HAT1. To the mixture was subsequently added SDS-PAGE loading buffer (Bio-Rad) for gel electrophoresis.

5.2.16 Immunofluorescence Microscopy

HEK293T cells were cultured on glass coverslips in a 12-well plate and fixed with 100% MeOH at -20°C for 20 min. The cells were rinsed briefly with wash buffer (1× PBS and 0.1% of Triton-X-100), blocked with 2% BSA in wash buffer at room temperature for 1 h, and subsequently incubated with primary antibodies, which were diluted in blocking solution, at 4°C overnight. Excess antibodies were removed by rinsing the cells with wash buffer. The cells were then incubated with secondary antibodies at room temperature in the dark for 1 h. After removal of excess secondary antibodies, the cells were rinsed briefly with water and air-dried. The coverslips were mounted with Antifade Mounting Medium with DAPI (Vector Laboratories). Antibodies used were H3K9Ac (Abcam, Ab4441, 1:100), Flag (Cell Signaling #8146S, 1:100), Alexa 488-mouse secondary antibody (Invitrogen #A32723TR, 1:200), and Alexa 594-rabbit secondary antibody (Invitrogen, A11037, 1:250).

5.2.17 H3K9Ac Chromatin Immunoprecipitation-Sequencing (ChIP-seq) and Data Analysis

HEK293T and *METTL3*$^{-/-}$ cells were plated in 100 mm dishes at ~35% confluency a day before transfection. On the next day, the cells were transfected with an empty vector, wild-type or D395A METTL3 plasmid. After 40 h, the cells were crosslinked with 1% formaldehyde at room temperature for 10 min with gentle shaking, and quenched with 125 mM glycine for 10 min. After washing with ice-cold PBS buffer for three times, the cells were resuspended in cell lysis buffer (10 mM Tris-HCl, pH 8.0, 10 mM NaCl, 0.5% NP-40, protease inhibitors) at 4°C on a rotator for 20 min. After centrifugation at 3000 g for 3 min, the pellets were resuspended in RIPA buffer (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 2 mM EDTA, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS, protease inhibitors) and rotated at 4°C for 20 min, followed by a brief sonication on a Qsonica Sonicator q125 (42% amplitude, 10s on/10s off, 80 s). After incubation for another 20 min, cell lysate was sonicated on a Covaris S220 Sonicator for 6 min with a peak incident power of 140 W, a duty cycle of 10.1% and 200 cycles per burst at 4°C. After centrifugation at 4°C for 10 min, 50 μL of supernatant was saved as "Input" sample, and the rest of supernatant was subjected to IP.

For the IP experiment, protein A/G beads (Santa Cruz, sc-2003) were pre-blocked as described above, followed by incubation with H3K9Ac antibody (Abcam, ab4441) at room temperature for 1 h. After washing extensively, the antibody-bound beads were incubated with the cell lysate, followed by adding 500 μL of ChIP-buffer (50 mM HEPES-KOH, pH 8.0, 140 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate,

0.1% SDS, protease inhibitors) and 300 µL of TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA). The mixture was incubated at 4°C overnight on a rotator. The bound fractions were washed twice with low-salt washing buffer (20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% SDS), twice with high-salt wash buffer (20 mM Tris-HCl, pH 8.0, 500 mM NaCl, 2 mM EDTA, 1% Triton X-100, 0.1% SDS), twice with LiCl buffer (10 mM Tris-HCl, pH 8.0, 0.25 M LiCl, 1 mM EDTA, 1% NP-40, 1% sodium deoxycholate), and once with TE buffer. The elution was carried out in a 120-µL elution buffer (1% SDS, 100 mM $NaHCO_3$) at 65°C for 30 min on a thermomixer (1000 rpm). After centrifugation, the supernatant was transferred into a new tube, followed by adding 4.8 µL of 5 M NaCl and 2 µL of RNase A. After crosslinking reversal and RNA removal at 65°C for over 8 h, proteins were digested with 2 µL of proteinase K (20 mg/mL) at 60°C for 1 h. Subsequently, DNA was purified using ChIP DNA Clean & Concentrator (Zymo Research, D5205). The reverse crosslinking, RNA digestion, protein digestion, and DNA purification were conducted for "Input" samples in the same way as the "IP" samples.

The DNA-sequencing library was prepared using NEBNext Ultra DNA Library Prep Kit for Illumina (NEB) following the manufacturer's instructions. The purified DNA libraries were subsequently quantified using an Agilent 2100 Bioanalyzer and multiplexed for sequencing on an MGI-seq 2000 instrument (BGI, China) with 100-bp paired-end sequencing.

Raw reads were aligned to human hg38 reference genome using Bowtie2 (v2.4.5)[27]. PCR duplication was removed by Picard (v2.27.3)[28] MarkDuplicates. Unique reads were filtered by MAPQ > 1 using samtools (1.15.1)[29] view with parameters '-hb -p 1'. For

improving quantification accuracy, a ChIP-seq normalization method, ChIPseqSpikeInFree[30], was employed to calculate the scale factors for all ChIP-seq samples. Genome coverage bigwig files for Integrative Genomics Viewer visualization were generated by deeptools (v3.5.1)[31] bamCoverage using scale factors for the normalization. Peaks were generated by macs2 (v2.2.7.1)[32] callpeak with parameters '--BROAD --BROAD-CUTOFF -p'. Peak annotation was conducted using Homer (v4.8.2)[33] annotatePeaks.pl. The overlap of ChIP-seq peaks was analyzed by bedtools (v2.30.0)[34] intersect. Heat maps were generated by deeptools (v3.5.1)[31] computeMatrix and plotHeatmap. ChIP–seq density was calculated by Homer (v4.8.2)[33] analyzeRepeats. Violin plots were generated by GraphPad Prism 9.4.

5.2.18 Strand-Specific RNA-Seq and Data Analysis

Cells were cultured in 6-well plates at ~60% confluency, and total RNA was extracted using with E.Z.N.A. Total RNA Kit I (Omega, Bio-tek). After removing residual genomic DNA using Turbo DNA-free Kit (Invitrogen), the resulting RNA was treated with NEBNext rRNA Depletion Kit (NEB) to remove rRNA. The rRNA-depleted RNA was subjected to strand-specific RNA library preparation using NEBNext Ultra II Directional RNA library Prep Kit for Illumina (NEB). The resulting libraries were quantified using Bioanalyzer, followed by sequencing on an MGI-seq 2000 instrument (BGI, China) with 100-bp paired-end sequencing.

Raw reads were aligned to human hg38 reference genome using HISAT2 (v2.2.1)[35] with parameters '--rna-strandness RF'. Strand-specific RNA density was calculated by

Homer (v4.8.2)[33] analyzeRepeats with parameter '-strand'. Violin plots were generated by GraphPad Prism 9.4.

5.2.19 Chromosome-Associated RNA Extraction and rRNA Depletion

The chromatin fractions were extracted from HEK293T and the isogenic *METTL3*$^{-/-}$ cells following previously reported procedures.[36,37] Briefly, 10 million cells were collected and washed with a buffer containing 1 mM EDTA in PBS. After centrifugation at 500 g for 5 min, the cell pellet was resuspended in 200 μL lysis buffer (10 mM Tris-HCl, pH 7.5, 0.05% NP40, 150 mM NaCl), and incubated on ice for 5 min. The cell lysate was gently overlaid on the top of 2.5 volumes of chilled sucrose cushion (24% RNase-free sucrose in lysis buffer), and then centrifuged at 3500 g for 10 min. The resulting pellet was collected as nuclei fraction. After washing once with PBS containing 1 mM EDTA without dislodging the pellet, the nuclei pellet was resuspended in 200 μL prechilled glycerol buffer (20 mM Tris-HCl, pH 7.9, 75 mM NaCl, 0.5 mM EDTA, 0.85 mM DTT, 0.125 mM PMSF, and 50% glycerol in DEPC water) with gentle mixing. After the intact nuclei were readily taken up into suspension, an equal volume of nuclei lysis buffer (10 mM HEPES, pH 7.6, 1 mM DTT, 7.5 mM $MgCl_2$, 0.2 mM EDTA, 0.3 M NaCl, 1 M urea, and 1% NP-40 in DEPC water) was added into the nuclei suspension, and then vortexed vigorously for 5 sec twice. After incubation on ice for 2 min, the mixture was centrifuged at 15,000 g for 2 min to pellet chromatin. The pellet was washed gently with 1 mM EDTA in PBS without dislodging, and then collected as the chromatin-associated fraction.

caRNAs were isolated with TRIzol reagent. Briefly, chromatin-associated fraction was pipetted extensively to resuspend in 1 mL of TRIzol Reagent, followed by several

passages through a 21-gauge needle. After incubation on ice for 2 min, 200 μL chloroform was added into the mixture, which was mixed well by vigorous vortexing. The mixture was subsequently incubated at room temperature for 10 min on a rotator, and centrifuged at 12,000 g for 15 min at 4°C. The upper aqueous phase was transferred into a new tube, followed by adding an equal volume of isopropanol and 10 ng of glycogen. The sample was placed at -20°C for 1 h and centrifuged at 12,000 g for 15 min at 4°C to pellet caRNAs. After washing once with 75% ethanol, the caRNA pellet was dried under a vacuum for 1 min, followed by dissolving it into 20 μL DEPC water.

The resulting caRNA was treated with Turbo DNA-free Kit (Invitrogen) and RiboMinus Human/Mouse Transcriptome Isolation Kit (Thermo Scientific) successively to remove residual genomic DNA and rRNA by following the manufacturer's protocols.

5.2.20 Chromosome-Associated RNA m^{6}A-Seq and Data Analysis

m^{6}A immunoprecipitation was performed with EpiMark® N6-Methyladenosine Enrichment Kit (NEB) with some modifications. Briefly, 1 μg rRNA-depleted caRNA was spiked with 1 μL 1:1000-diluted m^{6}A spike-in (from m^{6}A purification kit). The RNA samples were chemically fragmented into ~100 nt fragments by a 5-min incubation at 94 °C in a fragmentation buffer (10 mM $ZnCl_2$, 10 mM Tris-HCl, pH 7.0). After quenching with 0.05 M EDTA, the RNA fragments were purified by Zymo RNA Purification Kit (Zymo Research) and eluted into 30 μL RNase-free water. Two μL of the resulting RNA was saved as "Input" sample, and the rest was subjected to m^{6}A IP following the manufacturer's protocol. m^{6}A-seq library was prepared by using SMARTer Stranded Total RNA-Seq Kit v2 (Takara) according to the manufacturer's protocol. The resulting libraries were

quantified using Bioanalyzer, followed by sequencing on an MGI-seq 2000 instrument with 100-bp paired-end sequencing.

Raw reads were aligned to human hg38 reference genome and spike-in genomes (NEB) using HISAT2 (v2.2.1)[35] with parameters '--rna-strandness RF'. Strand-specific reads were separated by using samtools (1.15.1)[29] view with specific flags 99, 147, 83 and 163. m^6A Peaks on each strand were called by using macs2 (v2.2.7.1)[32] callpeak with parameters '--nomodel –keepdup 5 -q 0.01'. m^6A/input ratio bigwig files for metagene analysis were generated by deeptools (v3.5.1)[31] bamCompare using scale factors for the normalization. The scale factors for all samples were calculated based on the number of reads mapped to human and spike-in genomes. Metagene plots were generated by deeptools (v3.5.1)[31] computeMatrix and plotProfile.

Consensus motifs were called by using Homer (v4.8.2)[33] findMotifsGenome.pl.

5.2.21 Data Availability

The ChIP-seq, caRNA m^6A-seq and RNA-seq data generated in this study have been deposited into the NCBI GEO database with accession number GSE220870. The LC-MS/MS data were deposited to the ProteomeXchange (PXD-038545). Three published datasets, including GSE196563 (METTL3 ChIP-seq dataset for HEK293T cells), GSE140561 (caRNA m6A-seq datasets for wild-type and *METTL3*$^{-/-}$ HEC-1-A cells) and GSE49339 (RNA-seq datasets for wild-type and YTHDF2-depleted HeLa cells), were re-analyzed in this study.

5.3 Results

5.3.1 APEX-Labeling Uncovered Proximity Proteins of METTL3 and METTL14

METTL3-containing methyltransferase complex deposits m^6A in RNA, which was found to regulate numerous cellular processes; however, the protein interaction networks of this complex are largely unknown. To fill in this knowledge gap, we employed APEX-based proximity labeling to profile the proximity proteomes of METTL3 and METTL14, where bait proteins were fused to the N terminus of APEX through a V5 linker (Figure S5.1A) and EGFP-V5-APEX was used as a control to reveal non-specific labeling of abundant proteins (Figure 5.1A). Plasmids were titrated to achieve similar levels of expression for the bait proteins in HEK293T cells (Figure 5.1B and S5.1B). The resulting biotinylated proteins were enriched, digested with trypsin, and the ensuing peptides subjected to LC-MS/MS analysis in a data-dependent acquisition (DDA) mode.

The label-free quantification (LFQ) results showed that, with a 1.5-fold cutoff ratios of LFQ intensities, 351 and 154 proteins were enriched in the proximity proteomes of METTL3 and METTL14, respectively, relative to the EGFP control (Figure 5.1C). Additionally, approximately 250 proteins were under-represented in the proximity proteomes of METTL3 and METTL14 compared to that of EGFP (Figure S5.1C). In keeping with the fact that METTL14 forms a heterodimer with METTL3, the majority of the METTL14 proximity proteome overlapped with that of METTL3 (Figure 5.1D). Together, the APEX labeling results provide a list of candidate proteins that interact with the METTL3-containing m^6A methyltransferase complex.

To explore the functions of the interactome of the m^{6}A methyltransferase complex, we performed gene ontology (GO) analysis of biological processes associated with the proximity proteins of METTL3 and METTL14, and the results demonstrated that the proximity proteomes of METTL3 and METTL14 are mainly involved in mRNA splicing (Figure S5.2A-B), which is consistent with the subcellular localization (i.e., nuclear speckle) of the complex[1,38] and the previous finding that METTL3-mediated formation of m^{6}A could modulate mRNA splicing through binding with hnRNPC[39]. We further validated the LC-MS/MS results of selected splicing factors by analyzing the APEX-labeled samples using Western blot (Figure S5.2C).

The STRING protein network analysis also revealed adjacencies of METTL3-containing methyltransferase complex to proteins of gene expression, RNA transport, and translation initiation (Figure S5.2D). A number of these processes (e.g., RNA transport) are known to be regulated by METTL3/METTL14, whereas others (e.g., gene expression) are less well characterized. Consistent with the previous observation of an METTL14-independent function of METTL3 in promoting translation by tethering the 5′ and 3′UTRs of mRNA[2], the GO term of translation initiation was exclusively identified for the METTL3 proximity proteome (Figure S5.2). Together, our proximity proteomic results are in line with the current knowledge of the METTL3 methyltransferase complex.

5.3.2 METTL3 Regulates Post-Translational Modifications (PTMs) of Histones in Differentiated Cells

In light of the recent studies showing that METTL3 recruits the SETDB1 methyltransferase to catalyze the formation of H3K9me3 in mESCs[10], we asked whether

histone-modifying enzymes are also subjected to METTL3 regulation in HEK293T cells by assessing if histone regulatory proteins are enriched in the METTL3 proximity proteome. Our results showed that the mean enrichment ratio of histone regulatory proteins is significantly higher than the average enrichment ratio of all other quantified proteins (Figure 5.2A), suggesting a role of METTL3 in mediating histone PTMs in differentiated cells.

We next examined whether METTL3 modulates histone PTMs in HEK293T cells. To this end, we extracted core histones from HEK293T and the isogenic *METTL3*$^{-/-}$ cells for TMT-based quantitative proteomic analysis (Figure S5.3A). Interestingly, we observed an elevated H3K9me3, a repressive histone epigenetic mark, and decreased open chromatin marks, including H3K9Ac and H3K14Ac, in *METTL3*$^{-/-}$ cells (Figure 5.2B). Other common modifications were not altered upon genetic ablation of METTL3 (Figure S5.3B). In this vein, MS/MS of the peptide carrying H3K9Ac and H3K9me3, which exhibit the same nominal mass but different exact masses, provided unambiguous identification of the two types of modifications (Figure S5.3C-D). The results suggest that the regulatory mechanisms of METTL3 on histone methylation in HEK293T cells differ from what were previously observed in mESCs, where depletion of METTL3 led to decreased H3K9me3[10]. This finding is in keeping with the presence of distinct epigenetic regulatory mechanisms in differentiated cells vs. stem cells[40].

To confirm the LC-MS/MS results, we also assessed several histone modifications by Western blot analysis (Figure 5.2C-D). The results are in full agreement with the LC-MS/MS data and further substantiate the roles of METTL3 in regulating H3K9Ac. We also

found that shRNA- and siRNA-mediated knockdown of METTL3 led to reduced H3K9Ac in HeLa and HepG2 cells, respectively (Figure S5.4A-B), underscoring a general role of METTL3 in modulating H3K9Ac in differentiated cells.

We next sought to examine if the catalytic activity of the METTL3 methyltransferase complex is essential for its regulation of H3K9Ac. To this end, we reconstituted the *METTL3*$^{-/-}$ cells with wild-type METTL3 or its catalytically inactive D395A mutant[41,42]. Subsequent Western blot analysis revealed that complementation with wild-type METTL3, but not the D395A mutant, could restore H3K9Ac level (Figure 5.2E). In addition, knock-down of METTL14, another crucial component of the METTL3 m^{6}A methyltransferase complex, led to diminished H3K9Ac in HEK293T cells (Figure S5.4C). These results demonstrated that the enzymatic activity of the METTL3-containing methyltransferase complex is indispensable for METTL3-mediated H3K9Ac.

5.3.3 METTL3 Facilitates the Recruitment of HAT1 to Chromatin to Regulate H3K9Ac

HAT1, a histone acetyltransferase previously shown to regulate H3K9Ac[43,44], is among the proteins enriched in the proximity proteome of METTL3 (Figure S5.5A-B). To explore the physical and functional associations between HAT1 and METTL3, we expressed Flag-tagged METTL3 in HEK293T cells, immunoprecipitated METTL3 from lysates pre-treated with benzonase, and subjected the mixture to LC-MS/MS analysis, where the instrument was set up in the parallel-reaction monitoring (PRM) mode for assessing HAT1 tryptic peptides. Our results showed that HAT1 could be co-enriched with

Flag-METTL3 (Figure S5.5C-F), substantiating the interaction between METTL3 and HAT1.

To further explore the role of HAT1 in regulating H3K9Ac, we knocked down HAT1 in HEK293T and HepG2 cells, and monitored the level of H3K9Ac by Western blot analysis (Figure S5.6A-B). Our results revealed that genetic depletion of HAT1 in cells led to significant decreases in H3K9Ac.

Our aforementioned results established the interaction between METTL3 and HAT1, and the role of these two proteins in promoting H3K9Ac in chromatin. In light of the previous finding that METTL3 recruits histone methyltransferase to chromatin in mESCs[10], we hypothesized that METTL3 may recruit HAT1 to chromatin to promote H3K9 acetylation. To examine this possibility, we extracted chromatin fractions for LC-PRM and Western blot analysis of HAT1 (Figure S5.7A). Our results revealed a diminished chromatin occupancy of HAT1 upon genetic ablation of METTL3 (Figure 5.3A-B and Figure S5.7B) or upon shRNA-mediated knockdown of METTL14 (Figure S7C and S7D). The overall HAT1 protein level, however, was not altered upon genetic depletion of METTL3 or METTL14 (Figure S5.7E-F). Hence, these results support that the METTL3-containing methyltransferase complex regulates chromatin localization of HAT1 without altering its expression level.

To gain additional mechanistic insights and to determine if the chromatin localization of HAT1 requires the enzymatic activity of the METTL3 methyltransferase, we conducted subcellular fractionation of the *METTL3*$^{-/-}$ cells rescued with wild-type METTL3 or its catalytically inactive D395A variant (Figure 5.3C). Under optimized SDS-

PAGE conditions, we were able to resolve the two isoforms of HAT1. Strikingly, the long isoform, i.e., HAT1α, is localized exclusively in the chromatin fraction, suggesting a potential unique function of HAT1α in regulating histone acetylation in chromatin. Moreover, genetic depletion of METTL3 led to diminished chromatin occupancy of both isoforms of HAT1, which could be restored by complementing the cells with wild-type METTL3, but not its catalytically inactive counterpart. In line with this, siRNA-mediated knockdown of HAT1 resulted in decreased H3K9Ac in HEK293T cells, but not in the isogenic *METTL3*$^{-/-}$ cells (Figures 5.3C and S5.8A).

HAT1 was originally characterized as a type-B histone acetyltransferase that deposits acetyl group on newly synthesized histones H3 and H4 in the cytosol[14]. Recently, HAT1, however, was found to be localized predominantly in the nucleus[45], though its nuclear functions were largely unknown. To confirm the nuclear function of HAT1 in regulating H3K9Ac, we knocked down HAT1 with siRNA targeting its 3′UTR and subsequently complemented the cells with wild-type, enzymatically inactive HAT1 (E276Q), or the one fused with the nuclear export sequence (NES) of PKI to remove it from the nucleus[46] (Figure S5.8B). Immunofluorescence microscopy results confirmed the expression and subcellular localizations of the different variants of HAT1 (Figure S5.8C). Interestingly, reconstitution with wild-type HAT1α, but not its NES-fused counterpart, could restore H3K9Ac in HAT1 knockdown cells (Figure 5.3D). Moreover, *in vitro* acetylation assay results demonstrated that HAT1α could catalyze the acetylation of H3K9 in nucleosomes (Figure 5.3E and S5.8D). Together, our results revealed a novel nuclear

function of HAT1 in promoting H3K9 acetylation, and this function entails the enzymatic activity of METTL3.

In light of the above observations that the chromatin localization of HAT1 and the subsequent H3K9Ac depend on the enzymatic activity of METTL3, we reason that the HAT1-mediated H3K9Ac may also be modulated by m^6A demethylase(s). To test this hypothesis, we conducted similar experiments in HEK293T cells where *FTO* and *ALKBH5* genes were individually ablated with CRISPR-cas9 (Figure S5.9A-B)[47,48]. Interestingly, we observed an increased chromatin localization of HAT1 in $FTO^{-/-}$ cells, which is accompanied with an elevated H3K9Ac. This is in keeping with the nuclear localization of FTO and its function in regulating the expression of LINE1 elements[11]. On the other hand, HAT1 displays a diminished expression in $ALKBH5^{-/-}$ cells, which resulted in attenuated levels of HAT1 in the nucleoplasm and chromatin; H3K9Ac, however, exhibits a slight increase in $ALKBH5^{-/-}$ cells (Figure S5.9C). The results indicated that ALKBH5 might regulate H3K9Ac through a different mechanism.

The aforementioned observations that the chromatin occupancy of HAT1 is modulated by both m^6A writer (i.e., the METT3-METTL14 complex) and eraser (i.e., FTO) proteins suggest that m^6A reader protein(s) may also be involved in recruiting METTL3-HAT1 complex to chromatin. In this vein, we found that YTHDC1, a nuclear m^6A reader protein known to interact with METTL3[10], was also enriched in the METTL3 and METTL14 proximity proteomes (Figure S5.10A). Thus, we next examined whether HAT1's binding with chromatin is mediated by YTHDC1. Indeed our Western blot results showed that siRNA-mediated knock-down YTHDC1 led to diminished chromatin

occupancies of METTL3 and HAT1, which is accompanied with decreased H3K9Ac (Figure 5.3F). The expression levels of METTL3 and HAT1 proteins, however, were not impacted by genetic depletion of YTHDC1 (Figure S5.10B). Together, these results unveiled a crosstalk between epitranscriptome and epigenome, which involves the reader, writer and eraser proteins of m^6A, i.e., YTHDC1, METTL3 and FTO, respectively.

5.3.4 METTL3 Regulates H3K9Ac in Intron and Intergenic Regions

Our above results demonstrated that global levels of H3K9Ac could be regulated by the METTL3 methyltransferase complex. We next assessed how METTL3 modulates H3K9Ac at the genome-wide scale. To this end, we conducted ChIP-seq experiments for H3K9Ac in HEK293T cells, the isogenic *METTL3*$^{-/-}$ cells, and the knockout cells complemented with wild-type or catalytically inactive METTL3 (Figure S5.11A). Our results showed that, in HEK293T cells, H3K9Ac is primarily enriched in intron, intergenic and promoter regions (Figure 5.4A), which is consistent with the previous findings made for HepG2 and K562 cells (Figure S5.11B)[49,50]. The peak numbers and localizations of H3K9Ac in these four groups of samples were similar, except that there is an over 10% loss of H3K9Ac peaks in *METTL3*$^{-/-}$ background, and these loci with diminished H3K9Ac are mainly distributed in intron regions (Figure 5.4A).

Genetic depletion of METTL3 led to diminished density of H3K9Ac in intron, intergenic and exon regions, but augmented enrichment of H3K9Ac in promoters and 5′ UTRs (Figure 5.4B). In addition, these alternations of H3K9Ac density in *METTL3*$^{-/-}$ cells could only be restored by reconstitution with wild-type METTL3, but not its catalytically incompetent mutant (Figure 5.4B). Representative IGV plots showed the enrichment of

H3K9Ac in introns and intergenic regions (Figure 5.4C & Figure S5.11C). These results indicate that METTL3 facilitates efficient spreading of H3K9Ac in intron, intergenic and exon regions, and this process requires the catalytic activity of METTL3.

METTL3 was shown to be enriched at IAP repetitive elements in mESCs to enable the establishment of heterochromatin through augmented deposition of H3K9me3[10]. Thus, we next investigated H3K9Ac density in different types of repetitive elements. Our results showed that most repetitive elements display much lower H3K9Ac densities in *METTL3*$^{-/-}$ cells than parental HEK293T cells (Figure 5.5A). The diminution in H3K9Ac in these regions in *METTL3*$^{-/-}$ cells could again be rescued by wild-type METTL3, but not its catalytically inactive mutant (Figure 5.5A). Among these repeats, long terminal repeats (LTR), long interspersed nuclear elements (LINE), and short interspersed nuclear elements (SINE) are retrotransposons and are among the most widely studied repetitive elements[51]. We also observed that the effects of METTL3 on H3K9Ac deposition are more pronounced in LTR and LINE than SINE retrotransposons (Figure 5.5A).

We further compared the distribution of H3K9Ac in these repeats among different cell lines, and it turned out that H3K9Ac is evenly distributed in these repetitive elements (Figure 5.5B and Figure 5.11D). In addition, the densities of H3K9Ac around these elements decreased significantly upon genetic ablation of *METTL3*. Moreover, the loss of H3K9Ac at these elements was fully restored by complementation with wild-type METTL3, but not its catalytically inactive mutant, suggesting that the METTL3-mediated H3K9Ac deposition in these repetitive elements entails its catalytic activity. Together, these results underscored an association between METTL3 and H3K9Ac deposition at

repetitive elements, especially for LTR and LINE elements, which requires the catalytic activity of METTL3.

We next interrogated the co-localization between METTL3 and H3K9Ac loci on chromatin in HEK293T cells using publicly available METTL3 ChIP-seq dataset[52]. Strikingly, approximately 90% of METTL3-binding events overlap with H3K9Ac loci, and the co-localization occurs primarily in intron, promoter, and intergenic regions (Figure S5.11E).

Given that the catalytic activity of METTL3 is important for its role in promoting H3K9Ac, we hypothesized that m^{6}A-containing transcripts may play essential roles in the deposition of H3K9Ac in these regions. To test this hypothesis, we conducted m^{6}A-seq for rRNA-depleted caRNAs isolated from parental and *METTL3*$^{-/-}$ HEK293T cells. The sequencing results led to the identification of ~30,000 m^{6}A peaks in HEK293T cells, and the m^{6}A peak number exhibited a ~20% of decrease upon METTL3 depletion (Figure S5.12A). These m^{6}A peaks contain the RRACH (R= A/G, H= A/C/U) consensus sequences (Figure S5.12B), as well as an additional RAR (R= A/G) motif, which was also detected in previous studies (Figure S5.12C)[36,53,54]. Interestingly, the RAR and CAG are the predominantly enriched motifs in m^{6}A peaks detected in METTL3-depleted cells (Figures S5.12D-E). These results suggest that m^{6}A writer proteins other than METTL3 (e.g., METTL16) might also be involved in the deposition of m^{6}A on caRNAs, which requires further investigation.

Consistent with the previous findings[10,55], m^{6}A peaks were predominantly enriched at UTRs of caRNAs, and exhibited a global decrease in signal in *METTL3*$^{-/-}$ cells (Figure

5.6A). On the grounds that the regulatory role of H3K9Ac mainly occurs in intron and intergenic regions, we examined the alterations of m^6A level in these two regions, as well as retrotransposon elements. As expected, m^6A level in these regions is significantly attenuated upon METTL3 depletion (Figure 5.6B). Notably, m^6A signal exhibits much more pronounced enrichment in LTR and LINE transcripts than SINE transcripts, especially at the 5′ termini of these transcripts (Figure 5.6C and Figure S5.12D), suggesting that m^6A in LTR and LINE transcripts might be involved in gene regulation.

5.3.5 METTL3-Mediated H3K9Ac Modulates Transcript Abundance From Intron and Intergenic Regions

To explore the effects of METTL3-dependent histone H3K9Ac on gene expression, we conducted strand-specific RNA-seq for rRNA-depleted RNA isolated from HEK293T cells deficient in HAT1 or METTL3. To this end, we knocked down HAT1 in HEK293T cells using a mixture of three target-specific siRNAs, and confirmed the knock-down efficiency using Western blot analysis (Figure S5.13A). RNA-seq results showed that mRNA expression in intron and intergenic regions was significantly down-regulated in HAT1-depleted cells (Figure 5.6D), and the magnitudes of decrease in RNA reads for these regions were much more pronounced than those for exon and promoter regions (Figure S5.13B), indicating an essential role of H3K9Ac in activating transcription in intron and intergenic regions.

Given that METTL3 predominantly regulates H3K9Ac in intron and intergenic regions, the above results suggest an involvement of METTL3 in modulating transcript levels in these regions. However, different from what we observed for cells depleted of

HAT1, *METTL3*$^{-/-}$ cells do not exhibit declines in the overall transcript levels in intergenic and intron regions (Figure S5.14A). In light of the fact that m^{6}A in mRNAs is known to promote their decay[4], we reason that loss of METTL3 may also confer elevated transcript stability, which would counteract the transcriptional repression elicited by the loss of H3K9Ac. To test this possibility, we re-analyzed the published RNA-seq data for HeLa cells with or without depletion of YTHDF2[4], an m^{6}A reader protein that regulates RNA decay. We indeed observed that the half-lives of intronic and intergenic RNAs increase significantly in YTHDF2-depleted cells (Figure S5.14B). Together, these findings suggest that two METTL3-dependent mechanisms, i.e., m^{6}A-mediated RNA decay and H3K9Ac-regulated transcription activation, may act together to modulate transcript levels from intron and intergenic regions.

A careful analysis of repetitive elements showed that the transcription repression elicited by HAT1 knockdown was specific to LTR and LINE elements, but not SINE elements (Figure 5.6E), which mirrors the specific enrichment of m6A on LTR and LINE elements (Figure 5.6C). As depicted in the scatter plot, the vast majority of LTR RNAs exhibited diminished transcript levels upon HAT1 knockdown, and this alternation was positively correlated with the change in H3K9Ac density in *METTL3*$^{-/-}$ cells and *METTL3*$^{-/-}$ cells expressing METTL3^{D395A}, but not in *METTL3*$^{-/-}$ cells complemented with METTL3^{WT} (Figure 5.6F). Similar findings were made for LINE elements (Figure 5.6G). We next sought to explore the effect of METTL3 depletion on the transcription of LTR and LINE elements. Again, METTL3 removal led to slight rises in LTR and LINE

transcript levels (Figure S5.14A), which might be attributed to augmented RNA stability in *METTL3*$^{-/-}$ cells (Figure S5.14B).

Together, these data demonstrated that METTL3-dependent deposition of m^{6}A on caRNAs in specific genomic regions (i.e., LTR and LINE elements) recruits HAT1 to install H3K9Ac at these regions, which plays a critical role in transcriptional regulation of genes in these regions.

5.4 Discussion

METTL3 and METTL14, together with several other proteins, form a complex catalyzing the formation of m^{6}A in mRNA. In this study, we examined the proximity proteomes of METTL3 and METTL14 by using APEX labeling together with proteomic analysis, with a goal to uncover new functions of the m^{6}A methyltransferase complex. We were able to identify a large number of proteins in the proximity proteomes of METTL3 and METTL14, where GO analysis revealed a marked enrichment of proteins involved in RNA splicing. In particular, we identified 75 and 26 proteins in the proximity proteomes of METTL3 and METTL14, respectively, that are involved in mRNA splicing according to the spliceosome GO term (Figure S5.2A and S5.2B). APEX labeling followed by Western blot analysis confirmed the presence of hnRNPC, RBM17, and SF3B1 in the proximity proteomes of METTL3 and METTL14 (Figure S5.2C). Previous CLIP-Seq experiment revealed that METTL3-bound mRNAs are frequently present in multiple isoforms, and genetic depletion of METTL3 resulted in pronounced variations in isoform distributions of these transcripts[1,56]. Moreover, Pan and coworkers[57] demonstrated that m^{6}A could alter the local structure of mRNA to modulate its binding with hnRNPC, thereby

regulating alternative splicing. Our results suggest that, apart from depositing m^6A in mRNAs and influence their binding with hnRNPC, METTL3 may also regulate mRNA splicing through binding with splicing factors.

METTL3 was recently shown to localize in the chromatin fraction to regulate histone modifications in mESCs; however, it remains unclear whether a similar mechanism is at play in differentiated cells. Our proximity proteomic data revealed the enrichment of > 20 histone regulatory proteins, including those involved in histone acetylation and methylation, in the proximity proteome of METTL3 (Figure 5.2A). These results suggest that METTL3 may assume a broad role in modulating histone PTMs in differentiated cells.

Histone PTMs, DNA cytosine methylation and microRNA constitute crucial epigenetic mechanisms for regulating gene expression. In this vein, H3K4me3 and H3K9Ac are known to confer open chromatin structure to promote transcription, whereas other modifications, e.g., H3K9me3 and H3K27me3, forge heterochromatin to repress gene expression[58]. Our quantitative proteomics and Western blot analysis revealed attenuated H3K9Ac along with augmented H3K9me3 in cells depleted of METTL3 (Figure 5.2B), suggesting distinct functions of METTL3 in modulating histone PTMs in differentiated cells (HEK293T or HeLa cells) vs. mESCs.

Among the histone regulatory proteins enriched in the proximity proteome of METTL3, HAT1 is a histone acetyltransferase that can catalyze the formation of H3K9Ac[43,44]. Our immunoprecipitation followed by PRM analysis further confirmed the interaction between METTL3 and HAT1 (Figure S5.5F). We also found that the chromatin occupancy of HAT1 was diminished in HEK293T cells with the *METTL3* or *METTL14*

gene being depleted (Figure 5.3A and S7C), and the chromatin localizations of HAT1 and H3K9Ac entail the catalytic activity of METTL3 m^6A methyltransferase complex (Figure 5.2E and 5.3C).

HAT1 is considered a type-B histone acetyltransferase catalyzing the deposition of acetyl group on free histones H3 and H4[14]. However, later studies revealed two translation isoforms of HAT1 with different subcellular distributions in mammalian cells[16,59]. Here, we showed that the long isoform, HAT1α is predominantly localized in chromatin (Figure 5.3C) and the deposition of H3K9Ac in chromatin is promoted by nuclear localization of HAT1 (Figure 5.3D). *In vitro* assay further revealed the acetyltransferase activity of HAT1 on nucleosome histones (Figure 5.3F). In this regard, the previous finding about the inability of HAT1 in acetylating nucleosome histones was made with the use of HAT1β; however, HAT1α, which resides exclusively in chromatin (Figure 5.3C), was not examined. Thus, our work shifts the current paradigm about the functions of HAT1 by revealing its role in catalyzing the formation of H3K9Ac in the chromatin.

Interestingly, the chromatin localization of HAT1 and the subsequent H3K9Ac were diminished in *FTO*$^{-/-}$ cells (Figure S5.9A). Consistent with this, FTO was shown to regulate histone epigenetic marks in cells[11]. If METTL3 recruits HAT1 for H3K9Ac, how does the demethylase affect the localization of HAT1 and the subsequent H3K9Ac? It is plausible that, apart from the METTL3, m^6A-containing caRNAs may recruit HAT1 to enable acetylation of H3K9 in chromatin. This notion finds support from our observation that genetic depletion of YTHDC1, a nuclear m^6A reader protein, led to attenuated chromatin occupancy of METTL3 and HAT1, and diminished H3K9Ac (Figure 5.3F). In

this vein, YTHDC1 is known to interact with METTL3, which was proposed to amplify m^6A signal in caRNAs by enhancing the recruitment of METTL3 m^6A methyltransferase complex to chromatin to install additional m^6A in caRNAs[10]. Hence, METTL3 and its installed m^6A in caRNAs stimulates the chromatin localization of HAT1 through protein-RNA (i.e., the binding of YTHDC1 with m^6A in caRNAs) and protein-protein interactions (i.e., the interactions of METTL3 with YTHDC1 and HAT1). Together, our results demonstrate a dynamic and reversible crosstalk of epitranscriptomic and epigenetic regulations in HEK293T cells, and reveal the molecular mechanisms underlying the crosstalk.

By conducting ChIP-seq analysis of H3K9Ac in *METTL3*$^{-/-}$ cells reconstituted with wild-type or enzymatically inactive METTL3, we sought to identify the genomic loci in which H3K9Ac is regulated by METTL3. Consistent with the Western blot results, ChIP-seq results revealed diminished H3K9Ac in the majority of genomic regions, including intron, intergenic and exon regions in HEK293T cells depleted of METTL3 (Figure 5.4B). Moreover, the loss of H3K9Ac in these regions can only be restored by reintroducing wild-type METTL3, but not its catalytically inactive mutant. Given the importance of METTL3 in modulating repetitive elements, we further examined the occupancy of H3K9Ac in different types of repetitive elements, including LTR, SINE, LINE, satellite, etc. (Figure 5.4A). Altogether, ChIP-seq results support an important role of METTL3 in regulating H3K9Ac in chromatin, especially in intron and intergenic regions, which entails its catalytic activity.

We also confirmed the enrichment of METTL3-dependent m^6A modification in intron and intergenic regions through m^6A-seq of caRNAs (Figure 5.6). Specifically, we found that m^6A is enriched in LTR and LINE transcripts, but not SINE transcripts, suggesting that m^6A-modified transcripts play important roles in the recruitment of HAT1 to LTR and LINE elements in chromatin. Additionally, our RNA-seq results revealed an important regulatory role of HAT1 in transcriptional activity of intron and intergenic regions, as well as LTR ad LINE, but not SINE elements (Figure 5.6), where HAT1-mediated H3K9Ac activates transcription in these regions. METTL3's effect on transcript levels, however, did not parallel that of H3K9Ac, even though our ChIP-seq results strongly suggested an METTL3-dependent H3K9Ac in these regions. This can be attributed to the fact that METTL3 can modulate transcript levels through multiple mechanisms, including METTL3-dependent H3K9Ac that activates gene transcription, METTL3-installed m^6A that regulates mRNA decay,[4] and perhaps other METTL3-mediated histone modifications that may also influence gene expression[10].

In summary, we provided the first comprehensive maps of the proximal proteomes of METTL3 and METTL14. We also revealed a novel interaction between METTL3 and HAT1, uncovered a new function of HAT1 in acetylating H3K9 in chromatin, and unveiled a crosstalk between the epitranscriptome and epigenome in differentiated cells, where METTL3 and its methylation product in caRNAs promote the chromatin localization of HAT1 and H3K9Ac in chromatin.

5.5 Acknowledgements

This work was supported by the National Institutes of Health (R35 ES031707).

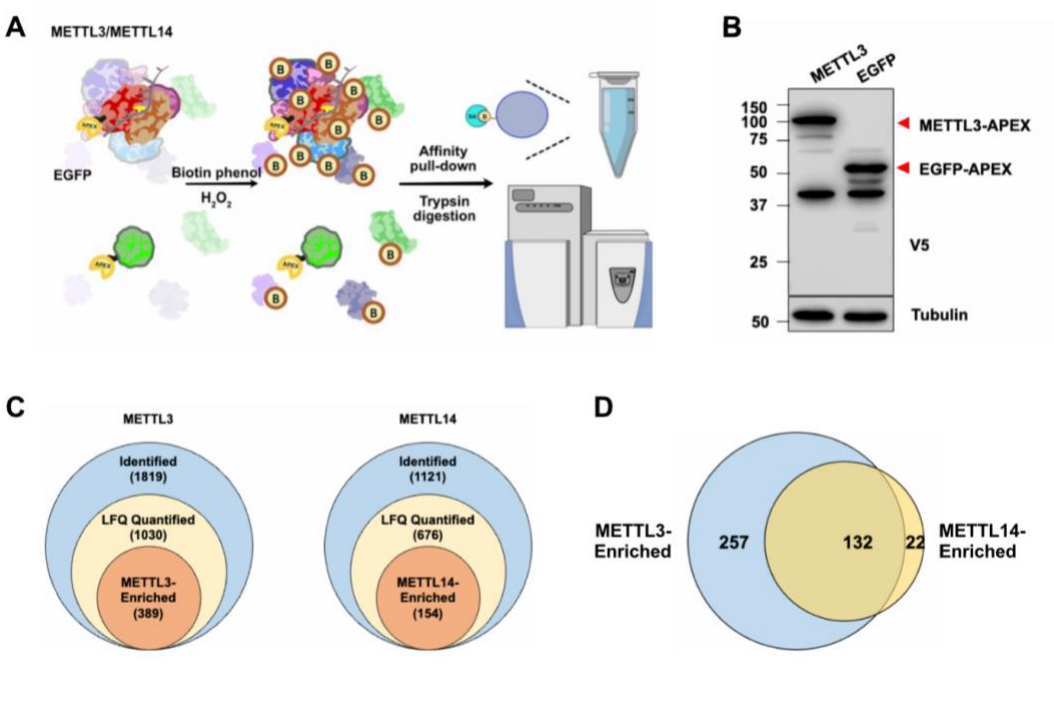


Figure 5.1 APEX-based proximity labeling for revealing the interactomes of METTL3 and METTL14.

(A) A schematic diagram depicting the experimental workflow. (B) Western blot analysis showing comparable expression levels of METTL3- and EGFP-APEX in HEK293T cells. (C) Venn diagrams displaying the numbers of proteins identified, quantified by label-free quantification, and enriched (> 1.5-fold) in the METTL3 (left) and METTL14 (right) proximity proteomes. (D) A Venn diagram illustrating the overlap of proteins enriched in METTL3 and METTL14 proximity proteomes.

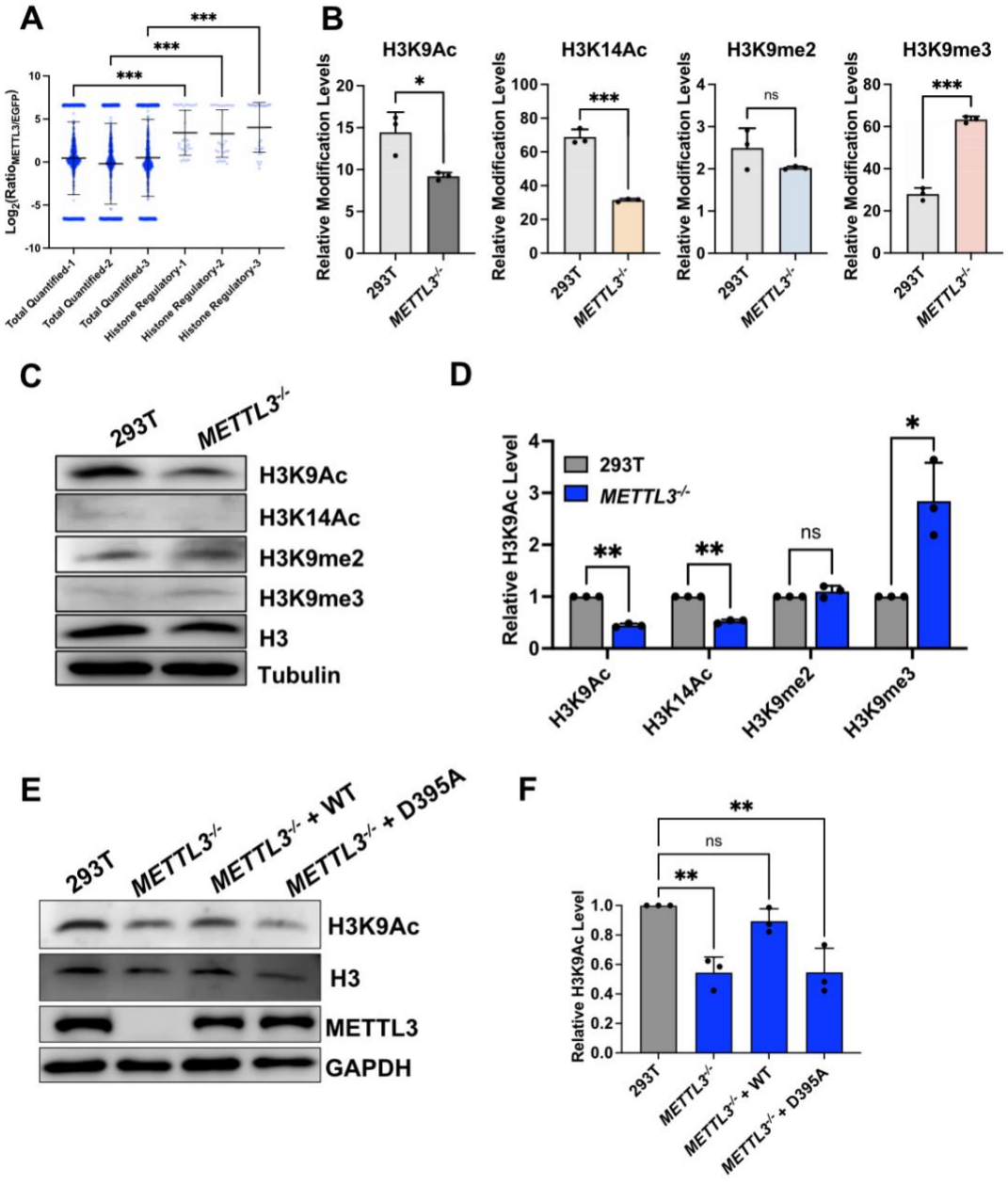

A
B
C
D
E
F
H3K9Ac
H3K14Ac
H3K9me2
H3K9me3
293T
METTL3-/-
Relative Modification Levels
Relative H3K9Ac Level
H3
Tubulin
METTL3
GAPDH
METTL3-/- + WT
METTL3-/- + D395A
Total Quantified-1
Total Quantified-2
Total Quantified-3
Histone Regulatory-1
Histone Regulatory-2
Histone Regulatory-3
ns

Figure 5.2 METTL3 and METTL14 regulate PTMs of core histone proteins.

(A) A scatter plot depicting the ratio of METTL3 versus EGFP for all quantified proteins or those involved in histone regulations in 3 biological replicates. Random arbitrary numbers of approximately 2^7 or -2^7 were assigned to proteins identified only in METTL3 or EGFP proximity proteome, respectively. (B) Bar charts showing the percentage of histone H3 modifications in *METTL3*$^{-/-}$ and parental HEK293T cells. Histones were extracted from cells for TMT-based quantitative proteomic analysis. The intensities of the modified peptides were normalized against the total intensities of that peptide sequence in all modified forms. (C-D) Western blot validations and the corresponding quantification results of selected histone H3 modifications in *METTL3*$^{-/-}$ cells. The signal intensities of the modified H3 bands were first normalized against H3, and the resulting ratio for *METTL3*$^{-/-}$ cells was further normalized against that of the HEK293T cells. (E-F) Western blot analysis of H3K9Ac in *METTL3*$^{-/-}$ cells complemented with wild-type METTL3 or the catalytically inactive D395A mutant. All data represent mean ± SD of results obtained from three biological replicates. The *p* values were calculated using two-tailed Welch's *t*-test (A and C) or one-way ANOVA (E) (*, $0.01 < p < 0.05$; **, $0.001 < p < 0.01$; ***, $p < 0.001$; ns, $p > 0.05$).

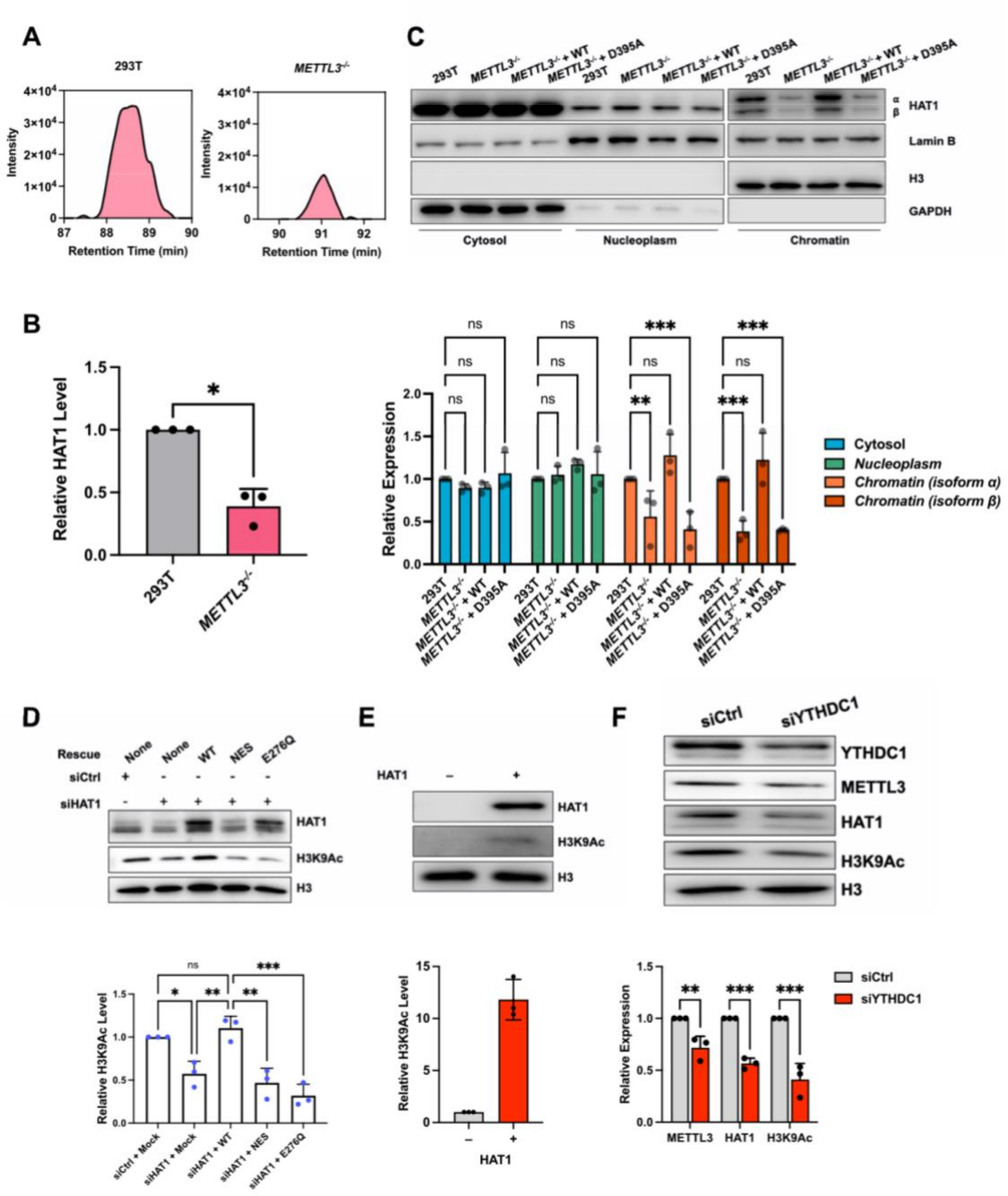

A
293T
METTL3-/-
Intensity
Retention Time (min)
B
Relative HAT1 Level
293T
METTL3-/-
Relative Expression
Cytosol
Nucleoplasm
Chromatin (isoform α)
Chromatin (isoform β)
C
293T
METTL3-/-
METTL3-/- + WT
METTL3-/- + D395A
HAT1
Lamin B
H3
GAPDH
Cytosol
Nucleoplasm
Chromatin
D
Rescue
None
None
WT
NES
E276Q
siCtrl
siHAT1
HAT1
H3K9Ac
H3
Relative H3K9Ac Level
siCtrl + Mock
siHAT1 + Mock
siHAT1 + WT
siHAT1 + NES
siHAT1 + E276Q
E
HAT1
H3K9Ac
H3
Relative H3K9Ac Level
F
siCtrl
siYTHDC1
YTHDC1
METTL3
HAT1
H3K9Ac
H3
Relative Expression
METTL3
HAT1
H3K9Ac

Figure 5.3 METTL3 regulates HAT1-dependent histone H3K9 acetylation.

Extracted-ion chromatograms (A) and the corresponding quantification results (B) of a unique HAT1 peptide obtained from LC-PRM analysis of the chromatin fractions extracted from HEK293T or the isogenic *METTL3*$^{-/-}$ cells. (C) Western blot analysis and the corresponding quantification results of HAT1 in different subcellular fractions of *METTL3*$^{-/-}$ cells reconstituted with wild-type or D395A METTL3. The signal intensities of HAT1 were first normalized to the controls (GAPDH; Lamin B, and histone H3 for cytoplasm, for nucleoplasm, and chromatin fractions, respectively), and the resulting ratios were further normalized against those obtained for HEK293T cells. (D) Western blot analysis and the corresponding quantification results of HAT1 knockdown cells rescued with wild-type HAT1α, enzymatically inactive mutant (E278Q), or HAT1α fused with the NES of PKI (NES). (E) Western blot analysis and the corresponding quantification results showing the *in vitro* acetylation of histone H3.1 in mononucleosomes with (+) or without (-) HAT1. The signal intensities of the H3K9Ac bands were first normalized against H3, and the resulting ratio was further normalized against that without HAT1. (F) Western blot analysis and the corresponding quantification results showing the chromatin occupancies of HAT1 and METTL3 and the level of H3K9Ac in YTHDC1 knock-down cells. The signal intensities of the HAT1, METTL3, and H3K9Ac bands were first normalized against H3, and the resulting ratio was further normalized against that without HAT1. All data represent mean ± SD obtained from three biological replicates. The *p* values were calculated using two-tailed Welch's *t*-test (**, $0.001 < p < 0.01$; ***, $p < 0.001$; ns, $p > 0.05$).

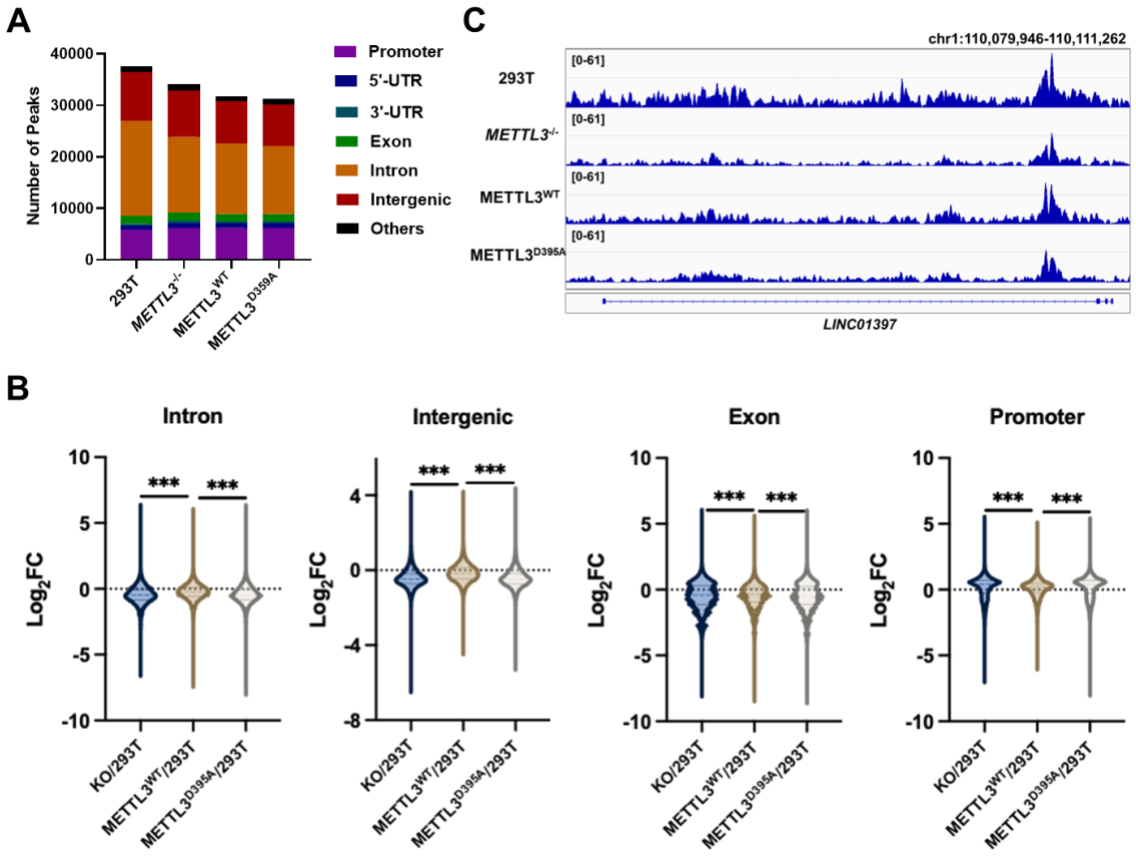


Figure 5.4 METTL3 regulates the H3K9Ac in introns and intergenic regions.

(A) Bar graphs showing the H3K9Ac distribution in HEK293T, *METTL3*$^{-/-}$, and the *METTL3*$^{-/-}$ cells complemented with wild-type or mutant METTL3. (B) Violin plots showing fold changes (Log_2FC) in H3K9Ac occupancy in intergenic, intron, exon and promoter regions in *METTL3*$^{-/-}$ and two rescued cell lines compared to the parental cells. (C) IGV plot showing H3K9Ac enrichment in intron region in four cell lines (y axis indicates ratios of IP over input).

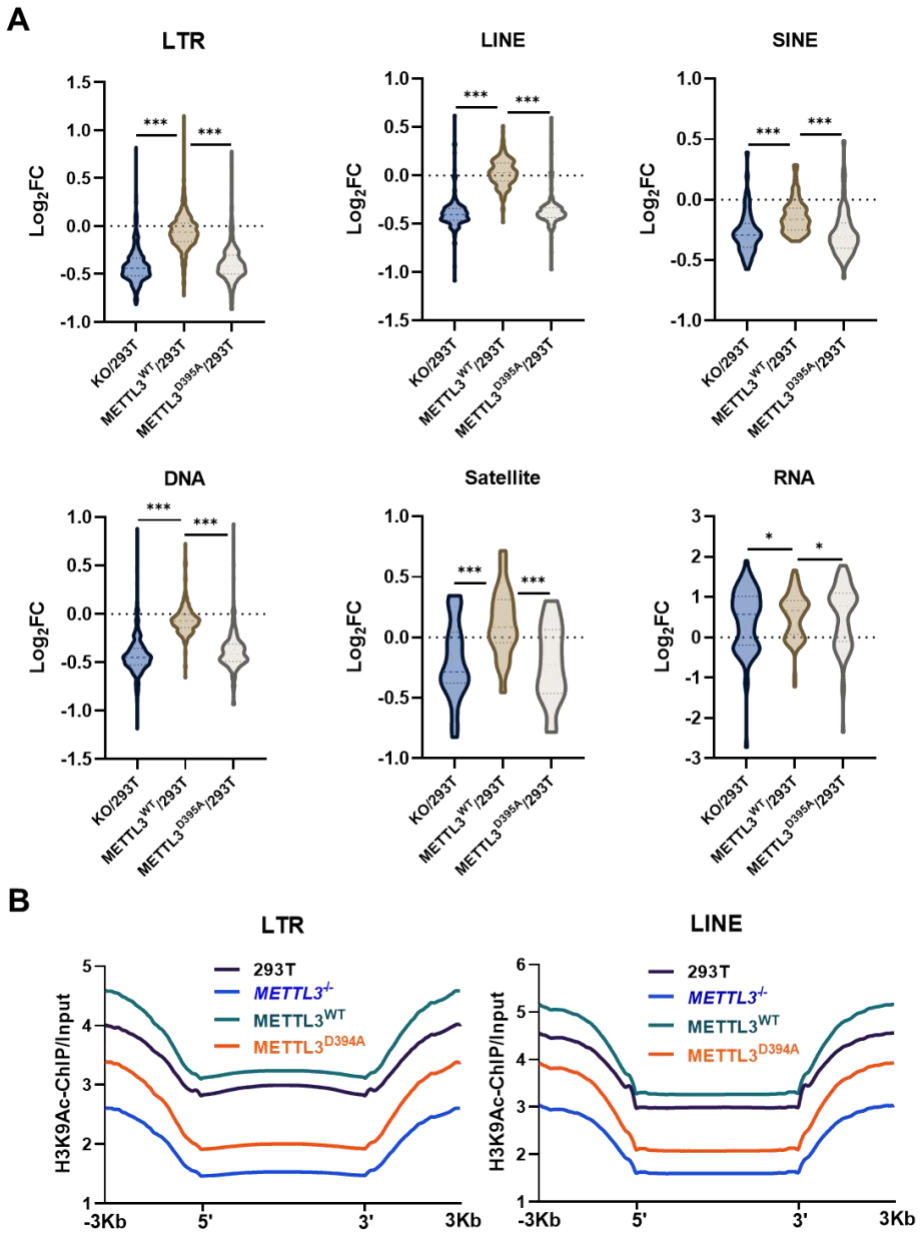


Figure 5.5 METTL3 regulates H3K9Ac in repetitive elements.

(A) Violin plots showing fold changes (Log_2FC) in H3K9Ac density at LTR, LINE, SINE, DNA and RNA elements in *METTL3*$^{-/-}$HEK293T cells and the two rescued cell lines compared to the parental cells. (B) Aggregation plots showing the distribution of H3K9Ac

on LTR and LINE elements in 293T, *METTL3*$^{-/-}$, wild-type or mutant METTL3 rescued cell lines.

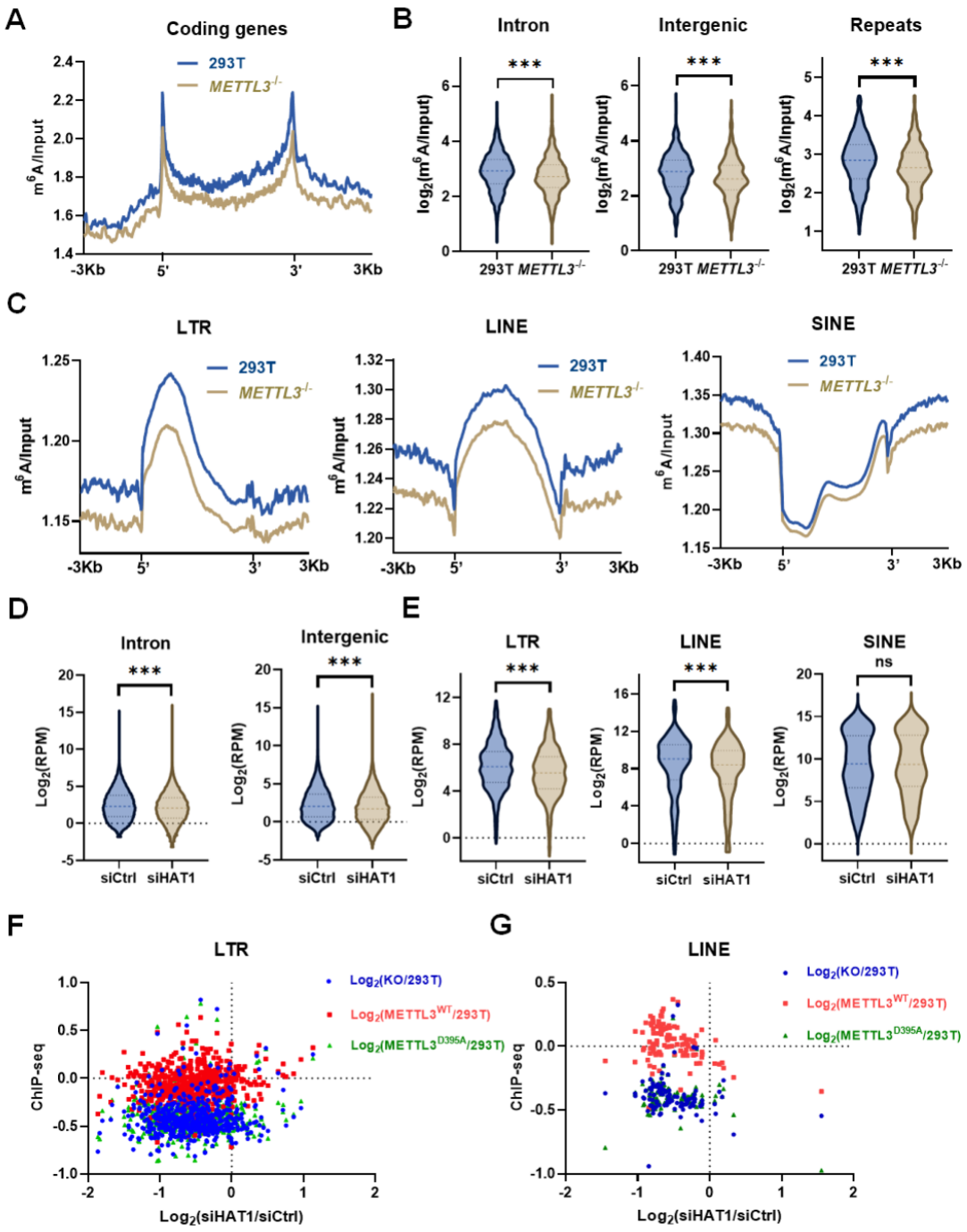

A
Coding genes
293T
METTL3-/-
m6A/Input
-3Kb
5'
3'
3Kb
B
Intron
Intergenic
Repeats
***
log2(m6A/Input)
293T METTL3-/-
C
LTR
LINE
SINE
293T
METTL3-/-
m6A/Input
-3Kb
5'
3'
3Kb
D
Intron
Intergenic
***
Log2(RPM)
siCtrl
siHAT1
E
LTR
LINE
SINE
***
ns
Log2(RPM)
siCtrl
siHAT1
F
LTR
Log2(KO/293T)
Log2(METTL3WT/293T)
Log2(METTL3D395A/293T)
ChIP-seq
Log2(siHAT1/siCtrl)
G
LINE
Log2(KO/293T)
Log2(METTL3WT/293T)
Log2(METTL3D395A/293T)
ChIP-seq
Log2(siHAT1/siCtrl)

Figure 5.6 m^{6}A on caRNA recruits HAT1 to install H3K9Ac in intron and intergenic regions, especially in LTR and LINE elements, thereby activating transcription in these regions.

(A) Metagene profile of m^{6}A peak density along caRNAs of coding genes. (B) m^{6}A levels on caRNAs of intron, intergenic, and repetitive regions in parental and *METTL3*$^{-/-}$ 293T cells. (C) Aggregation plots revealing the enrichment of METTL3-dependent m^{6}A of caRNAs on LTR and LINE but not SINE transcripts (crRNAs). (D-E) Violin plots showing the RNA densities in intron and intergenic regions (D); LTR, LINE and SINE elements (E) in siCtrl and siHAT1 cells. RPM indicates reads per million. (F-G) Comparison between RNA-seq and H3K9Ac ChIP-seq data in LTR (F) and LINE elements (G). The *p* values were calculated using two-tailed Welch's t-test, ***, $p < 0.001$; ns, $p \geq 0.05$.

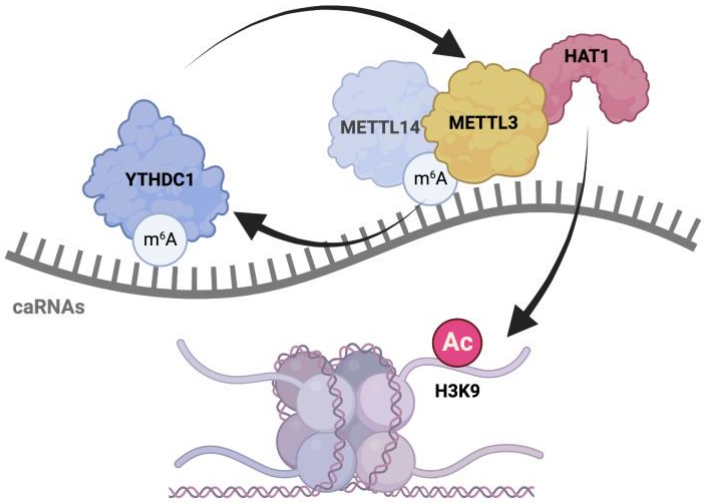


Figure 5.7 Proposed model of METTL3 promotes H3K9Ac through recruitment of HAT1 to chromatin.

5.6 References

1. Ping, X.L. et al. Mammalian WTAP is a regulatory subunit of the RNA N^6-methyladenosine methyltransferase. *Cell Res.* **24**, 177-89 (2014).

2. Choe, J. et al. mRNA circularization by METTL3-eIF3h enhances translation and promotes oncogenesis. *Nature* **561**, 556-560 (2018).

3. Ries, R.J. et al. m^6A enhances the phase separation potential of mRNA. *Nature* **571**, 424-428 (2019).

4. Wang, X. et al. N^6-methyladenosine-dependent regulation of messenger RNA stability. *Nature* **505**, 117-20 (2014).

5. Li, A. et al. Cytoplasmic m^6A reader YTHDF3 promotes mRNA translation. *Cell Res.* **27**, 444-447 (2017).

6. Zhao, X. et al. FTO-dependent demethylation of N^6-methyladenosine regulates mRNA splicing and is required for adipogenesis. *Cell Res.* **24**, 1403-19 (2014).

7. Zheng, G. et al. ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol. Cell* **49**, 18-29 (2013).

8. Jia, G. et al. N^6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat. Chem. Biol.* **7**, 885-7 (2011).

9. Liu, J. et al. N^6-methyladenosine of chromosome-associated regulatory RNA regulates chromatin state and transcription. *Science* **367**, 580-586 (2020).

10. Xu, W. et al. METTL3 regulates heterochromatin in mouse embryonic stem cells. *Nature* **591**, 317-321 (2021).

11. Wei, J. et al. FTO mediates LINE1 m^6A demethylation and chromatin regulation in mESCs and mouse development. *Science* **376**, 968-973 (2022).

12. Yuan, C. et al. Crosstalk of histone and RNA modifications identified a stromal-activated subtype with poor survival and resistance to immunotherapy in gastric cancer. *Front. Pharmacol.* **13**, 868830 (2022).

13. Xu, W. et al. Dynamic control of chromatin-associated m^6A methylation regulates nascent RNA synthesis. *Mol. Cell* **82**, 1156-1168 e7 (2022).

14. Parthun, M.R., Widom, J. & Gottschling, D.E. The major cytoplasmic histone acetyltransferase in yeast: links to chromatin replication and histone metabolism. *Cell* **87**, 85-94 (1996).

15. Verreault, A., Kaufman, P.D., Kobayashi, R. & Stillman, B. Nucleosomal DNA regulates the core-histone-binding subunit of the human Hat1 acetyltransferase. *Curr. Biol.* **8**, 96-108 (1998).

16. Tafrova, J.I. & Tafrov, S.T. Human histone acetyltransferase 1 (Hat1) acetylates lysine 5 of histone H2A in vivo. *Mol. Cell. Biochem.* **392**, 259-72 (2014).

17. Hung, V. et al. Spatially resolved proteomic mapping in living cells with the engineered peroxidase APEX2. *Nat. Protoc.* **11**, 456-75 (2016).

18. Paek, J. et al. Multidimensional tracking of GPCR signaling via peroxidase-catalyzed proximity labeling. *Cell* **169**, 338-349 e11 (2017).

19. Liu, J. et al. A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. *Nat. Chem. Biol.* **10**, 93-5 (2014).

20. Shevchenko, A., Tomas, H., Havlis, J., Olsen, J.V. & Mann, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* **1**, 2856-60 (2006).

21. Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc.* **11**, 2301-2319 (2016).

22. Cox, J. et al. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol. Cell. Proteomics* **13**, 2513-26 (2014).

23. Jiao, X. et al. DAVID-WS: a stateful web service to facilitate gene/protein list analysis. *Bioinformatics* **28**, 1805-6 (2012).

24. Aygun, O., Svejstrup, J. & Liu, Y. A RECQ5-RNA polymerase II association identified by targeted proteomic analysis of human chromatin. *Proc. Natl. Acad. Sci. U. S. A.* **105**, 8580-4 (2008).

25. Sidoli, S., Bhanu, N.V., Karch, K.R., Wang, X. & Garcia, B.A. Complete workflow for analysis of histone post-translational modifications using bottom-up mass spectrometry: from histone extraction to data analysis. *J Vis Exp* (2016).

26. Brown, J.A.L. In vitro histone acetylation assay. *Curr Protoc Pharmacol* **79**, 3.14.1-3.14.16 (2017).

27. Langmead, B. & Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nature Methods* **9**, 357-359 (2012).

28. Institute, B. Picard toolkit. *Broad Institute, GitHub repository* (2019).

29. Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078-9 (2009).

30. Jin, H. et al. ChIPseqSpikeInFree: a ChIP-seq normalization approach to reveal global changes in histone modifications without spike-in. *Bioinformatics* **36**, 1270-1272 (2020).

31. Ramírez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Research* **44**, W160-W165 (2016).

32. Zhang, Y. et al. Model-based analysis of ChIP-seq (MACS). *Genome Biology* **9**, R137 (2008).

33. Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Molecular Cell* **38**, 576-589 (2010).

34. Quinlan, A.R. & Hall, I.M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841-842 (2010).

35. Kim, D., Paggi, J.M., Park, C., Bennett, C. & Salzberg, S.L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology* **37**, 907-915 (2019).

36. Liu, J. et al. m6A methyladenosine of chromosome-associated regulatory RNA regulates chromatin state and transcription. *Science* **367**, 580-586 (2020).

37. Conrad, T. & Ørom, U.A. Cellular fractionation and isolation of chromatin-associated RNA. *Methods Mol Biol* **1468**, 1-9 (2017).

38. Han, D. et al. Dynamic assembly of the mRNA m^6A methyltransferase complex is regulated by METTL3 phase separation. *PLoS Biol.* **20**, e3001535 (2022).

39. Liu, N. et al. *N^6*-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature* **518**, 560-4 (2015).

40. Wu, H. & Sun, Y.E. Epigenetic regulation of stem cell differentiation. *Pediatr. Res.* **59**, 21R-5R (2006).

41. Wang, P., Doxtader, K.A. & Nam, Y. Structural basis for cooperative function of Mettl3 and Mettl14 methyltransferases. *Mol. Cell* **63**, 306-317 (2016).

42. Jambhekar, A., Dhall, A. & Shi, Y. Roles and regulation of histone methylation in animal development. *Nat. Rev. Mol. Cell. Biol.* **20**, 625-641 (2019).

43. Gruber, J.J. et al. HAT1 coordinates histone production and acetylation via H4 promoter binding. *Mol. Cell* **75**, 711-724 e5 (2019).

44. Nagarajan, P. et al. Histone acetyl transferase 1 is essential for mammalian development, genome stability, and the processing of newly synthesized histones H3 and H4. *PLoS Genet.* **9**, e1003518 (2013).

45. Yang, X. et al. Histone acetyltransferase 1 promotes homologous recombination in DNA repair by facilitating histone turnover. *J. Biol. Chem.* **288**, 18271-82 (2013).

46. Mersfelder, E.L. & Parthun, M.R. Involvement of Hat1p (Kat1p) catalytic activity and subcellular localization in telomeric silencing. *J. Biol. Chem.* **283**, 29060-8 (2008).

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

48. Miao, W. et al. HSP90 inhibitors stimulate DNAJB4 protein expression through a mechanism involving N^6-methyladenosine. *Nat. Commun.* **10**, 3613 (2019).

49. Sethi, A. et al. Supervised enhancer prediction with epigenetic pattern recognition and targeted validation. *Nature Methods* **17**, 807-814 (2020).

50. Zhang, J. et al. An integrative ENCODE resource for cancer genomics. *Nature Communications* **11**, 3696 (2020).

51. Payer, L.M. & Burns, K.H. Transposable elements in human genetic disease. *Nature Reviews Genetics* **20**, 760-772 (2019).

52. Xu, W. et al. Dynamic control of chromatin-associated m(6)A methylation regulates nascent RNA synthesis. *Mol Cell* **82**, 1156-1168.e7 (2022).

53. Louloupi, A., Ntini, E., Conrad, T. & Ørom, U.A.V. Transient N-6-methyladenosine transcriptome sequencing reveals a regulatory role of m6A in splicing efficiency. *Cell Reports* **23**, 3429-3437 (2018).

54. Wang, Y. et al. m6A methylation analysis of differentially expressed genes in skin tissues of coarse and fine type Liaoning Cashmere Goats. *Frontiers in Genetics* **10**(2020).

55. Wei, J. et al. FTO mediates LINE1 m(6)A demethylation and chromatin regulation in mESCs and mouse development. *Science* **376**, 968-973 (2022).

56. Feng, Z., Li, Q., Meng, R., Yi, B. & Xu, Q. METTL3 regulates alternative splicing of MyD88 upon the lipopolysaccharide-induced inflammatory response in human dental pulp cells. *J. Cell. Mol. Med.* **22**, 2558-2568 (2018).

57. Liu, N. & Pan, T. N6-methyladenosine–encoded epitranscriptomics. *Nat Struct Mol Biol* **23**, 98-102 (2016).

58. Bannister, A.J. & Kouzarides, T. Regulation of chromatin by histone modifications. *Cell Res.* **21**, 381-95 (2011).

59. Lebel, E.A., Boukamp, P. & Tafrov, S.T. Irradiation with heavy-ion particles changes the cellular distribution of human histone acetyltransferase HAT1. *Mol. Cell. Biochem.* **339**, 271-84 (2010).

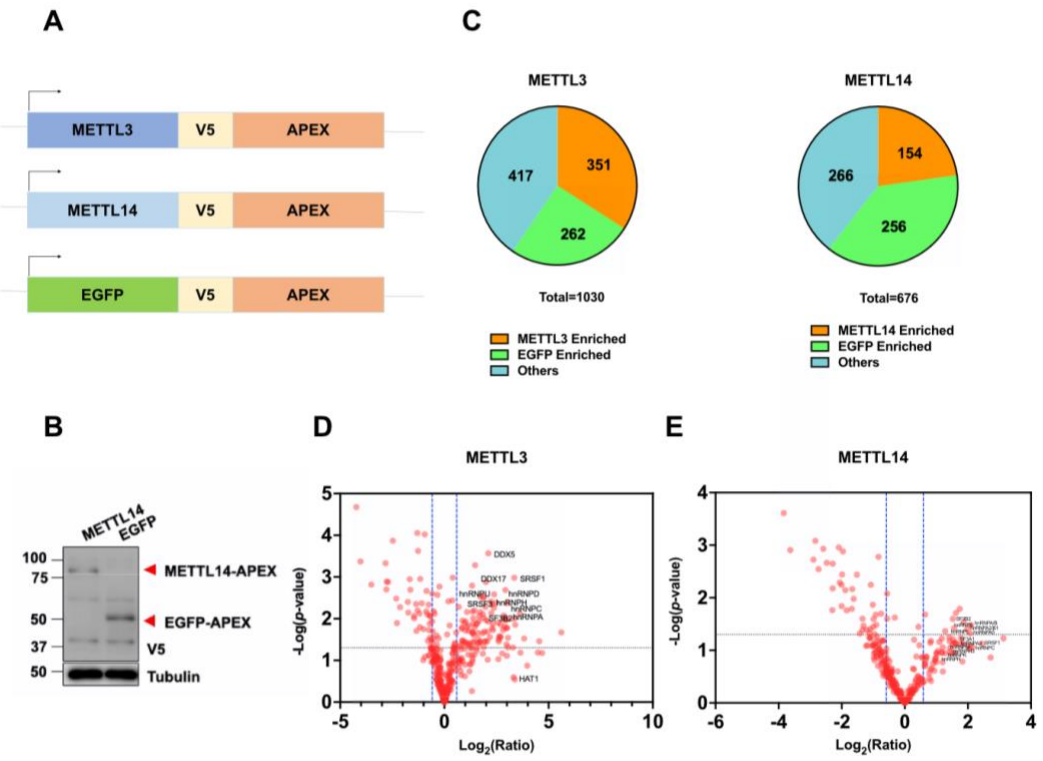


Figure S5.1 APEX-based proximity proteomics of METTL3 and METTL14.

(A) A schematic diagram showing the constructs expressing METTL3-, METTL14-, or EGFP-APEX fusion proteins with a V5 tag in between. (B) Western blot analysis showing the expression of METTL14- and EGFP-APEX in HEK293T cells. (C) Pie chart showing the number of METTL3-, METTL14-, or EGFP-enriched proteins in HEK293T cells. (D-E) Volcano plots for METTL3 or METTL14 versus EGFP. The horizontal gray lines represent a *p*-value of 0.05.

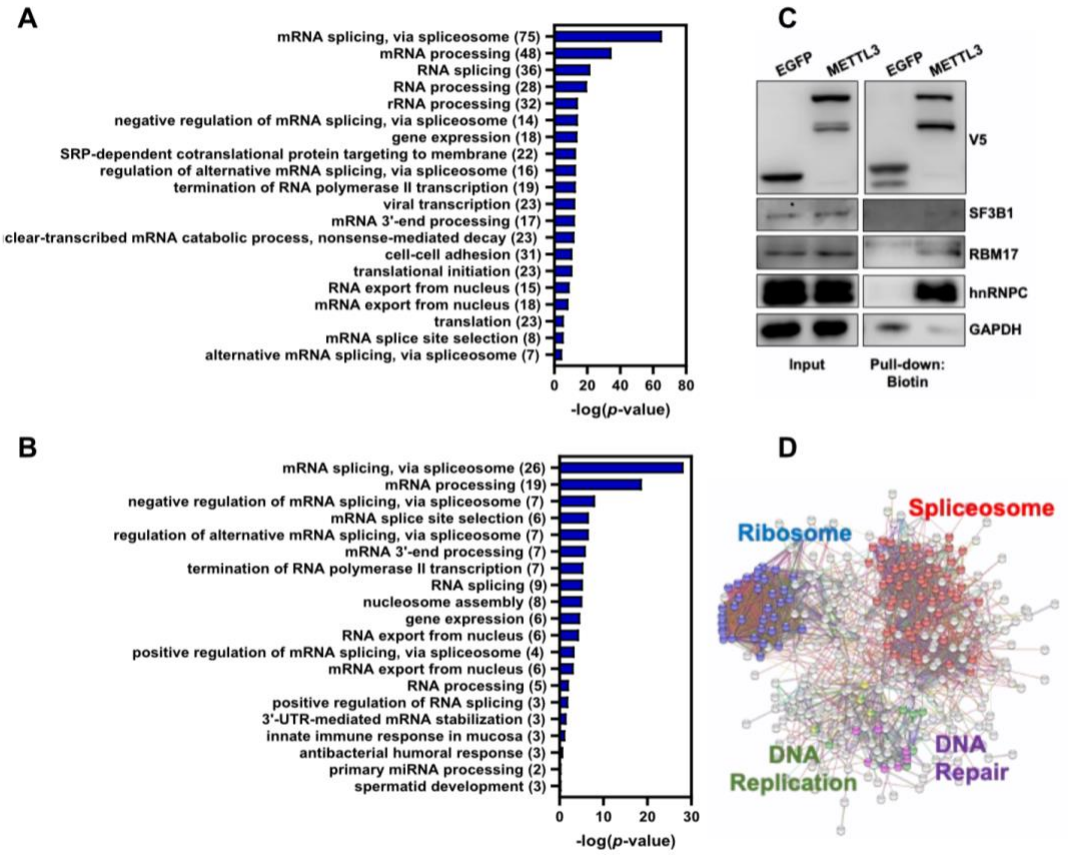


Figure S5.2 Gene ontology analysis of METTL3 and METTL14 proximity proteomes. (A-B) Gene Ontology (GO) enrichment analysis for the biological processes of METTL3- or METTL14-enriched proteins. (C) APEX-based proximity labeling of EGFP and METTL3 followed by Western blot analysis for the validation of selected splicing factors. (D) STRING protein-protein interaction diagram of METTL3. KEGG pathway enrichment analysis revealed that METTL3 is in close proximity to protein clusters involved in splicing, translation, and protein degradation, and DNA replication.

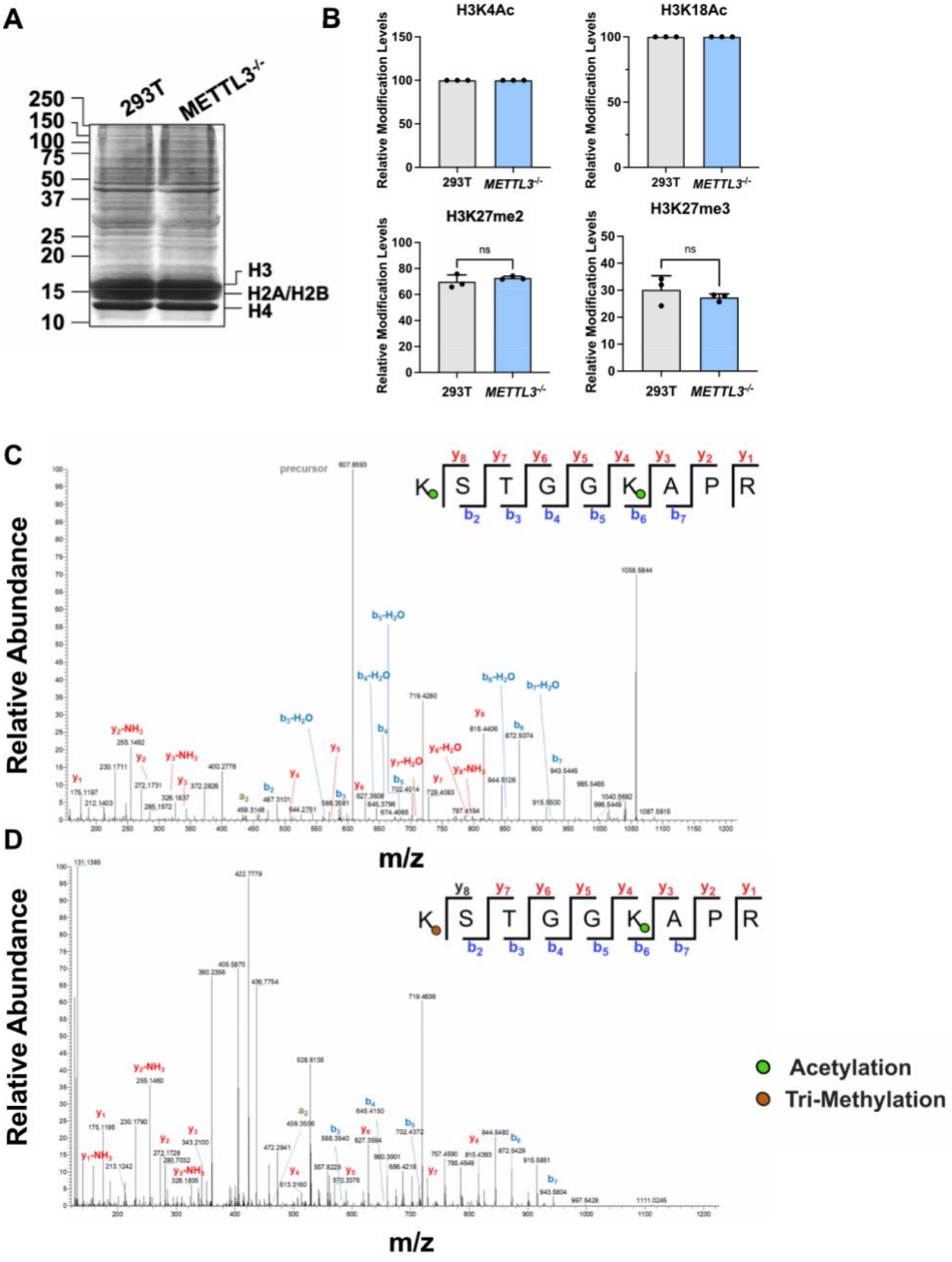


Figure S5.3 Histone post-translational modifications in the isogenic *METTL3*$^{-/-}$ cells. (A) Coomassie blue staining showing the enrichment of histone proteins. (B) Bar charts showing the percentages of histone H3 modifications in HEK293T and the isogenic

METTL3$^{-/-}$ cells. Histones were extracted from the cells for TMT-based quantitative proteomic analysis. The intensities of the modified peptides were normalized against to the total intensities of that peptide sequence in unmodified and all of its modified forms. (C-D) MS/MS of the H3K9Ac- and H3K9me3-containing peptides.

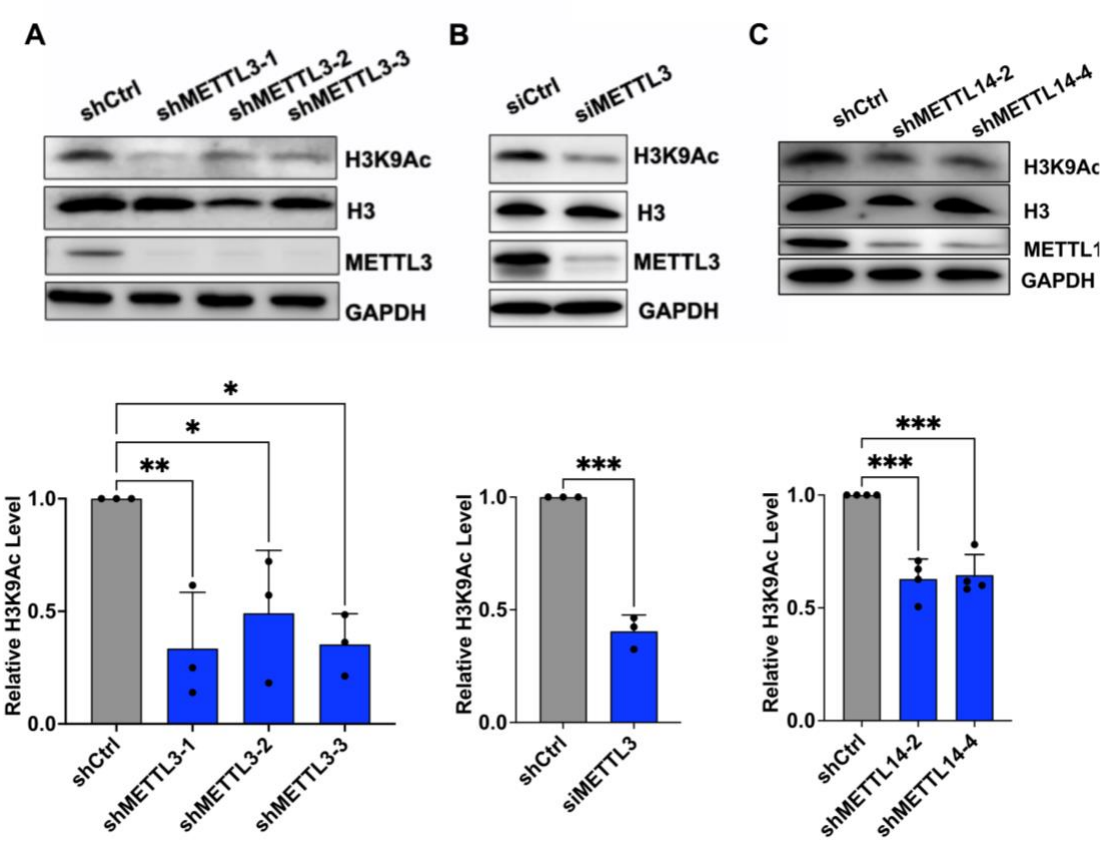


Figure S5.4 METTL3 m^{6}A methyltransferase complex regulates H3K9Ac in differentiated cells.

Western blot analysis and the corresponding quantification results of H3K9Ac in (A) HeLa cells knocked-down of METTL3 by shRNA or (B) HepG2 cells knocked-down of METTL3 by siRNA. (C) Western blot analysis and the corresponding quantification results of H3K9Ac in HEK293T cells where *METTL14* gene was knocked down with shRNA. The signal intensities of H3K9Ac were first normalized against H3, and the resulting ratio was further normalized against that of scrambled control (shCtrl) for shRNA or siCtrl for siRNA. All data represent mean ± SD of results obtained from three biological replicates.

The p values were calculated using one-way ANOVA (*, $0.01 < p < 0.05$; **, $0.001 < p < 0.01$; ***, $p < 0.001$).

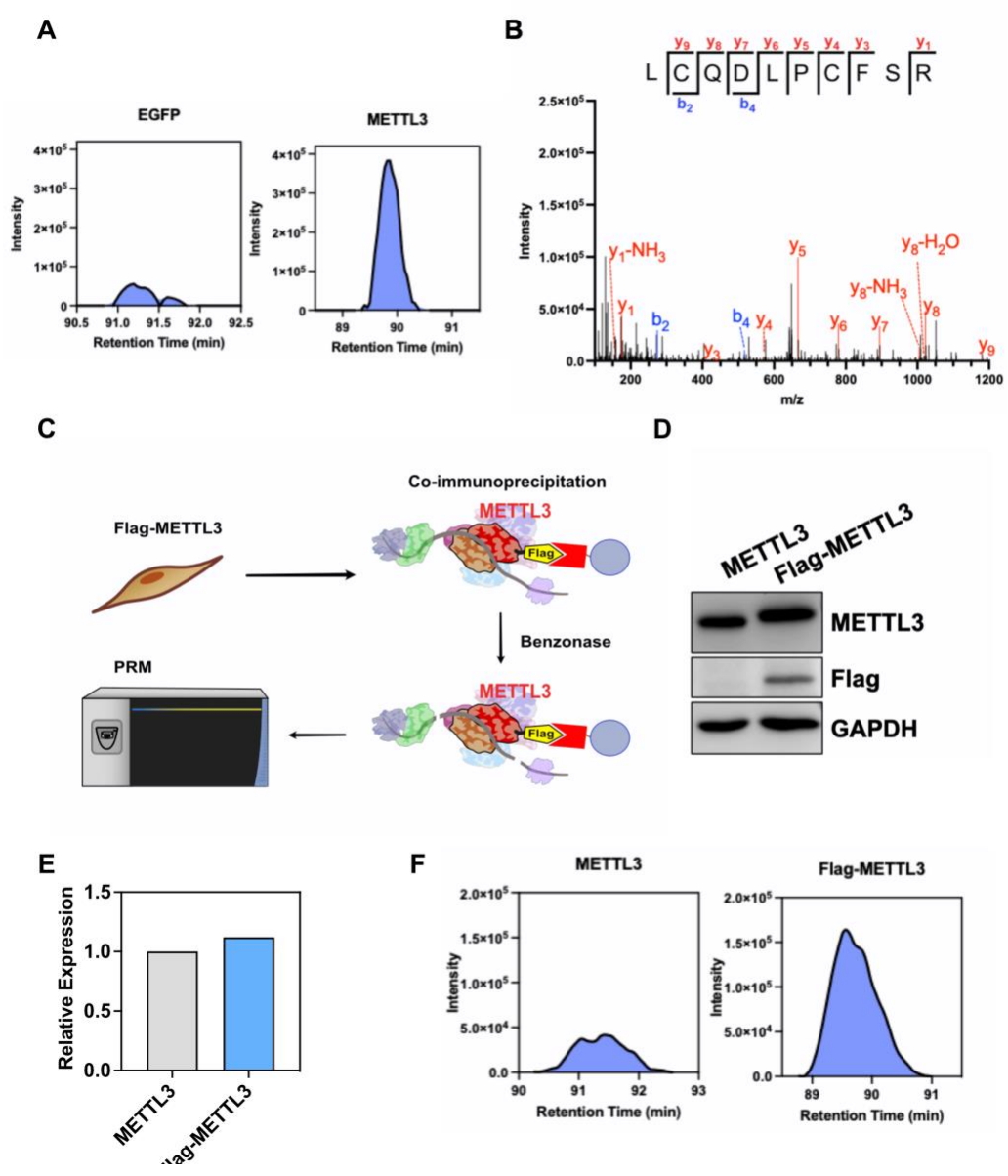


Figure S5.5 Physical interaction between METTL3 and HAT1 in HEK293T cells.

Extracted-ion chromatograms (A) and MS/MS (B) of a unique HAT1 peptide from EGFP and METTL3 proximity proteome. (C) A schematic diagram illustrating the experimental design of co-immunoprecipitation of Flag-tagged METTL3 followed by LC-PRM analysis. Cells expressing METTL3 without Flag tag were used as control. (D-E) Western blot

analysis and the corresponding quantification results showing the comparable levels of expression METTL3 with or without Flag tag. The signal intensities of METTL3 were first normalized against that of GAPDH, and the resulting ratio was further normalized against that of the control (METTL3). (F) Extracted-ion chromatograms of the HAT1 unique peptide from co-immunoprecipitation of Flag-tagged METTL3 followed by LC-PRM analysis. Cells expressing a comparable level of untagged METTL3 (METTL3) were used as control (Figure S4C).

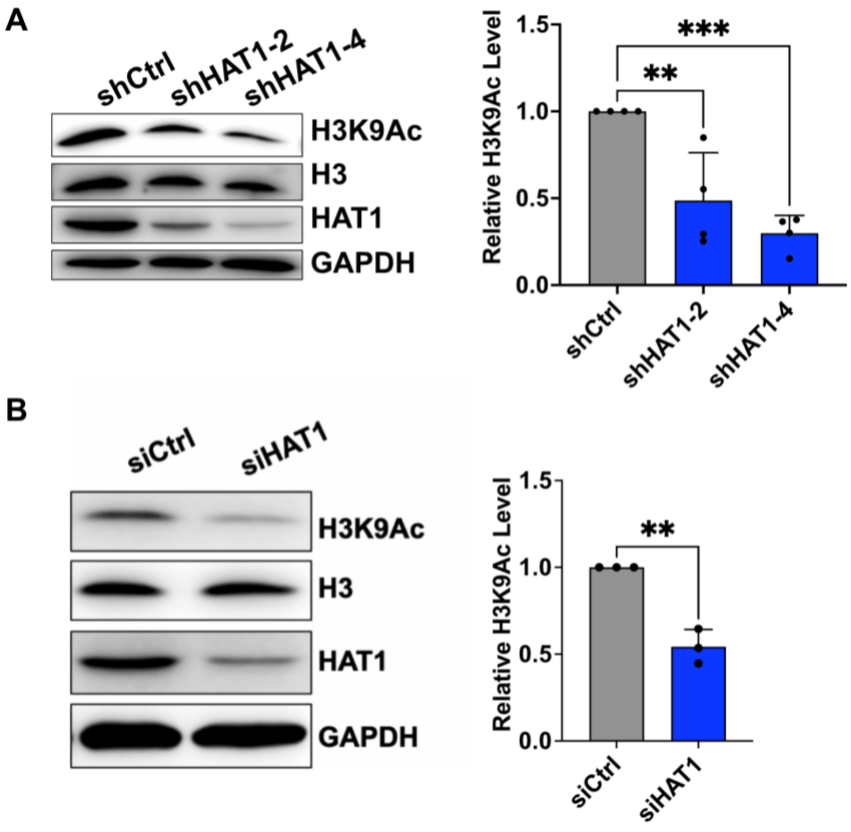


Figure S5.6 HAT1 is involved in H3K9Ac.

Western blot analysis and the corresponding quantification results of H3K9Ac in (A) HEK293T cells where *HAT1* gene was knocked down with shRNA or (B) HepG2 cells where HAT1 gene was knocked down by using a mixture of three target-specific siRNAs (B). The signal intensities of H3K9Ac were first normalized against that of H3, and the resulting ratio was further normalized against that of scrambled control (shCtrl) for shRNA or siCtrl for siRNA. (The *p*-values were calculated using one-way ANOVA (**, $0.001 < p < 0.01$; ***, $p < 0.001$).

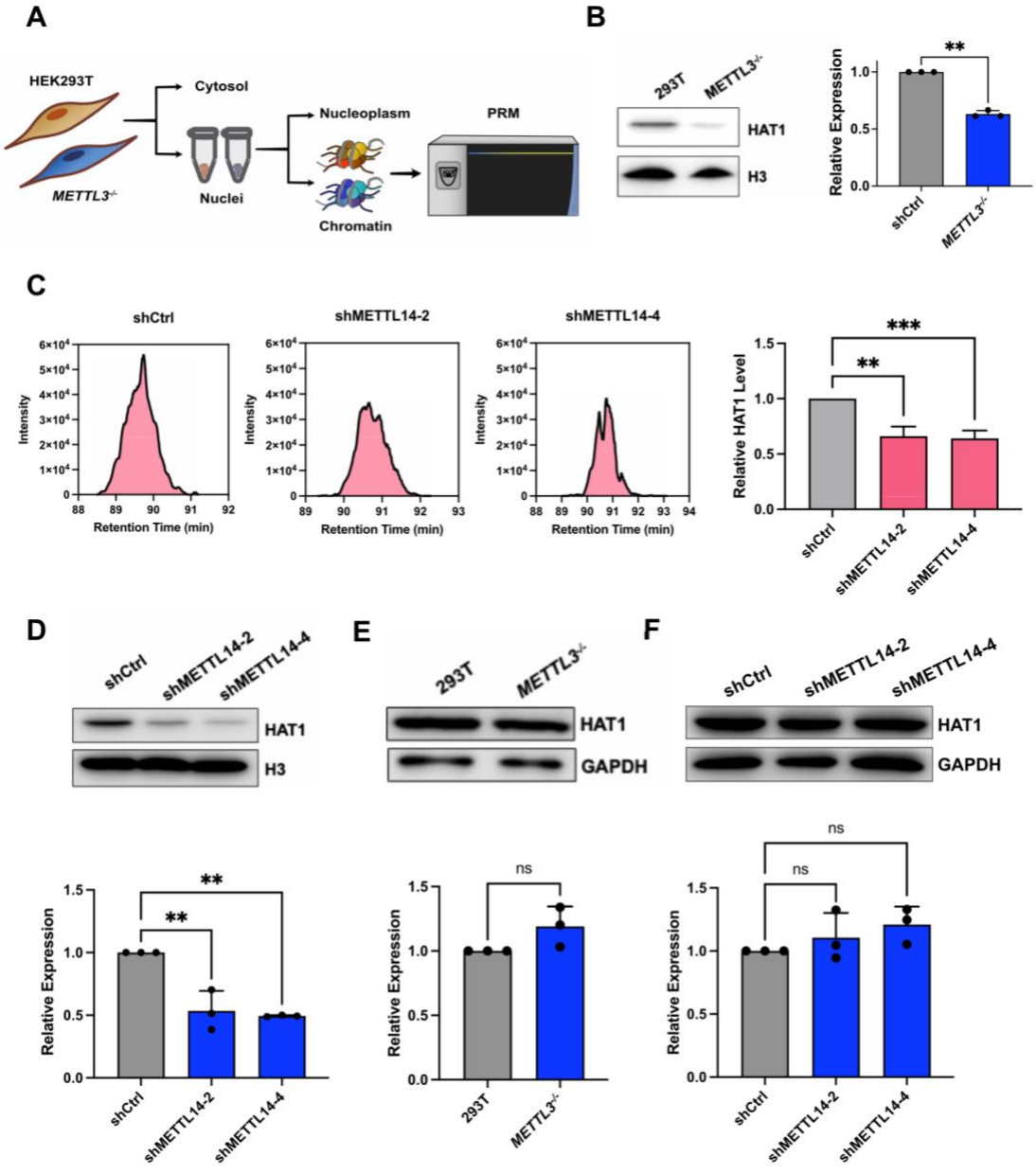

A
HEK293T
METTL3-/-
Cytosol
Nuclei
Nucleoplasm
Chromatin
PRM
B
293T
METTL3-/-
HAT1
H3
Relative Expression
shCtrl
METTL3-/-
C
shCtrl
shMETTL14-2
shMETTL14-4
Intensity
Retention Time (min)
Relative HAT1 Level
D
HAT1
H3
E
293T
METTL3-/-
HAT1
GAPDH
F
HAT1
GAPDH
ns

Figure S5.7 METTL3-containing m^{6}A methyltransferase regulates chromatin localization of HAT1.

(A) Schematic diagram of subcellular fractionation followed by LC-PRM analysis. (B) Western blot analysis and the corresponding quantification results of HAT1 to validate decreased HAT1 in the chromatin fraction of *METTL3*$^{-/-}$ cells. (C) Extracted-ion chromatograms and the corresponding quantification results of the HAT1 unique peptide in the chromatin fractions extracted from METTL14 knock-down cells followed by LC-PRM analysis. (D) Western blot analysis and the corresponding quantification results of HAT1 to validate decreased HAT1 in the chromatin fraction in METTL14 knock-down cells. (E-F) Western blot analysis of HAT1 from total lysates of *METTL3*$^{-/-}$ or METTL14 knock-down cells. All intensities of H3K9Ac were first normalized against that of H3, and the resulting ratios were further normalized to the of parental HEK293T or scrambled control (shCtrl) for *METTL3*$^{-/-}$ and METTL14 knock-down cells. The data represent mean ± SD of results obtained from three biological replicates. The *p* values were calculated using two-tailed Welch's *t*-test (**, $0.001 < p < 0.01$; ***, $p < 0.001$; ns, $0.05 < p$).

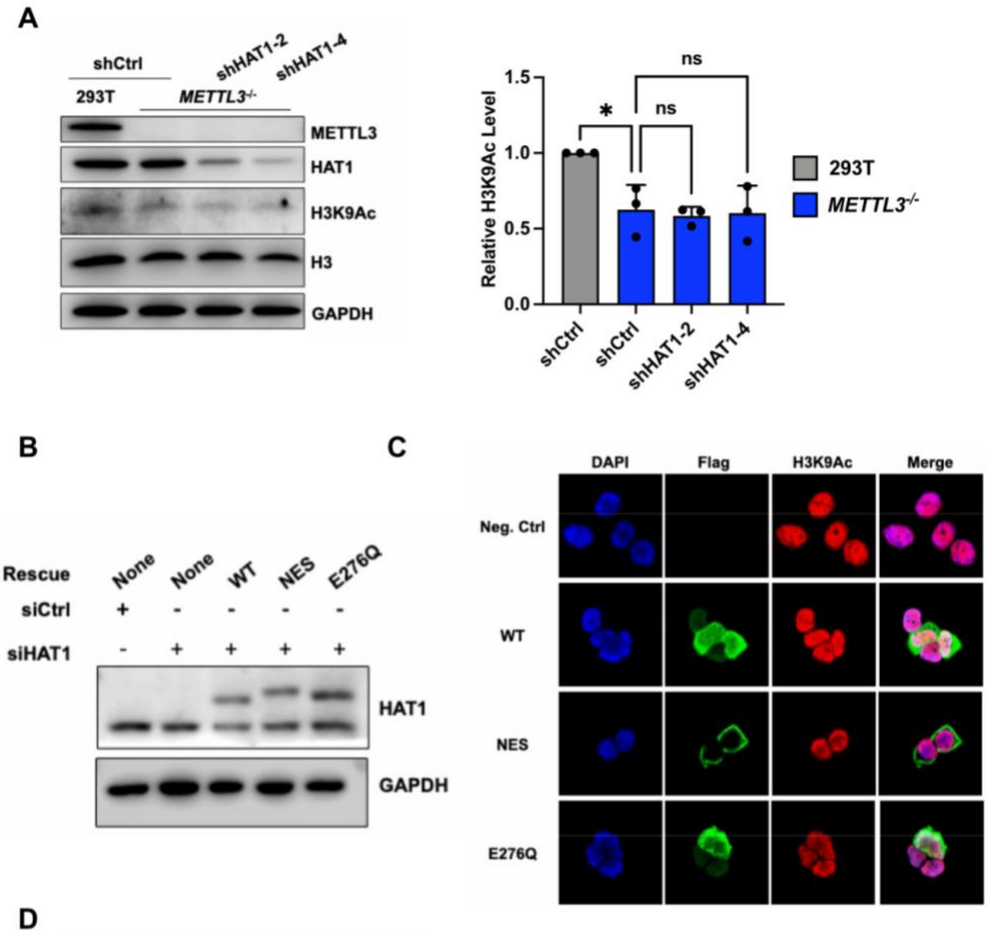

A
shCtrl
shHAT1-2
shHAT1-4
293T
METTL3-/-
METTL3
HAT1
H3K9Ac
H3
GAPDH
Relative H3K9Ac Level
1.5
1.0
0.5
0.0
*
ns
ns
293T
METTL3-/-
shCtrl
shCtrl
shHAT1-2
shHAT1-4
B
Rescue
None
None
WT
NES
E276Q
siCtrl
siHAT1
HAT1
GAPDH
C
DAPI
Flag
H3K9Ac
Merge
Neg. Ctrl
WT
NES
E276Q
D
BSA (ug)
0.1
0.2
0.4
Cell lysate
Flow through
Flag-HAT1
Elution 1
Elution 2
Elution 3

Figure S5.8 HAT1-mediated H3K9Ac is modulated by METTL3.

(A) Western blot analysis and the corresponding quantification results of H3K9Ac in *METTL3*$^{-/-}$ cells that were further knocked-down of HAT1. The signal intensities of H3K9Ac were first normalized against that of H3, and the resulting ratios were further normalized against that of the HEK293T cells. All data represent mean ± SD of results obtained from three biological replicates. The *p* values were calculated one-way ANOVA (C-E) (*, $0.01 < p < 0.05$; ns, $p > 0.05$). (B) Western blot analysis showing the comparable ectopic expression of different variants of HAT1α, including wild-type (WT), the one fused with the nuclear export sequence of PKI (NES), or the enzymatically inactive mutant (E276Q). (C) Immunofluorescence staining to show the subcellular localization of ectopically expressed HAT1 variants, including wild-type HAT1α, enzymatically inactive mutant (E278Q), HAT1α fused with the NES of PKI (NES). (D) Coomassie Blue staining showing the purity of purified Flag-tagged HAT1 used for *in vitro* acetylation experiment.

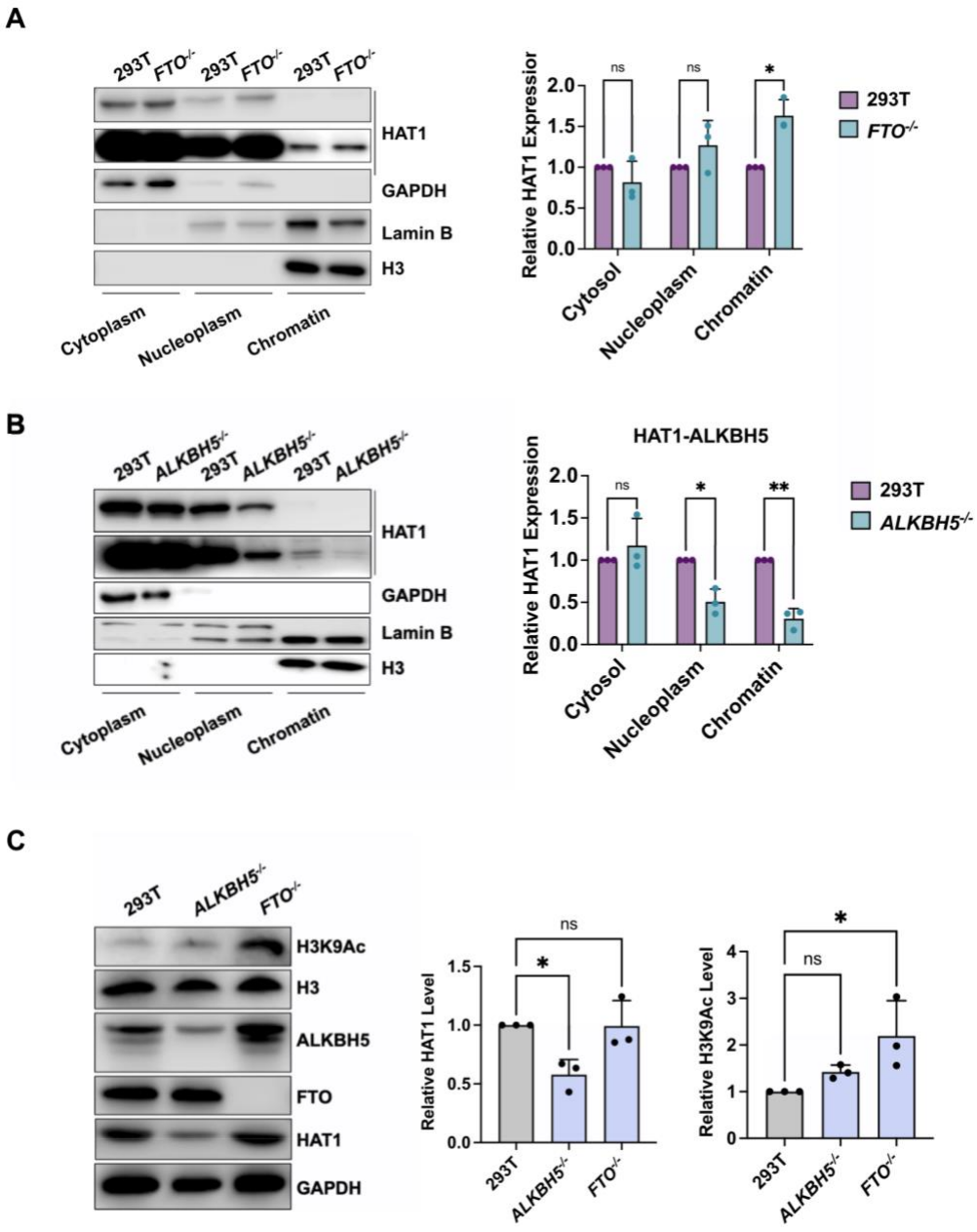


Figure S5.9 m^{6}A demethylase FTO regulates H3K9Ac in HEK293T cells.

(A-B) Western blot analysis and the corresponding quantification results of HAT1 in different subcellular fractions in cells depleted of demethylases, including ALKBH5 and FTO. The signal intensities of HAT1 were first normalized against the control (GAPDH for cytoplasm, Lamin B for nucleoplasm, and H3 for chromatin), and the resulting ratio

was further normalized against that of parental HEK293T cells. (C) Western blot analysis of H3K9Ac in *FTO*$^{-/-}$ and *ALKBH5*$^{-/-}$ HEK293T cells. The H3K9Ac intensity was first normalized against H3, and the resulting ratio was further normalized against to that of parental HEK293T cells. The data represent mean ± SD of results obtained from three biological replicates. The *p* values were calculated using two-tailed Welch's *t*-test (*, $0.01 < p < 0.05$; ns, $0.05 < p$).

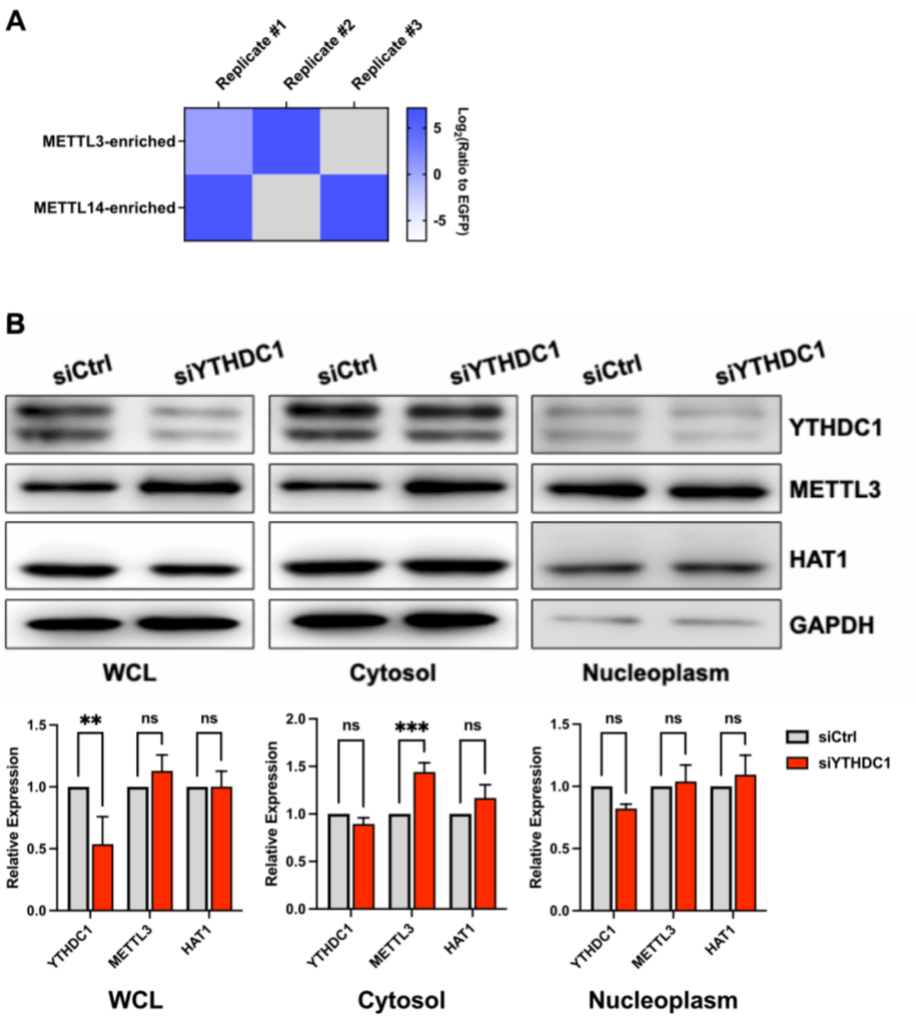


Figure S5.10 YTHDC1 promotes chromatin occupancy of METTL3 and HAT1 to install acetylation at H3K9.

(A) A heatmap showing the enrichment of YTHDC1 in the METTL3 and METTL14 proximity proteome. Missing replicates are indicated in gray. (B) Western blot analysis and the corresponding quantification results of whole cell lysate (WCL), cytoplasm, and nucleoplasm fractions of cells depleted of YTHDC1.

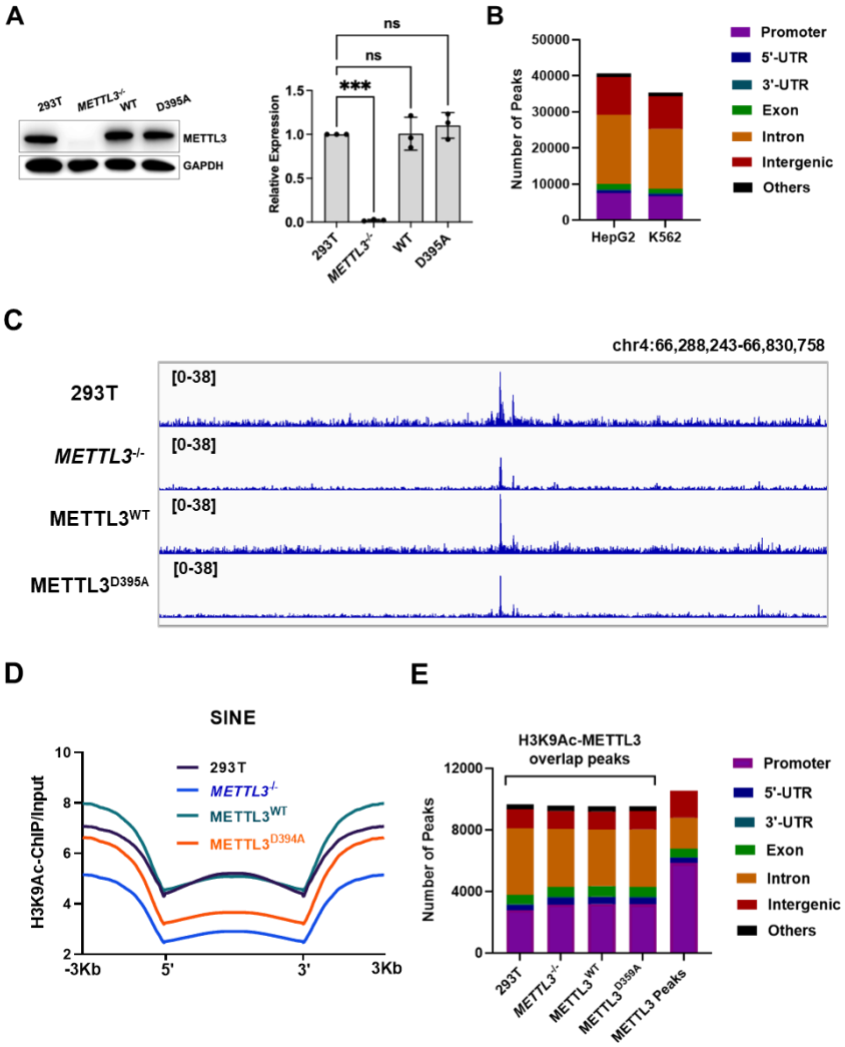


Figure S5.11 The distribution of H3K9Ac in the human genome.

(A) Western blot analysis showing the comparable levels of expression of METTL3 in HEK293T cells used for the H3K9Ac ChIP-seq experiments. The *METTL3*$^{-/-}$ cells were complemented with wild-type or catalytically inactive METTL3. (B) Bar plots showing the H3K9Ac distribution in HepG2 and K562 cells. (C) IGV plot showing H3K9Ac enrichment in intergenic regions of 293T, *METTL3*$^{-/-}$, METTL3^{WT} or METTTL3^{D395A} rescued cell lines (y axis indicates ratios of IP over input). (D) Aggregation plots showing the distribution of H3K9Ac in SINE elements. (E) Bar plots showing the distribution of

H3K9Ac-METTL3 overlap peaks in 293T, *METTL3*$^{-/-}$, METTL3^{WT} or METTL3^{D395A} rescued cell lines.

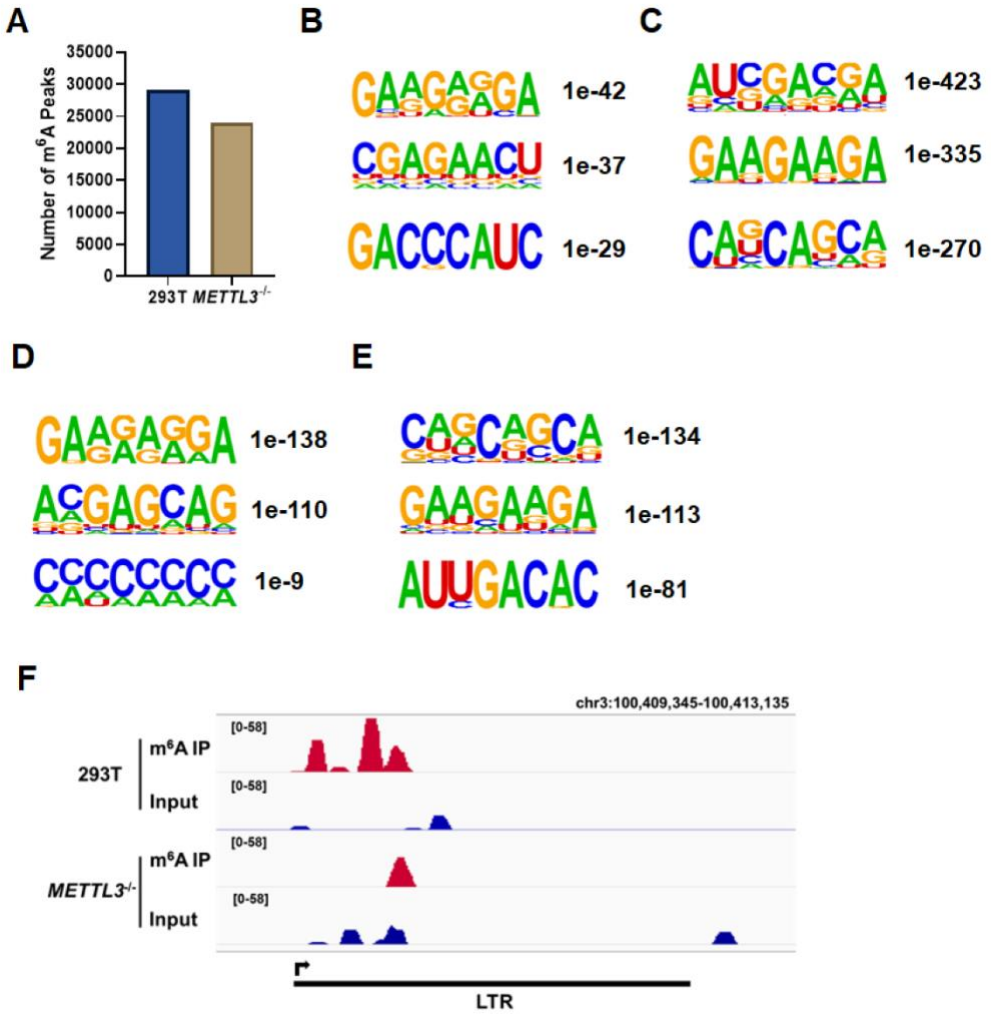


Figure S5.12 m^{6}A levels on caRNA are different in parental and *METTL3*$^{-/-}$ HEK293T cells.

(A) Number of m^{6}A peaks identified in WT and *METTL3*$^{-/-}$ cells. (B-C) The most enriched motifs of caRNA m^{6}A peaks identified in wild-type (B) and *METTL3*$^{-/-}$ HEK293T cells in this study (C). (D-E) The most enriched motifs of caRNA m^{6}A peaks identified in wild-type (D) and METTL3-KD HEC-1-A cells in a previous study[1] (E). (F) IGV plots showing m^{6}A enrichment close to 5′ end of LTR transcript, which is depleted in *METTL3*$^{-/-}$ cells.

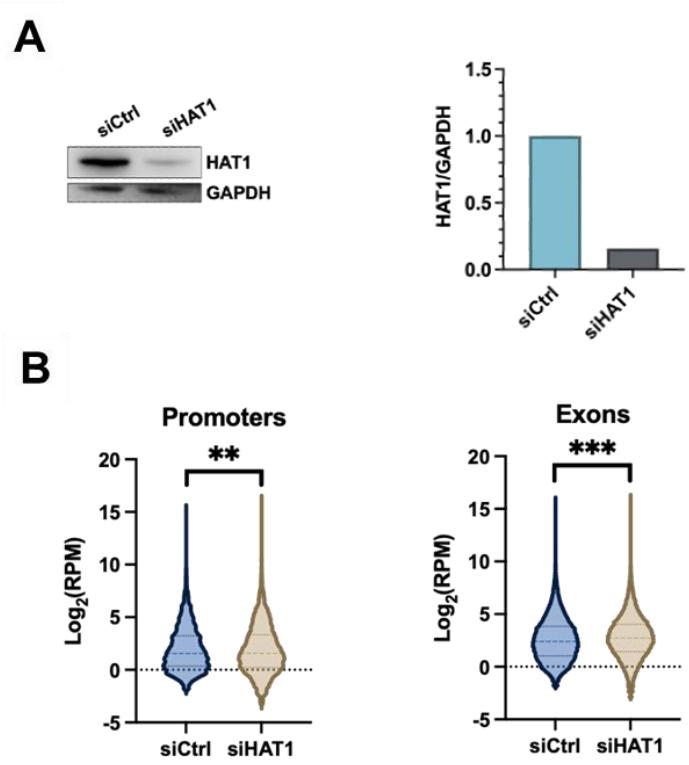


Figure S5.13 HAT1 regulates RNA expression.

(A) Western blot analysis showing the knock-down efficiency of HAT1 in HEK293T cells used for the RNA-seq experiment. (B) Violin plots showing the RNA densities in promoter region and exon regions in siCtrl and siHAT1 cells. RPM indicates reads per million. The *p* values were calculated using two-tailed Welch's *t*-test, **, $0.001 < p < 0.01$; ***, $p < 0.001$.

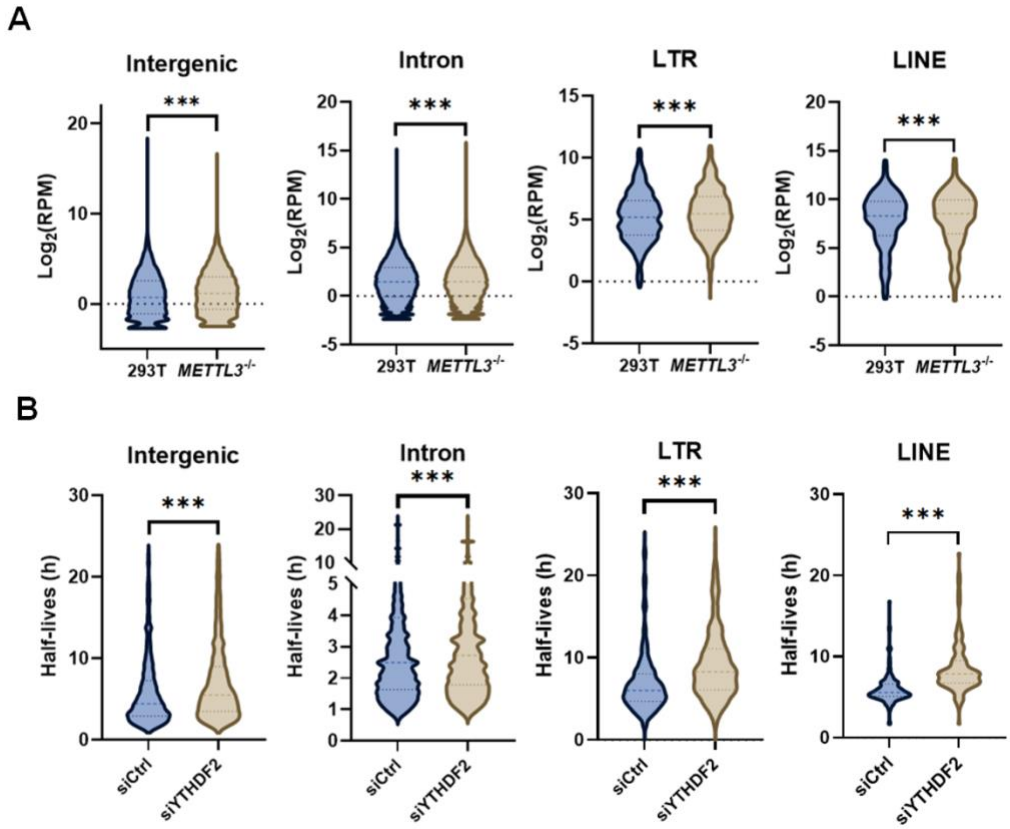


Figure S5.14 The depletion of METTL3 leads to increased transcript levels in intergenic, intron, LTR and LINE regions, which could be due to the extension of RNA half-lives in these regions.

(A) RNA densities in intergenic region, intron region, LTR and LINE elements of parental and *METTL3*$^{-/-}$ HEK293T cells. RPM indicates reads per million. (B) Half-lives of RNAs in intergenic region, intron region, LTR and LINE elements of siCtrl and siYTHDF2 cells. The *p* values were calculated using two-tailed Welch's *t*-test, ***, $p < 0.001$.

6 Chapter 6. Concluding Remarks

The research presented in this dissertation is mainly divided into two parts, namely the applications of MRM-based targeted proteomics coupled with scheduled LC to profile the expression of small GTPase proteome, with the focus being placed on stem cell differentiations and the involvements of epitranscriptomic regulations in modulating the expression of the entire protein superfamily; and APEX-based proximity labeling to profile the protein-protein interactions of METTL3 and METTL14.

In Chapter 2, we investigated the differential expression of small GTPases during the adipogenesis of murine MSCs. The scheduled LC-MRM method enables the quantification of 55 and 49 small GTPases accompanied with adipogenic differentiation in 3T3-L1 and C3H10T1/2 cells, respectively, in a single LC-MS/MS analysis. More importantly, a role of Rab32 in the adipogenesis was unveiled from the targeted proteomic results, together with follow-up genetic manipulation studies in both cell lines. In addition, the quantitative proteomic results provided the first comprehensive interrogation of the small GTPase superfamily during the adipogenesis of MSCs and documented a subset of small GTPases involved in modulating adipogenic differentiation. The proteomic results lay the foundation for investigating the molecular regulations upon adipogenesis in the future.

Inspired by the discoveries in adipogenesis study, we hypothesized that the small GTPases may also be involved during differentiation of stem cells to other cell types. Thus, in Chapter 3, we employed this LC-MRM method to assess the differential expression of small GTPase proteins during the course of osteogenic differentiation of H9 human

embryonic stem cells (ESCs). We were able to quantify 83 small GTPases, which represent approximately 60% of the human small GTPase proteome. Importantly, the results provide the first comprehensive analysis of the small GTPase superfamily during ESC differentiation in general and osteogenesis in particular.

Among these quantified proteins, we discovered diminished expression of KRAS upon mesenchymal differentiation to mature osteoblast. Biochemical assays together with genetic studies further revealed KRAS as a novel osteogenesis regulator that attenuates the activity of secreted matrix metalloproteinase-9 (MMP9) to regulate the accumulation of extracellular matrix for mineralization, a crucial process of ossification. Moreover, principal component analysis revealed that the expression of autophagy-regulating small GTPases are systemically coordinated during osteogenic differentiation. Together, in Chapters 2 and 3, the results revealed the involvements of small GTPase superfamily in directing the adipogenesis and osteogenesis of stem cells and build a strong foundation for future mechanistic studies of mesenchymal differentiation.

In mRNA, the m^6A methylation is the most abundant internal modifications installed post-transcriptionally and was found to dynamically and reversibly regulate the stability and translational efficiency of the transcripts. In Chapter 4, I investigated how m^6A methylation modulates the expression of small GTPase superfamily. I employed the LC-MRM analysis to explore the differential expression of small GTPases proteins in cells genetically depleted of m^6A reader (YTHDC1, YTHDC2, and YTHDC3), writer (METTL3), and eraser (ALKBH5 and FTO) proteins. Approximately 50% of the small

GTPase proteome in human was quantified across all the CRISPR-Cas9-engineered knockout cells, showing the sensitivity and reproducibility of the method.

The results from the quantitative proteomic experiments revealed that the expression of small GTPases, including several known tumor drivers or suppressors, were substantially perturbed in the cells depleted of m^6A modulators, especially the methyltransferase METTL3 and demethylase ALKBH5. Consistently, through analysis of publicly available data of paired primary/metastatic melanoma and colorectal cancer cells, we found that those small GTPases targeted by METTL3 and ALKBH5 exhibited higher discrepancy between the transcriptome and proteome. Given that small GTPases are difficult to target through pharmacological approaches due to their limited binding pockets and structural homology of the proteins in the superfamily, the results documented in this dissertation afforded an important basis for alternative therapeutic approach to targeting small GTPases for modulating tumor metastasis.

METTL3 comprises the core catalytic subunit of the major m^6A methyltransferase to regulate numerous cellular functions. However, there is so far no systematic study about the protein-protein interaction of METTL3 and its binding partner METTL14 in cells. To advance our knowledge in epitranscriptomic regulations, in Chapter 5, we employed APEX-based proximity labeling followed by LC-MS/MS analysis to assess the proximity proteomes of METTL3 and METTL14 methyltransferases. Interestingly, I identified over 20 histone PTM reader, writer, and eraser proteins to be in close proximity of METTL3, among which HAT1 was found, through co-immunoprecipitation followed by PRM analysis, to interact physically with METTL3. The subsequent cell-based and sequencing

results further revealed that METTL3 methylates chromatin-associated RNAs (caRNAs), of which the methylation could be recognized by nuclear m^6A reader protein YTHDC1 to recruit more METTL3 to amplify the regulation. The METTL3 subsequently promoted the chromatin localization of HAT1 to acetylate H3K9 in chromatin to regulate gene expression.

While the crosstalk was previously reported in mouse ESCs, whether the differentiated cells adapted similar approach to fine-tune gene expression was unclear. Here, we revealed the crosstalk between epigenetic and epitranscriptomic regulations in the differentiated cells and the results provided in-depth mechanism underlying the crosstalk. We anticipate that the METTL3-HAT1 interaction-regulated H3K9Ac of caRNAs is only part of the whole crosstalk between epitranscriptome and epigenome in the differentiated cells, and it will be important to further characterize how diverse epitransctiptomic regulations (e.g., METTL16, another reported m^6A methyltransferase) could reshape epigenetic marks in the human cells to regulate gene expression.

Together, in this dissertation, we presented the application of scheduled LC-MRM analysis for profiling the expression of small GTPases. The results demonstrated the excellent throughput, sensitivity, and reproducibility of the platform in studying the small GTPases during the adipogenesis (Chapter 2) and osteogenesis (Chapter 3) in stem cells, and how the m^6A-modulating proteins alter the expression of the members in this superfamily (Chapter 4). Moreover, we demonstrated the effectiveness of APEX-based proximity labeling in identifying the interactome of a protein of interest with high selectivity. With the APEX-based approach, we revealed a novel crosstalk between

epitranscriptome and epigenome. We envision that the developments of modern proteomics will continue to allow for profiling proteome of interest with robust speed, sensitivity, and selectivity.