UNIVERSITY OF CALIFORNIA
RIVERSIDE

Targeted Quantitative Proteomic Profiling of Small GTPases in Cultured Cells and Tissues

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

Doctor of Philosophy

in Environmental Toxicology

by

Ming Huang

September 2019

Dissertation Committee:
Dr. Yinsheng Wang, Chairperson
Dr. Ryan Julian
Dr. Joseph Genereux

The Dissertation of Ming Huang is approved:

_____

_____

_____

Committee Chairperson

University of California, Riverside

Acknowledgements

This has been a wonderful journey and it just set at a milestone, which would not have been made possible without the incredible persons that shaped me into what I am today.

I wish to express my deepest gratitude and appreciation to my Ph.D. mentor, Dr. Yinsheng Wang, for his untiring guidance, encouragement and support throughout the entire graduate school life. He has always been patient and constructive whenever I need help and support in all my pursuits of research. As a great mentor, he kept instilling in me, and many others, the passion for exploring, the curiosity for learning, the dedication to excellence, and a culture of scientific integrity. Thanks to him, the door to the truly fascinating world of mass spectrometry and proteomics was opened to me. Benefited from the working environment that fostered critical thinking, problem solving and self-motivation, I developed and matured to be a scientist and continue learning to be a better one.

I would also like to acknowledge Dr. Joseph Genereux and Dr. Ryan Julian for being my dissertation committee, and Dr. Cynthia Larive, Dr. Huiwang Ai and Dr. Zhenbiao Yang for serving as my guidance committee, giving me precious suggestions and comments on my research.

I gratefully acknowledge financial support from the Environmental Toxicology Graduate Program, National Institutes of Health (NIH), and NIH National Research Service Award (NRSA) Institutional Training Grant (T32). My dissertation work would not have been made possible without their support.

My life in Riverside would not have been so pleasant without the friendship and help from current and former members of the Wang group. I would like to thank Dr. Yongsheng Xiao and his wife Dr. Lei Guo for their help and guidance with proteomics when I first started my experiments. Especially, I would like to thank Dr. Xiaogang Jiang, Dr. Preston Williams, Dr. Yang Yu, Dr. Rong Cai, Yuxiang Cui and Gwendolyn Gonzales for their tremendous help and instructions on instrument maintenance and troubleshooting. I owe a debt of gratitude for Dr. Nathan Price in his immense help in revising the dissertation and friendship. I would also like to offer my profound thanks to Dr. Xiaoxia Dai, Dr. Lin Li and Dr. Xiaochuan Liu for sharing their experience in guiding my biological experiments. A big thanks must go to our lab manager Shuli Zhai for her great help and efforts in organizing our research environment. Besides work, a special thanks goes to my "coffee buddies": Dr. Yang Yu, Dr. Pengcheng Wang, Drs. Nicole and Preston Williams, Y. C., Jiabin Wu, and Jun Yuan. Another special thanks go to my "badminton buddies": Dr. Hua Du, Dr. Jun Wu, Y. C., and Zi Gao. I wish to give my sincere thanks to Dr. Lijuan Fu, Dr. Shuo Liu, Dr. Ji Jiang, Dr. Tianlu Wang, Dr. Xiaomei He, Dr. Kailin Yu, Xuejiao Dong, David Bade, Tianyu Qi, Jiekai Yin, Ying Tan, Su Guo and Feng Tang for their consistent support and friendship. Finally, I would like to extend my gratitude to all other former and current group members: Dr. Qianqian Zhai, Dr. Debin Ji, Dr. Tao Bing, Dr. Qian Cai, Dr. Changjun You, Dr. Jiapeng Leng, Dr. Yuxiang Sun, Eric Stephens, Zi Wang, Weili Miao, Lok-Ming Tam, Yenyu Yang and Yinan Wang.

My heartfelt thanks to other friends I was so lucky to have met in Riverside: Dr. Cui Zhang, Dr. Jiayi Chen, Dr. Qing Tang, Dr. Wenrong He, Dr. Yingnan Hou, Dr. Shiliu Tian,

Dr. Weimin Zhou, Dr. Zhisheng Lin, Ji Wang, Andy Smith, Anqi Yu, Tushar Jain, and Md Abid Hasan. It was you that made my graduate life filled with beautiful moments to cherish in this unique journey of my life.

At the end, I want to give thanks to my parents, sister, brother, sister-in-law and parents-in-law, for their love and support that I feel forever indebted to. There are no proper words to convey my gratitude and love for them. To another mass spectrometrist in my family − Dr. Zhihao Yu – thanks for always standing by my side as my husband/soulmate, my closest friend, my peer mentor, my reliable travel buddy, and my lifelong badminton and workout partner. Long-distance relationship was tough, but distance means so little when you mean so much in letting me grow stronger, thrive and become my better self.

Acknowledgements of Previously Published Materials

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

Chapter 2: Huang, M., Qi, T. F., Li, L., Zhang, G., Wang, Y. "A Targeted Quantitative Proteomic Approach for Examining the Reprogramming of Small GTPases during Melanoma Metastasis". Cancer Research 2018, 78, 5431-5445.

Chapter 3: Huang, M.; Wang, Y., "Roles of Small GTPases in Acquired Tamoxifen Resistance in MCF-7 Cells Revealed by Targeted, Quantitative Proteomic Analysis". Anal Chem 2018, 90 (24), 14551-14560.

Chapter 4: Huang, M.; Wang, Y., "Targeted Quantitative Proteomic Approach for Probing Altered Protein Expression of Small GTPases Associated with Colorectal Cancer Metastasis". Anal Chem 2019, 91 (9), 6233-6241.

Dedication

To my parents,

Guorong Huang and Dai Liang,

who granted me knowledge and strength along the way,

with unconditional love;

to my husband Zhihao Yu,

who never stops loving, inspiring and supporting me.

I love you all.

ABSTRACT OF THE DISSERTATION

Targeted Quantitative Proteomic Profiling of Small GTPases in Cultured Cells and Tissues

by

Ming Huang

Doctor of Philosophy, Graduate Program in Environmental Toxicology
University of California, Riverside, September 2019
Dr. Yinsheng Wang, Chairperson

Mass spectrometry (MS)-based bottom-up proteomics and quantitative proteomic labeling strategies have led to unprecedented insights into systems biology and provided invaluable resources as a multifaceted analytical tool. We have utilized such techniques to analyze small GTPases of the Ras superfamily, which represent a class of crucial signaling molecules in cells, and the aberrant regulation of their expressions is implicated with various types of human diseases. In this dissertation, I report the development and applications of novel targeted quantitative proteomic methods for high-throughput and reproducible profiling of small GTPases in cultured human cells and patient-derived brain tissues that carry disease-related changes.

In Chapter 2, I describe the development of a novel scheduled multiple-reaction monitoring (MRM)-based targeted quantitative proteomic method, in conjunction with stable isotope labeling by amino acids in cell culture (SILAC) for the quantification of more than 90 small GTPases in the paired primary/metastatic melanoma cell lines. The data reveal previously unrecognized roles of RAB38 in promoting melanoma metastasis *in vitro*.

In Chapter 3, the established scheduled MRM-based method was further applied to assess the differential expression of small GTPases in wild-type MCF-7 and the paired tamoxifen-resistant breast cancer cells. The method facilitated robust quantification of 96 small GTPases, among which down-regulation of RAB31 was analyzed further and demonstrated to play a role in the development of acquired tamoxifen resistance.

In Chapter 4, we extended the use of the scheduled MRM method to comprehensively investigate the differential expression of small GTPases in paired primary/metastatic

colorectal cancer cell (CRC) lines SW480 and SW620. With this approach, 83 small GTPases were robustly quantified, leading to the identification of SAR1B as a potential suppressor for CRC metastasis. We also showed that diminished SAR1B expression could stimulate epithelial–mesenchymal transition (EMT), thereby promoting motility and *in vitro* metastasis of SW480 cells.

In Chapter 5, I describe the development of a novel targeted quantitative proteomic assay based on MRM and the use of crude synthetic stable isotope-labeled (SIL) peptides as internal standards (IS) and surrogate standards (SS). By using this approach, we quantified ~80 small GTPases from lysates of frontal cortex from post-mortem Alzheimer's disease (AD) patient brain tissue samples. The method displayed excellent throughput, sensitivity and reproducibility. Furthermore, we observed that the protein expression levels of Rab3A/C, Rab4A/B and Rab27B proteins, which are involved with synaptic and secretory vesicles, increased with degree of disease severity. The MRM quantification results were further verified by Western blotting.

Table of Contents

Table of Figures

Chapter 1 Introduction

1.1 General Overview

The past two decades have witnessed increasingly widespread applications of mass spectrometry (MS)-based proteomics in systems biology by facilitating accurate, high-throughput and specific interrogation of either, the entire or a subset, of the proteome from biological samples.[1, 2] This boom in MS-based proteomics is largely due to the development of liquid chromatography-tandem mass spectrometry (LC-MS/MS), which is the method of choice for large-scale identification and quantification of proteins. Among the various proteomic techniques, bottom-up or shotgun proteomics is the most commonly used, where large proteins are proteolytically digested into mixture of peptides for downstream LC-MS/MS analysis.[3] In this chapter, common ion detection methods including shotgun proteomics (also known as discovery proteomics) using data-dependent acquisition (DDA), data-independent acquisition (DIA) and targeted proteomics using multiple-reaction monitoring (MRM; also referred to as selected reaction monitoring, SRM) as well as parallel-reaction monitoring (PRM) will be discussed.

I will move on to review common labeling strategies in quantitative proteomics, with specific emphasis on the following: label-free quantification (LFQ), chemical labeling approaches such as isotope-coded affinity tag (ICAT), isobaric tag for relative and absolute quantitation (iTRAQ), tandem mass tags (TMT); metabolic labeling approach such as stable isotope labeling by amino acids in cell culture (SILAC); and lastly, absolute quantification of proteins (AQUA) using stable isotope-labeled (SIL) internal standards.

Next, I will elaborate on the classification, biological functions and public health relevance of GTP-binding proteins and small GTPases. At last I will detail the major focus of this dissertation, i.e. the development and applications of novel targeted quantitative proteomic methods in studying small GTPases of the Ras superfamily in cultured human cancer cells and patient-derived tissues, together with investigation of the roles of small GTPases in several human diseases including melanoma, breast cancer, colorectal cancer, and Alzheimer's disease by using cell lines or tissue samples as study models.

1.2 Detection Strategies in Bottom-up Proteomics

1.2.1 Data-dependent Acquisition (DDA)

A traditional MS-based bottom-up proteomic method is often performed in a "shotgun" format, where LC-MS/MS is operated in the DDA mode and the N most abundant precursor ions of peptides (N = 5–20) are selected from full or MS1 scans for the subsequent tandem mass spectrometry (MS/MS) analysis (Figure 1.1).[4] The depth of proteome coverage in DDA analysis is largely dependent on sample complexity resulting from the background proteome, and the highest coverage is achieved from the least complex sample. Consequently, protein or peptide fractionation techniques prior to peptide separation on reverse-phase (RP) columns and downstream MS analysis are essential to obtaining a high-degree of protein coverage.

Among the many prefractionation schemes such as strong cation exchange (SCX) or two-dimensional sodium dodecyl sulfate polyacrylamide gel electrophoresis (2D-SDS-PAGE), multi-dimensional protein identification technology (MudPIT) is advantageous over gel-based methods with respects to the ease of sample handling, sensitivity and

dynamic range.[5-7] The concept of MudPIT was built upon the earlier work reported by Link *et al*.,[6, 8] where complex peptide mixtures were loaded onto a biphasic SCX-RP microcapillary column and separated in a multidimensional manner. In this vein, MudPIT allows for an automated and high-throughput online two-dimensional separation pipeline that ultimately leads to large-scale global proteomic studies with simplified sample handling.[6] As a predominant shotgun proteomic approach, MudPIT has achieved remarkable progress in resolving complex biological samples to facilitate protein identification and quantification. While achieving great proteome coverage, shotgun proteomics cannot guarantee that every precursor will be fragmented in each run and therefore, the peptides identified by shotgun proteomic analysis only display 35–60% overlap between technical replicates.[9]

1.2.2 Data-independent Acquisition (DIA)

One of the biggest challenges in discovery or shotgun proteomics is the unbiased identification and precise quantification of protein abundances in highly complex samples, which can consist of thousands of proteins. To address the growing needs for delivering accurate, complete and reproducible large-scale proteomic datasets, an emerging strategy termed sequential window acquisition of all theoretical mass spectra (SWATH-MS) on TripleTOF platforms was proposed by Gillet *et al*. in 2012, and this technology is capable of providing deep proteome coverage with quantitative consistency and accuracy.[10, 11] SWATH-MS is a specific variant of DIA methods, from which DIA generally refers to MS acquisition methods that continuously acquire MS/MS spectra of all ions in the target m/z range in an unbiased fashion (Figure 1.1). In 2015, a similar DIA method implemented on

Orbitrap platforms was termed hyper reaction monitoring (HRM).[12] In DIA mode, all peptide precursors of the entire mass range are sequentially isolated in broader m/z windows and fragmented together in a systematic and unbiased fashion.[10] One advantage provided by such a data collection scheme is a comprehensive and reliable digital fingerprint or data archive of the peptide mixture. However, the tremendously complex fragment ion spectra derived from multiple precursor ions render the data analysis extremely complicated and challenging. To enable unbiased peptide identification and SWATH-MS data interpretation, a commonly used strategy is data interrogation after construction of assay libraries.[13] For example, the SWATH-MS raw data can be re-interrogated over and over using updated peptide spectral libraries. Recent improvements in software tools enable the identification of 10,000 human protein groups from a single DIA run.[14] Compared to conventional DDA workflows, this is an improvement of more than 30% without compromising quantification accuracy.

1.2.3 Multiple-Reaction Monitoring (MRM)

In recent years, targeted proteomics techniques have emerged as a powerful toolset in systems biology to comprehensively study a predefined set of proteins involved in distinct signaling pathways or gene families. Unlike shotgun proteomic methods, the targeted proteomics approach aims to deliver highly reproducible and sensitive measurement of target peptides and thus requires information about the analytes *a priori*. The most widely used targeted technique is MRM, also referred to as selected-reaction monitoring (SRM) performed on a triple quadrupole (QqQ) mass spectrometer. As depicted in Figure 1.1, when LC-MS/MS is performed in the MRM mode, the instrument is programmed to

monitor a pre-selected series of transitions (precursor/product ion pairs).[15] MRM-based protein quantification is superior in that it provides unparalleled ability to characterize and quantify a set of proteins reproducibly, completely, and with high sensitivity. Therefore, MRM contributes to the routine quantitative pipeline for sensitive, reproducible and high-throughput large-scale biomarker discovery, measurement and validation.[16] Although MRM is capable of eliminating interference signals from background and highly abundant contaminating species, one major drawback lies in the low-resolution nature of the Q1 and Q3 mass filters in a QqQ instrument for precursor ion and fragment ion selections, respectively; this results in concomitant elution of interfering ions which cannot be circumvented, especially for complicated biological samples.[17]

With automated or dynamic adjustment of the scheduled retention time (RT) windows, the analytical robustness of MRM can be dramatically enhanced to facilitate large-scale study.[18] Previously prediction of retention times largely relied on intrinsic properties of peptides such as hydrophobicity, which can be calculated from peptide sequences, representative of the sequence-specific retention calculator (SSRCalc) algorithms.[19] In 2012, normalized retention time (iRT), which is an empirically determined RT scale for peptides, was introduced by Escher *et al*.[20] They showed that using empirical data (iRT) instead of a calculated parameter (SSRCalc) provided significantly more reliable results for target peptide RT prediction.[20] The RT scheduling in MRM based on iRT is schematically illustrated in Figure 1.2.

1.2.4 Parallel Reaction Monitoring (PRM)

PRM is a recently developed paradigm for targeted quantitative proteomics typically performed on high-resolution hybrid quadrupole-Orbitrap (Q-OT) or time-of-flight (TOF) instruments.[21, 22] As illustrated in Figure 1.1, a peptide precursor is first isolated by the quadrupole mass filter and then fragmented in the higher-energy collisional dissociation (HCD) cell, and the resulting fragment ions are then analyzed in a high-resolution mass analyzer that permits parallel detection of all target product ions (typically 10–20 ppm).[21] These parallel monitoring features contrast with MRM, allowing all detectable product ions scanned by the Orbitrap mass analyzer in the PRM mode to be used for accurate peptide quantification.[21] As a result, there is no need for *a priori* selection or optimization of target peptide transitions. Furthermore, PRM offers higher specificity or selectivity than MRM on QqQ instruments due to the reduced presence of interfering ions, which are readily discriminated by the high-resolution MS/MS acquisition. In addition to higher specificity, sensitivity is also increased by the trapping capabilities of the C-trap; this is advantageous for detecting low-abundance species in the complex background through the use of longer fill times in the process of fragmentation/accumulation of fragments to increase the signal-to-noise ratio.[23] Several studies showed that PRM and MRM display comparable linearity, dynamic range, precision, and repeatability for protein quantification.[24-26] Some papers reported that PRM exhibited a wider dynamic range than MRM in the presence of complex background matrix owing to higher selectivity.[21]

Despite the immense benefit obtained from high-resolution mass analyzers, PRM analysis performed in an unscheduled mode is only confined to a limited set of analytes

(typically 50 peptides per analysis) when performing measurements in the upper range of acquisition parameter settings (i.e. a maximum fill time exceeding 100 ms at a resolving power of at least 35,000).[23] To overcome such shortcoming, Gallien *et al*. proposed internal standard triggered (IS)-PRM for low-scale PRM experiments.[23, 27] Collectively, PRM represents a promising new addition to the quantitative proteomics toolbox and an attractive alternative to MRM.

1.3 Labeling Strategies in Quantitative Proteomics

1.3.1 Label-free Quantification

Modern proteomics has necessitated accurate, reproducible and rapid quantification of peptides/proteins with elevated sampling throughput and reduced labor. Therefore, label-free quantification (LFQ) emerged as a frequently adopted strategy that offers reasonable sensitivity, high dynamic range and obviates the needs of chemical or metabolic labeling.[28] There are three major LFQ strategies: spectral counting, where relative quantification is achieved based on the number of identified MS/MS spectra for each protein;[29] DIA approaches, which utilize product ion intensity;[30, 31] and MS1 peak intensity, utilizing chromatographic peak area extracted from the selected peptide precursor.[32] In the past decade, LFQ has received broad implementations in a myriad of shotgun or targeted proteomic studies. However, the main disadvantages include reliance on technical reproducibility, since LFQ is more prone to batch-to-batch variation due to differences in sample preparation and instrumentation (Figure 1.3).[33] Moreover, LFQ requires more intense computational processing to normalize such differences. Therefore, tremendous efforts have been made in the field to improve the algorithms of LFQ.[34, 35]

1.3.2 Chemical Labeling

Covalently attaching a small functional group will chemically label peptides, which retain similar chemical and physical properties yet distinguishable m/z values, compared to their natural counterparts; because comparison to the labeled standard could increase precision and accuracy, diverse labeling strategies were incorporated into MS-based proteomic workflows. Among them, isotope-coded affinity tag (ICAT) was developed in 1999 by Gygi *et al*. as the first chemical tagging strategy in quantitative proteomics.[8] Typically, the ICAT reagent consists of three elements: an affinity tag (biotin), a light or heavy isotope-labeled (diethylene glycol) linker containing either hydrogens or deuterons, and a thiol reactive moiety which will specifically derivatize cysteine residues. By targeting different amino acid residues for active-site labeling, ICAT represents a flexible labeling strategy for proteome-wide quantification, in conjunction with subsequent affinity-based enrichment (Figure 1.3). Other similar chemical modification-based approaches were later developed such as stable-isotope dimethyl labeling, which utilizes formaldehyde to globally label the N-terminus and the side chain ε-amino group of lysine residues through reductive amination.[36]

In addition to ICAT-based quantitation, isobaric mass tagging is another popular implementation of chemical derivatization in quantitative proteomics. In 2004, the isobaric tag for relative and absolute quantitation (iTRAQ) strategy was developed by Applied Biosystems (Ross *et al*.), in a 4-plex fashion.[37] The design of the reagent consists of a reporter group (*N*-methylpiperazine), a mass balance group (carbonyl) and an amine-reactive moiety (NHS ester). The amine-reactive NHS ester tag can covalently modify the

N-terminus and the side chain of lysine residue (Figure 1.4). After loss of the balance group during CID fragmentation of the isobaric peptide precursors, the resulting cluster of reporter ions (ranging from *m/z* 114.1 to 117.1 for 4-plex) can be used for relative protein quantification (Figure 1.4). One major advantage for iTRAQ lies in the simultaneous detection of isobaric peptides with identical molecular mass and chromatographic elution behavior during separation, which may reduce background interference. In 2012, Pottiez *et al*. [38] developed an 8-plex version of the iTRAQ reagents (Figure 1.4). Despite structural difference between the two iTRAQ tagging strategies, the 8-plex iTRAQ provides more consistent ratios without compromising protein identification.

Earlier in 2003, Thermo Fisher Scientific also released a product similar to iTRAQ, tandem mass tags (TMT) 6-plex isobaric mass tagging kit, which can provide simultaneous quantification of up to 6 different samples.[39] The structures of TMT reagents are composed of a mass reporter group (1,5-dimethylpiperidine), a mass balance group and an amine-reactive group (NHS-ester) (Figure 1.5). In 2012, McAlister *et al*. expanded the multiplexing capacity of the TMT reagents to at least 10-plex.[40] Recently, the TMT11-131C reagent was developed to further expand the multiplexity to 11-plex.[41]

In addition to the commercialized iTRAQ (SCIEX) and TMT (Thermo Scientific) reagents, individual research groups have made great efforts to design new isobaric tags for chemical labeling of peptides. In 2015, Li *et al*. developed a novel 12-plex DiLeu isobaric tags for quantitative proteomic analysis.[42] In 2018, Virreira Winter *et al*. synthesized and utilized a new version of amine-derivatizing and sulfoxide-containing

isobaric labeling reagents, termed easily abstractable sulfoxide-based isobaric-tag (EASI-tag), for highly accurate and multiplexed quantitative proteomics analysis.[43]

Although the iTRAQ- or TMT-based approaches allow for highly multiplexed and precise quantitative analyses of proteins, quantification of complex proteomic samples could be biased due to the "ratio compression" issue originated from co-fragmentation of peptides.[44] This is usually linked with reporter ion pattern distortion caused by inevitable co-isolated precursor interference during the MS/MS (MS2)-based quantitation.[44] To circumvent this issue, the SPS-MS3 technique, namely isolation waveforms with multiple frequency notches (i.e., synchronous precursor selection, SPS) coupled to MS/MS/MS, can effectively enhance quantitative accuracy and sensitivity.[45, 46] That said, MS2-based TMT outperforms MS3-based TMT since its higher precision and larger identification numbers allow detection of a greater number of significantly altered proteins.[47] Lastly, the introduction of heavy labels at a late stage of sample preparation is more susceptible to experimental errors, compared to metabolic labeling (Figure 1.3).[48]

1.3.3 Metabolic Labeling

In vivo metabolic labeling approaches include stable isotope labeling by amino acids in cell culture (SILAC),[49] stable isotope labeling in mammals (SILAM),[50] and neutron-encoded (NeuCode) SILAC or SILAM.[51] Classical SILAC was first introduced by Ong *et al*. in 2002, which enables comparative analysis of the global proteome from two cell populations fully labeled with "light" or "heavy" isotope-labeled amino acids, respectively.[49] The typical choice of labeled amino acids is lysine and arginine, which cam combine with trypsin digestion to ensure labeling of every peptide except for the C-

terminal peptide of the protein.[52] After equi-mass mixing of the "light" and "heavy" cell lysates followed by tryptic digestion, the resulting peptide mixtures are purified or fractionated prior to LC-MS/MS analysis, and each isotope-labeled peptide usually appears as a "doublet" in the mass spectra with easily distinguishable *m/z* shifts. By comparing the differential signal intensities of the paired isotope-derived peptide peaks, the differences in protein abundance between sample groups can be directly calculated.[49] Naturally, the multi-Da spacing of isotopes in SILAC confines the quantitative capacity to triplex, or triple-SILAC, namely SILAC with three isotope labeling states can also be employed.[53]

Lately, multiple SILAC-derived technical modifications have been developed to enlarge its practicability in the field of quantitative proteomics. The use of SILAC as an internal or "spike-in" standard was proposed by Geiger *et al.*, wherein SILAC is only used to produce heavy labeled reference proteomes and is not incorporated in the early steps of biological experiments.[54] Another similar strategy named "super-SILAC" has been expanded to five SILAC-labeling cell lines which serve as the internal standards for tissue proteome quantification.[55] In another variation of the classical SILAC methodology, the pulse-chase or pulsed SILAC (pSILAC) technique, after the pulse labeling with heavy amino acids in cells, all newly synthesized proteins will incorporate heavy isotopes; this has been successfully applied to study global cellular proteostasis by assessing protein degradation, synthesis/translation and turnover.[56, 57] With this strategy, the transient changes in protein abundance reflective of either protein degradation or synthesis can be investigated by MS analysis. To enhance the multiplexing capabilities of metabolic labeling, NeuCode SILAC was described in 2013 as a novel strategy to improve duplex or

triplex SILAC method by reducing the mass window between isotopologs.[51] NeuCode produces precursor partners spaced by as little as 6 mDa and as much as ~40 mDa apart; this enables amino acid counting and peptide quantification without the need to increase MS1 spectral complexity.

The analytical advantages of SILAC lie in its greater ease of implementation, quantitative accuracy, and reproducibility compared with chemical labeling or LFQ approaches. These advantages arise because SILAC allows earlier introduction of labeling during sample processing and thereby greatly reduces variation and errors (Figure 1.3).[49, 52] That said, traditional SILAC is only applicable to cultured cells or mouse models, whereas clinical samples such as tissue samples and biological fluids are not amenable to metabolic labeling. Furthermore, to achieve similar breadth of quantitative multiplexity enabled by chemical labeling and/or LFQ, traditional duplex-SILAC scheme requires not only time-consuming and labor-intensive metabolic labeling steps but also vastly tedious efforts in sample preparation.

1.3.4 Absolute QUAntitation (AQUA) Using Isotope-labeled Internal Standards

Increased sample multiplexing capacity enabled by chemical labeling strategies has broadened opportunities for conducting large-scale global proteomic studies. These techniques, nonetheless, only provide relative quantification for peptides of interest. In contrast, absolute quantitation (AQUA) is performed by spiking complex samples with stable isotope-labeled (SIL) synthetic peptides that act as internal standards for a specific subset of peptides (Figure 1.3). These peptides are synthesized with incorporated stable isotopes and serve as ideal internal standards by mimicking native peptides formed by

proteolysis (e.g. trypsin digestion). Hence, they are designed to be structurally identical to tryptic peptides generated by sample preparation and can therefore ensure co-elution in LC with target peptide and simultaneous MS analysis.

Once the assay is optimized for a predetermined set of peptides, AQUA used in conjunction with MRM offers the highest level of reproducibility and sensitivity in targeted profiling from multiple samples. This approach has been reported to detect proteins with concentrations less than 50 copies per cell in unfractionated lysates, demonstrating that this quantitative approach largely unaffected by sample complexity. As a gold standard in delivering precise and reliable absolute quantity of target proteins and post-translational modifications (PTMs), AQUA has received wide applications in a plethora of clinical studies, such as quantifying disease biomarkers in patient-derived samples from plasma or serum by coupling the use of MRM with stable isotope dilution.[58-61] Although using SIL peptides for absolute quantification is very precise, availability and costs for such reference peptides limit their applications in assays with a great number of proteins.

1.4 GTP-binding Proteins and Small GTPases

1.4.1 Overview

In cells, guanosine mono-, di-, and triphosphate (GMP, GDP, GTP) constitute fundamental building blocks and messengers for a broad spectrum of cellular processes. Guanine nucleotide binding proteins, or GTP-binding proteins, include septins (SEPT family), tubulins (TUBB family), eukaryotic translation initiation/elongation factors (eIF/eEF family), heterotrimeric G protein alpha subunit (Gα), and so on. By shuffling between the active GTP-bound form and the inactive GDP-bound form these proteins act

as molecular switches, as such they play essential roles in various cellular processes and orchestrate a diversity of signaling networks. For large guanosine triphosphatases (GTPases), the hydrolysis of GTP fuels organelle re-organization, while for small GTPases, the GTP hydrolysis induces protein conformational changes and subsequently the interaction with downstream effectors to transmit extracellular signals. Guanine nucleotide exchange factors (GEFs) and GTPase-activating proteins (GAPs) are two distinct classes of molecular chaperones that mediate the activity of small GTPases (Figure 1.6).[62] In general, GEFs turn on signaling by catalyzing the exchange from GTPase-bound GDP to GTP, whereas GAPs terminate signaling by facilitating GTP hydrolysis.[62, 63] For certain small GTPases that carry C-terminal farnesyl or geranylgeranyl modifications, GDP/GTP switch involves cytosol/membrane alternation and hence is modulated by guanine dissociation inhibitors (GDIs).[63]

1.4.2 Small GTPases

As the largest gene family of monomeric GTP-binding proteins, the Ras superfamily of small GTPases is comprised of over 150 human members (Figure 1.7), with highly evolutionarily conserved orthologs in *Drosophila, C. elegans*, *S. cerevisiae*, *S. pombe*, and plants.[64-66] Based on structural similarities, they can be further classified into five families, Ras, Rho, Rab, Ran, and Arf, as well as the "orphan" or atypical GTPases RhoBTB1/2/2. The Ras family of GTPases responds to extracellular stimuli to regulate cellular gene transcription, proliferation and survival; while the Rho family of GTPases couples the same stimuli to mediate gene expression and cytoskeletal organization.[67-69] The Rho family of GTPases are also known for their role in regulating cell shape and plasticity of cell

migration.[70] The Rab and the Arf families of GTPases control receptor internalization, intracellular vesicular trafficking and actin remodeling.[71, 72] The Ran protein, the single member of the Ran subfamily, is the most abundant small GTPase in cells and is responsible for microtubule stability and nucleocytoplasmic transport.[73]

Growing lines of evidence suggest that small GTPases and their regulators (i.e. GAPs and GEFs) may be potential therapeutic targets for drug discovery in treating a wide variety of diseases, including cancer.[74] Aberrant regulation of small GTPase expression has been reported in various types of cancer including hepatocellular carcinoma (RAB1B, RAB4B, RAB10, RAB22A, and RAB24), non-small lung carcinoma (RAB14, RHOA, RAC1, and CDC42), pancreatic carcinoma (RAB20, RAC1) colorectal cancer (RAB22A, RAC1B) and prostate cancer (RAB3B).[75, 76] Given the important functions of these proteins in signal transduction and trafficking, a better mechanistic understanding of their roles in disease development and progression may provide new insights into strategies for therapeutic intervention.

1.4.3 Heterotrimeric G Proteins

Heterotrimeric G proteins consist of two functional units, an α subunit (Gα) and a tightly associated βγ complex (Gβγ), which play pivotal roles in signal transduction involved with G-protein-coupled receptor (GPCR) activation. The Gα subunit harbors the guanine nucleotide-binding site and is associated with the βγ complex in its GDP-bound inactive state. Agonist-receptor binding triggers GDP/GTP exchange, a conformational change of Gα, subunit dissociation from the βγ complex, and ultimately downstream signaling cascades.[77] It has been estimated that ~700 approved drugs target GPCRs,

suggesting that approximately 35% of approved drugs target GPCRs.[78] GPCRs and GPCR-related proteins, i.e. those proteins upstream or downstream in the GPCR-related pathways, represent ~17% of all protein targets for approved drugs, with GPCRs themselves accounting for ~12%. As such, GPCRs constitute the largest family of proteins targeted by approved drugs.[78]

1.4.4 Proteomic Profiling of GTP-binding Proteins

1.4.4.1 Acyl-phosphate GTP Affinity Probes

Several structurally distinct chemical probes have been developed for targeting GTP-binding proteins. One example is the commercialized lysine-reactive desthiobiotin-GTP probes that contains an acyl phosphate anhydride (ActivX Biosciences)[79]; these probes target the conserved lysine residue in the GTP-binding GXXXXGK motif (P-loop)[80] and have been employed by several studies examining the GTPase proteome. Patricelli *et al*.[81] and Qiu *et al*.[82] reported a lysine-reactive nucleotide acyl phosphate probe for the proteome-wide profiling of nucleotide-binding proteins, including kinases and GTPases. By exploiting a similar strategy, Xiao *et al.* identified 66 GTP-binding proteins in HL-60 cell lysates[83], and later a similar approach utilizing 6-thioguanosine triphosphate (^{S}GTP) acyl-phosphate probe combined with SILAC was devised for proteome-wide quantification of ^{S}GTP-binding proteins in Jurkat-T cell lysates[84], in which 91 GTP-binding proteins were quantified. Hunter *et al.* used the desthiobiotin-GTP probe for characterizing an active site inhibitor of oncogenic KRAS-G12C in MIA PaCa cell lysates[85], which led to the detection of over 100 GTP-binding proteins. Recently, Cai *et. al* further extended this approach by combining the use of isotope-coded desthiobiotin-GTP

acyl-phosphate probes with MRM and quantified 91 GTP-binding proteins in SW480/SW620 cell lysates.[86] However, these chemical proteomic probes which incorporate adenine/guanine cofactor-based chemical scaffolds may suffer from off-target engagement of adenosine-binding proteins due to their widespread reactivity.[87]

1.4.4.2 Photoreactive GTP Affinity Probes

Another type of GTP affinity probe relies on the photoactive covalent modifications that occur at the GTP-binding sites. Kaneda *et al*.[88] first synthesized a GTP probe that allows simple and efficient photoaffinity-based proteomic profiling of GTP-binding proteins. Another example is the GTP-BP-yne probe reported by Cisar *et al.*,[89] which facilitated active-site labeling of more than 30 annotated GTP-binding proteins subsequently identified by MudPIT.

1.4.4.3 Gel Electrophoresis (Ge) Coupled to LC-MRM (GeLC-MRM)

In addition to small-molecule probes, other profiling techniques take advantage of the GTPase-binding domain of individual downstream effectors and their distinct molecular weights (15–37 kDa) to enrich specific active small GTPases. Zhang *et al.*[90] first combined such pull-down strategies with gel electrophoresis (Ge) coupled to LC-MRM (GeLC-MRM) for the development of quantitative multiplexed small GTPase activity assay, and they were able to detect 12 active isoforms of small GTPases simultaneously. Halvey *et al*.[91] also demonstrated that GeLC-MRM is a robust and sensitive approach to quantify KRAS mutant variants in complex biological samples.

1.5 Scope of the Dissertation

Compared to conventional discovery-based proteomic approaches, i.e. shotgun proteomics, targeted proteomic techniques show clear advantages in achieving better reproducibility, higher sensitivity and superior quantitative accuracy. In this dissertation, we reported the development and application of novel targeted quantitative proteomic methods for high-throughput and reproducible profiling of small GTPases in cultured cells and patient-derived brain tissues that carry disease-related changes. We believe this is an important area of study because small GTPases of the Ras superfamily represent a class of crucial signaling molecules in cells, and the aberrant regulation of their expressions is implicated with various types of human diseases.

In Chapter two, we describe the development of a novel targeted quantitative proteomic method that involves metabolic labeling by SILAC, total protein quantitation, prior enrichment of small GTPases by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) in the low molecular weight (15–37 kDa) protein fraction and the subsequent scheduled LC-MS/MS analysis in the multiple-reaction monitoring (MRM) mode. Owing to the largely reduced complexity of the proteome analyzed, this method allows reliable high-throughput quantification of small GTPases of the Ras superfamily with relatively low protein inputs (5–100 μg) and without the need for chemoaffinity or immunoaffinity enrichment. Taking advantage of both the effective enrichment of small GTPases by gel-based fractionation and the analytical robustness of the scheduled MRM analysis, over 90 small GTPases were robustly quantified in two scheduled LC-MRM runs. We also demonstrated that this MRM-based assay displayed

much better sensitivity, reproducibility and accuracy than the discovery-based shotgun proteomic method. We applied this method to probing the differential protein abundance of small GTPases in three pairs of primary/metastatic human melanoma cell lines: WM-115/WM-266-4, IGR39/IGR37 and WM793/1205Lu. Bioinformatic analyses of The Cancer Genome Atlas (TCGA) data and other publicly available data as well as cell-based assays revealed previously unrecognized roles of RAB38 in promoting melanoma metastasis. Diminished promoter methylation and the subsequent augmented binding of transcription factor MITF contributed to elevated expression of the *RAB38* gene in metastatic versus primary melanoma cells. Moreover, RAB38 promoted invasion of cultured melanoma cells by modulating the expression and activities of matrix metalloproteinases-2 and -9. Together, these data establish a novel targeted proteomic method for interrogating the small GTPase proteome in human cells and identify epigenetic reactivation of RAB38 as a contributing factor to metastatic transformation in melanoma.

In Chapter three, we applied the established quantitative proteomic method in probing the differential expression of small GTPases associated with acquired tamoxifen resistance in the estrogen-receptor (ER)-positive MCF-7 breast cancer cells. Briefly, we employed the established quantitative proteomic method to assess the differential expression of small GTPases in wild-type MCF-7 and the paired tamoxifen-resistant breast cancer cells. The method displayed superior sensitivity and reproducibility over the shotgun-proteomic approach, and it facilitated the quantification of 96 small GTPases. Among them, 13 and 10 proteins were significantly down- and up-regulated (with >1.5-fold change) in the tamoxifen-resistant line relative to the parental line, respectively. Notably, we observed a

significant down-regulation of RAB31 in tamoxifen-resistant cells, which, in combination with bioinformatic analysis and downstream validation experiments, supported a role for RAB31 in tamoxifen resistance in ER-positive breast-cancer cells. Together, our results demonstrate that the targeted proteomic method constituted a powerful approach for revealing the role of small GTPases in therapeutic resistance.

In Chapter four, we sought to conduct a systematic study of the implications of small GTPases in the metastatic transformation of colorectal cancer (CRC). By utilizing the established MRM method, we investigated comprehensively the relative expression of the small GTPase proteome in a pair of matched primary/metastatic CRC cell lines (SW480/SW620). Among the 83 quantified small GTPases, 25 exhibited at least a 1.5-fold difference in protein expression in metastatic SW620 relative to primary SW480 cells. Bioinformatic analyses revealed that diminished SAR1B mRNA expression is significantly associated with higher CRC stages and unfavorable patient prognosis, which supports a potential role of SAR1B in suppressing CRC metastasis. In addition, diminished SAR1B expression could stimulate epithelial–mesenchymal transition (EMT), thereby promoting motility and *in vitro* metastasis of SW480 cells. In summary, we profiled systematically, by employing an MRM-based targeted proteomic method, the differentially expressed small GTPases in a matched pair of primary/metastatic CRC cell lines. Our results revealed the potential roles for SAR1B in suppressing CRC metastasis and in the prognosis of CRC patients.

In Chapter five, we present the development of a novel targeted quantitative proteomic assay aimed at switching from metabolic labeling to the use of stable isotope-labeled (SIL)

peptides. We applied this MRM method, which was built upon in-house shotgun proteomics data and online MRM assay design tool, for simultaneous assessment of the relative expression of small GTPases from Alzheimer's disease (AD) patient derived brain tissues. The scheduled GeLC-MRM analyses provided robust quantification of more than 80 small GTPases in tissue samples by using a 90-min LC-MRM run, with excellent throughput and reproducibility. Interestingly, levels of RAB27B, RAB3A and RAB3D proteins, which were previously shown to be involved with synaptic and secretory vesicles, were found to increase in brain tissue samples with higher degree of disease severity. This facile and accurate assay provides invaluable knowledge to probe the potential roles for small GTPases in the development and progression of AD.

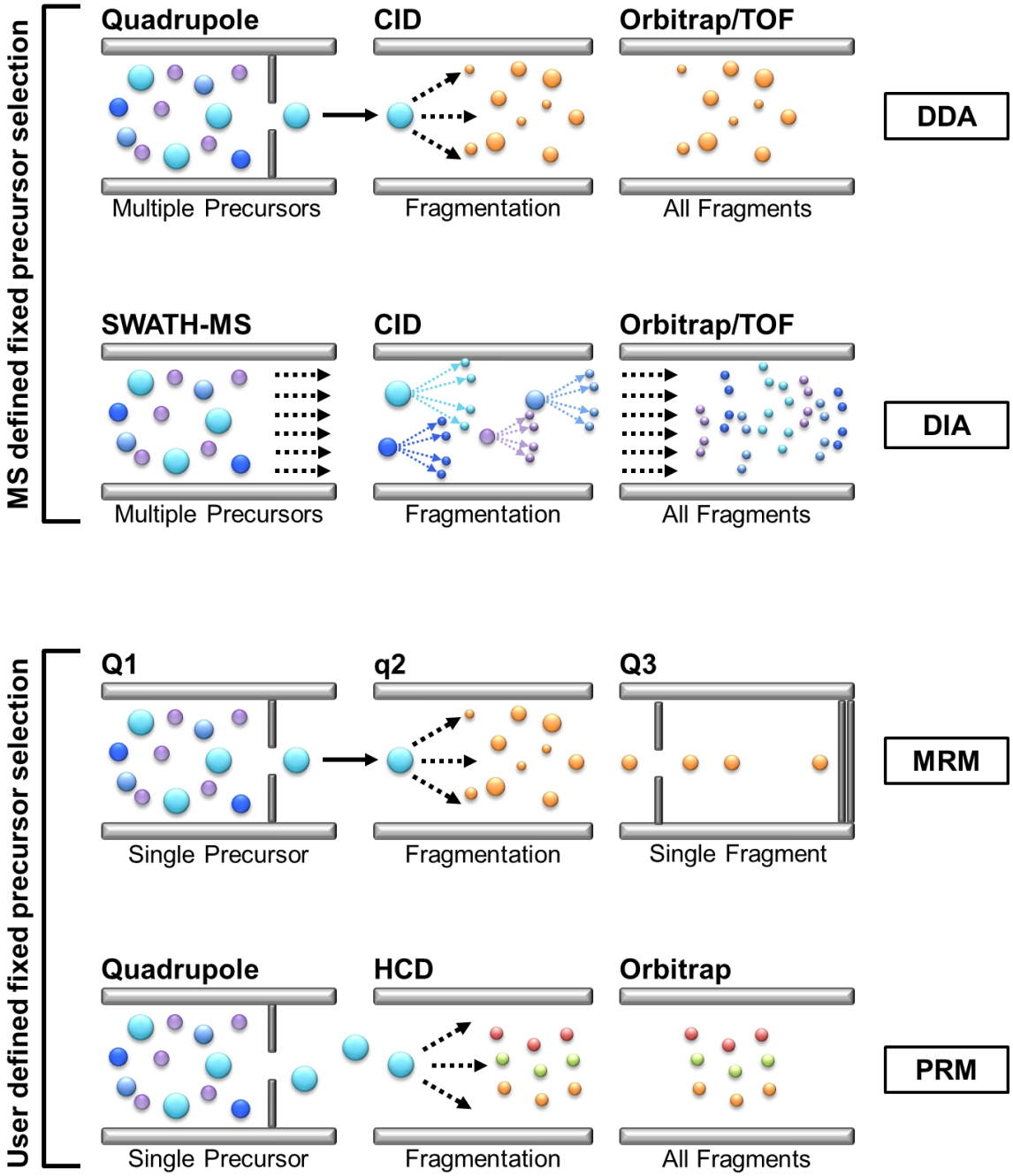


Figure 1.1 Representative scanning modes used in MS-based bottom-up proteomics. Shown is a schematic diagram representing common ion detection methods in MS.

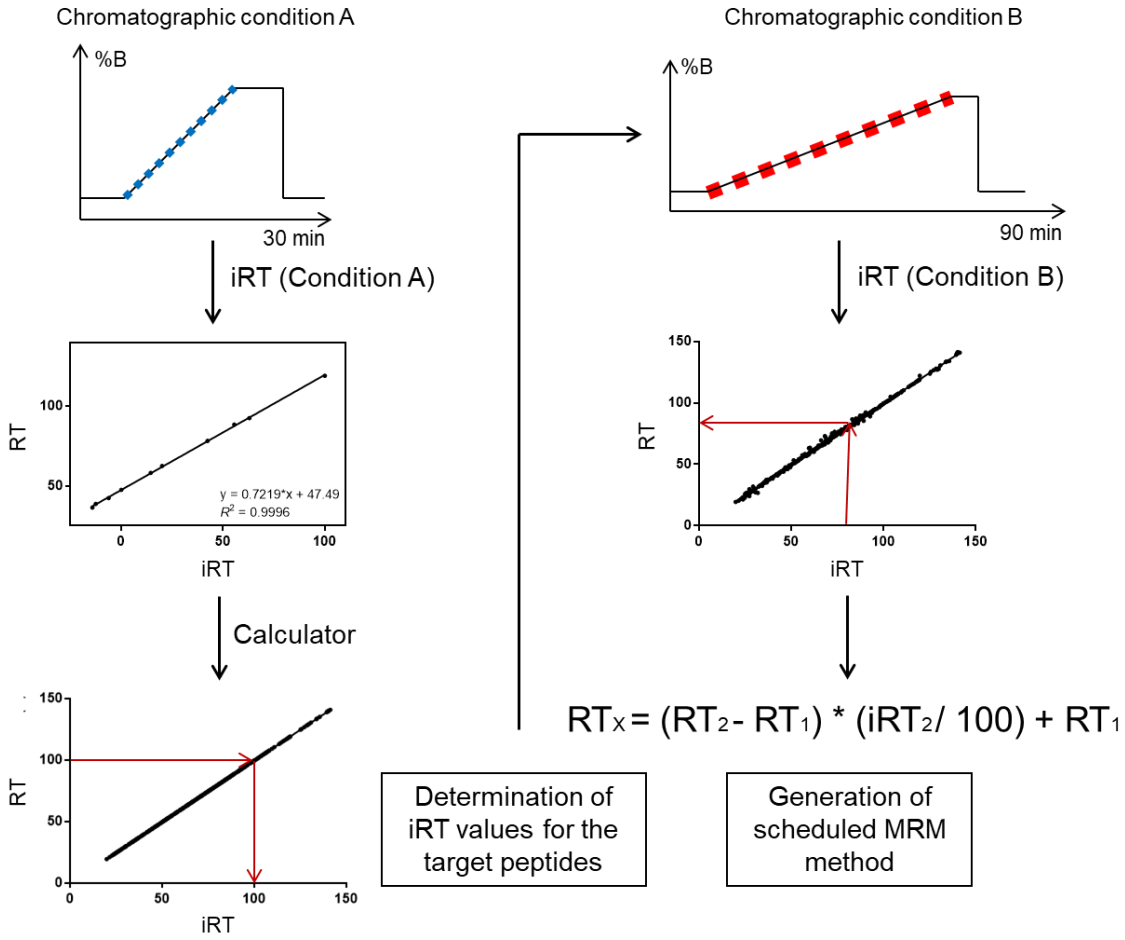


Figure 1.2 A schematic workflow that outlines the empirical determination of iRT scale, conversion of RTs for targeted peptides into iRTs, and RT scheduling for MRM.

To simplify, RT_1 and RT_2 represent 2 of the 10 reference peptides (iRT-peptides) used for iRT transformation based upon linear regression. The retention time of the target peptide (RT_x) is can be empirically transformed into iRT (iRT_x) using the established linear regression of the iRT-peptides. iRT_x values can then be transferred to a different chromatographic setup by using an RT calibration of the iRT-peptides.

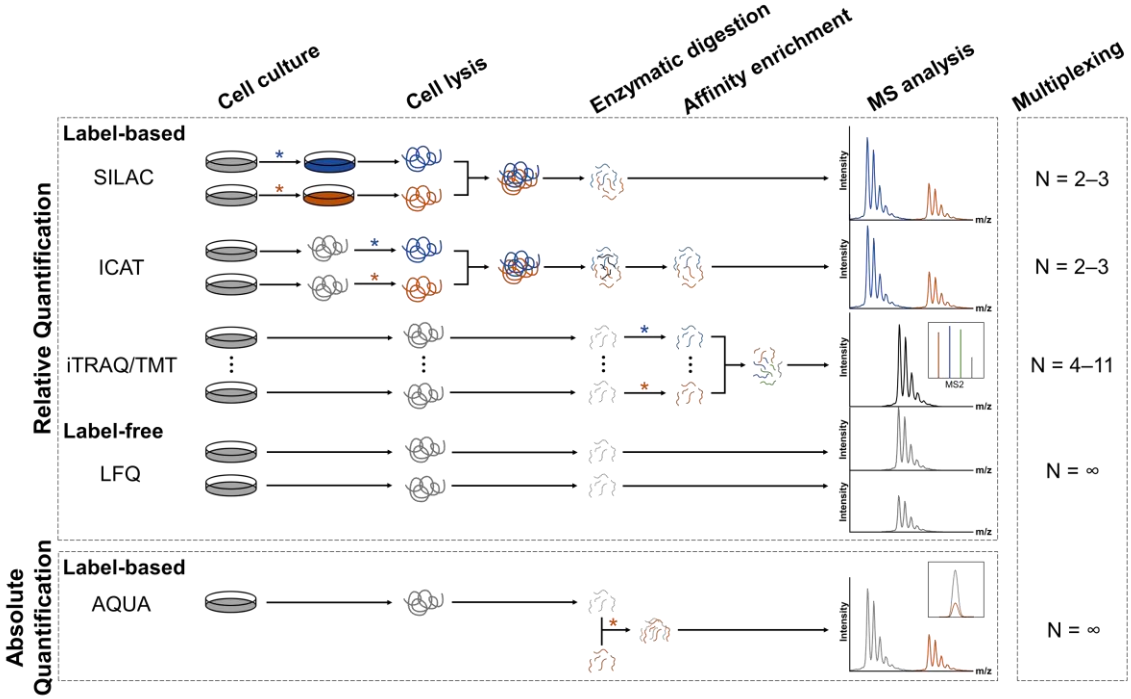


Figure 1.3 Comparison of commonly used labeling strategies in quantitative proteomics.

The blue and orange asterisk indicates the introduction of light and heavy stable isotope labeling, respectively.

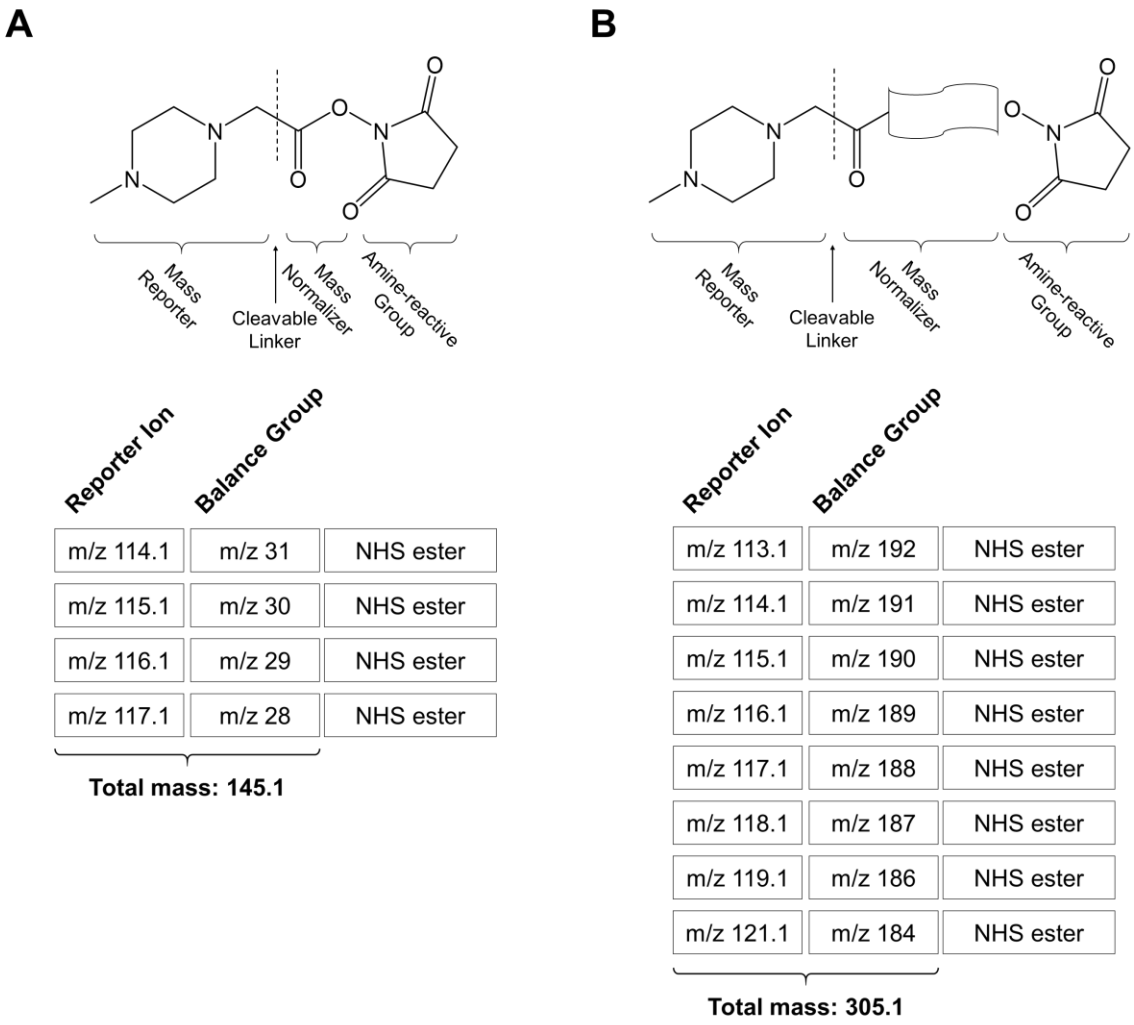


Figure 1.4 Design and concept of iTRAQ.

(A) The backbone structure and the detailed description of the 4-plex iTRAQ reagents. (B) The backbone structure and the detailed description of the 8-plex iTRAQ reagents. The dashed line indicates the fragmentation site of CID.

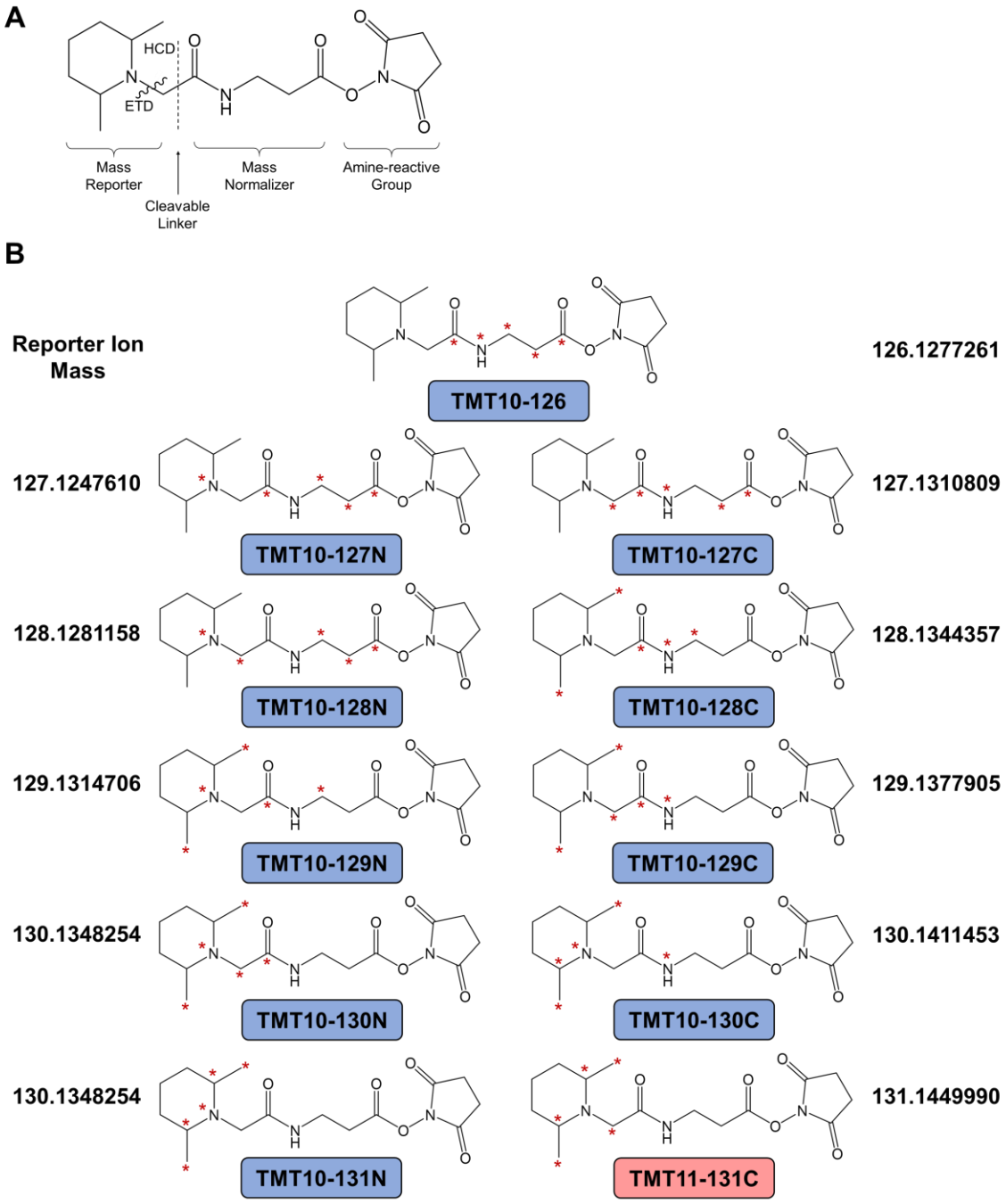


Figure 1.5 Design and concept of TMT.

(A) The backbone structure of TMT reagent. (B) The detailed molecular structures of the 11-plex version of the TMT reagent, where asterisk denotes the position of isotopic labelled ^{13}C or ^{14}N, the zig-zag line indicates the fragmentation site of ETD, and the dashed line indicates the fragmentation site of HCD.

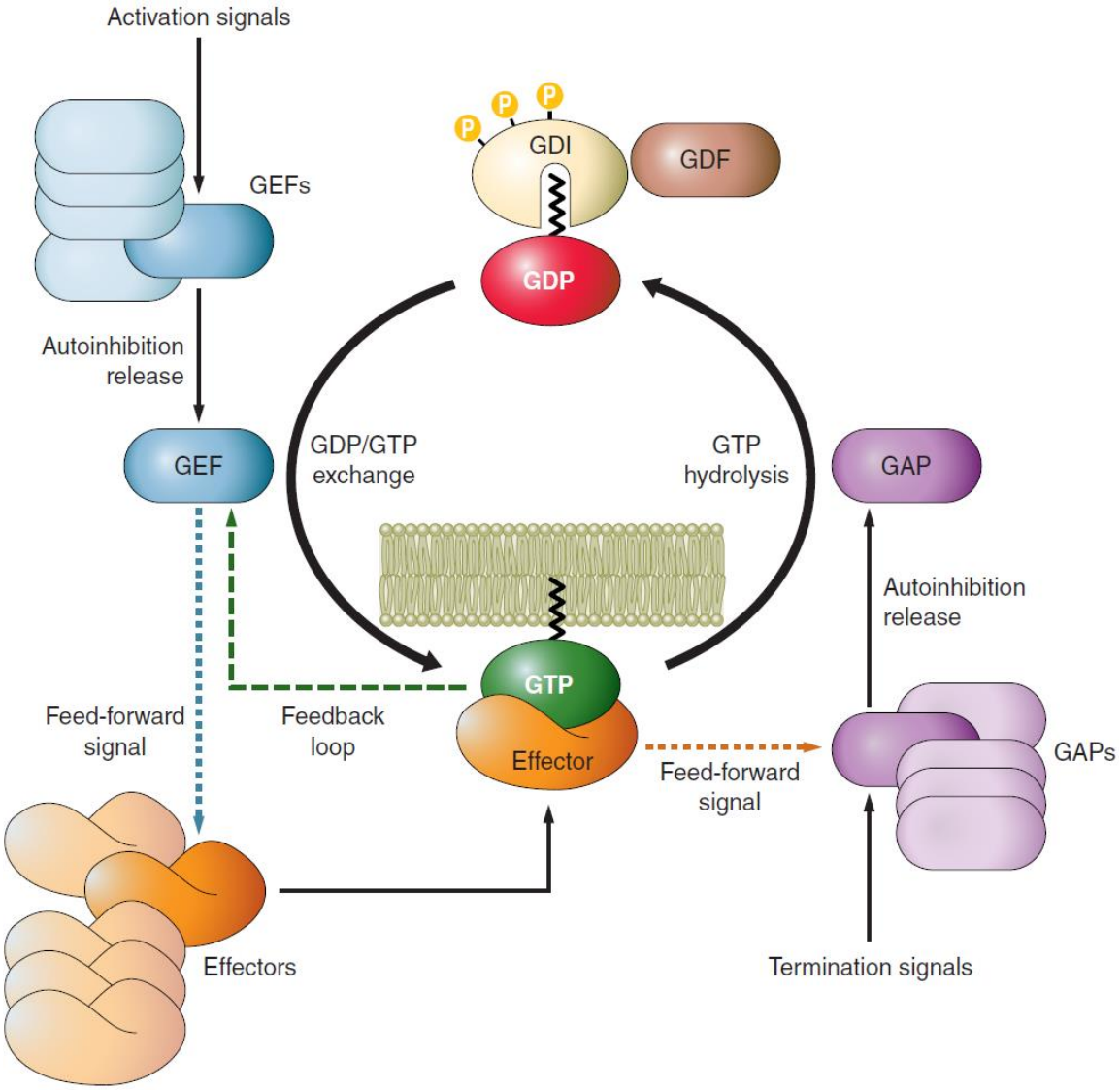


Figure 1.6 Regulation of small GTPases by GEFs, GAPs, and GDIs.

A diagram depicting the regulation of the GDP/GTP switch of small GTPases modulated by GEFs, GAPs, and GDIs (adopted from Ref[63]).

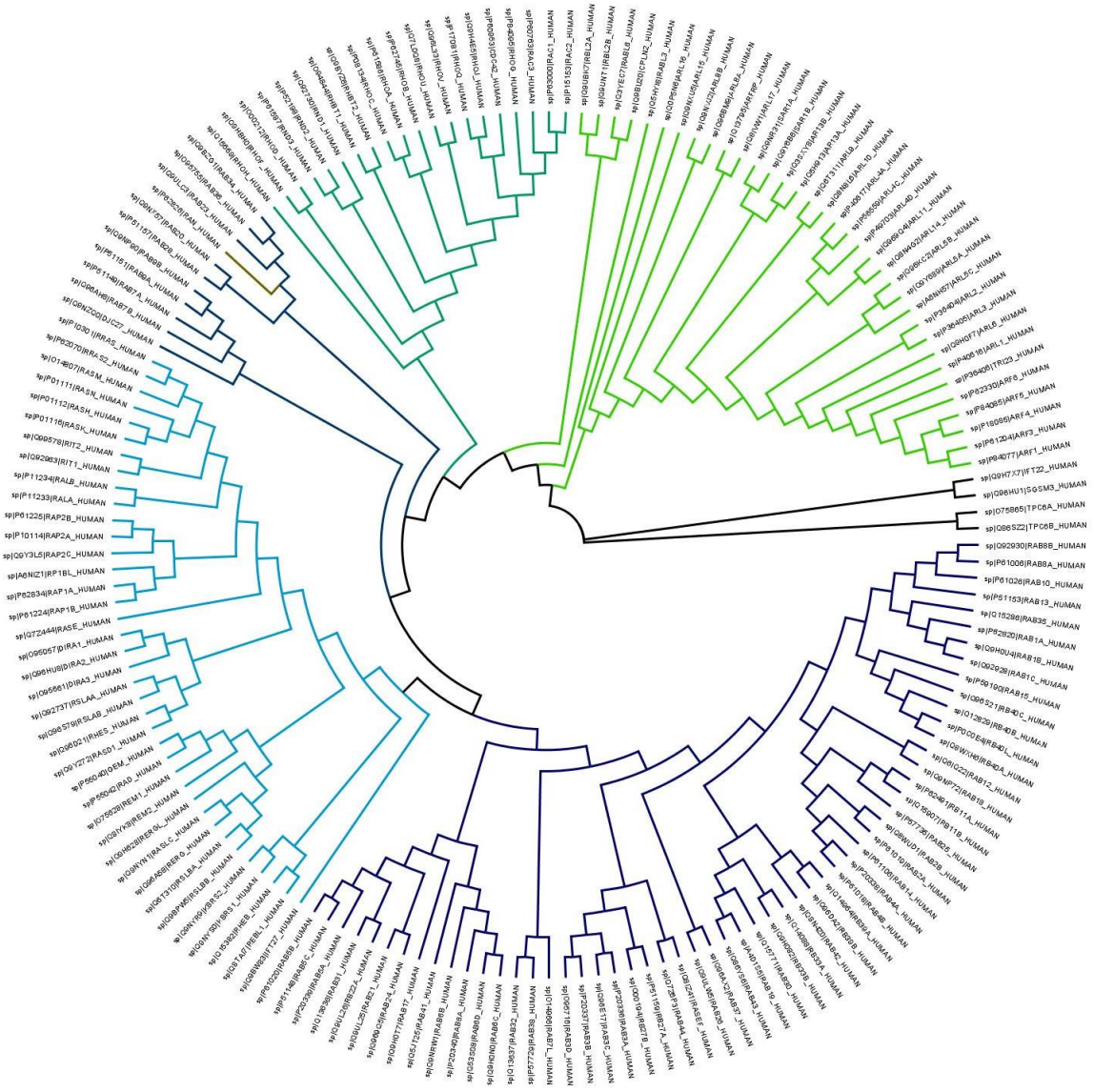


Figure 1.7 Phylogenetic relationships and gene structure of the Homo Sapiens small GTPase genes.

The unrooted tree was generated using the MEGA v7.0 software with the full-length amino acid sequences of the *Homo Sapiens* small GTPase proteins using a Neighbor-Joining (NJ) method, including 1,000 boot-strap replications. All the protein sequences were aligned using ClustalW. The phylogenetic tree was visualized using the FigTree v1.4.4 software. The five sub-families of small GTPase genes are highlighted with different colored tree branches.

1. Aebersold, R.; Mann, M., Mass spectrometry-based proteomics. *Nature* **2003,** *422* (6928), 198-207.
2. Nilsson, T.; Mann, M.; Aebersold, R.; Yates, J. R., 3rd; Bairoch, A.; Bergeron, J. J., Mass spectrometry in high-throughput proteomics: ready for the big time. *Nat Methods* **2010,** *7* (9), 681-5.
3. Zhang, Y.; Fonslow, B. R.; Shan, B.; Baek, M. C.; Yates, J. R., 3rd, Protein analysis by shotgun/bottom-up proteomics. *Chem Rev* **2013,** *113* (4), 2343-94.
4. Wu, C. C.; MacCoss, M. J., Shotgun proteomics: tools for the analysis of complex biological systems. *Curr Opin Mol Ther* **2002,** *4* (3), 242-50.
5. Wolters, D. A.; Washburn, M. P.; Yates, J. R., 3rd, An automated multidimensional protein identification technology for shotgun proteomics. *Anal Chem* **2001,** *73* (23), 5683-90.
6. Washburn, M. P.; Wolters, D.; Yates, J. R., 3rd, Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol* **2001,** *19* (3), 242-7.
7. McDonald, W. H.; Yates, J. R., 3rd, Shotgun proteomics and biomarker discovery. *Dis Markers* **2002,** *18* (2), 99-105.
8. Gygi, S. P.; Rist, B.; Gerber, S. A.; Turecek, F.; Gelb, M. H.; Aebersold, R., Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat Biotechnol* **1999,** *17* (10), 994-9.
9. Tabb, D. L.; Vega-Montoto, L.; Rudnick, P. A.; Variyath, A. M.; Ham, A. J.; Bunk, D. M.; Kilpatrick, L. E.; Billheimer, D. D.; Blackman, R. K.; Cardasis, H. L.; Carr, S. A.; Clauser, K. R.; Jaffe, J. D.; Kowalski, K. A.; Neubert, T. A.; Regnier, F. E.; Schilling, B.; Tegeler, T. J.; Wang, M.; Wang, P.; Whiteaker, J. R.; Zimmerman, L. J.; Fisher, S. J.; Gibson, B. W.; Kinsinger, C. R.; Mesri, M.; Rodriguez, H.; Stein, S. E.; Tempst, P.; Paulovich, A. G.; Liebler, D. C.; Spiegelman, C., Repeatability and reproducibility in proteomic identifications by liquid chromatography-tandem mass spectrometry. *J Proteome Res* **2010,** *9* (2), 761-76.
10. Gillet, L. C.; Navarro, P.; Tate, S.; Rost, H.; Selevsek, N.; Reiter, L.; Bonner, R.; Aebersold, R., Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol Cell Proteomics* **2012,** *11* (6), O111 016717.
11. Ludwig, C.; Gillet, L.; Rosenberger, G.; Amon, S.; Collins, B. C.; Aebersold, R., Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial. *Mol Syst Biol* **2018,** *14* (8), e8126.
12. Bruderer, R.; Bernhardt, O. M.; Gandhi, T.; Miladinovic, S. M.; Cheng, L. Y.; Messner, S.; Ehrenberger, T.; Zanotelli, V.; Butscheid, Y.; Escher, C.; Vitek, O.; Rinner, O.; Reiter, L., Extending the limits of quantitative proteome profiling with data-independent acquisition and application to acetaminophen-treated three-dimensional liver microtissues. *Mol Cell Proteomics* **2015,** *14* (5), 1400-10.
13. Rost, H. L.; Rosenberger, G.; Navarro, P.; Gillet, L.; Miladinovic, S. M.; Schubert, O. T.; Wolski, W.; Collins, B. C.; Malmstrom, J.; Malmstrom, L.; Aebersold, R., OpenSWATH enables automated, targeted analysis of data-independent acquisition MS data. *Nat Biotechnol* **2014,** *32* (3), 219-23.

14. Rosenberger, G.; Koh, C. C.; Guo, T.; Rost, H. L.; Kouvonen, P.; Collins, B. C.; Heusel, M.; Liu, Y.; Caron, E.; Vichalkovski, A.; Faini, M.; Schubert, O. T.; Faridi, P.; Ebhardt, H. A.; Matondo, M.; Lam, H.; Bader, S. L.; Campbell, D. S.; Deutsch, E. W.; Moritz, R. L.; Tate, S.; Aebersold, R., A repository of assays to quantify 10,000 human proteins by SWATH-MS. *Sci Data* **2014,** *1*, 140031.
15. Lange, V.; Picotti, P.; Domon, B.; Aebersold, R., Selected reaction monitoring for quantitative proteomics: a tutorial. *Mol Syst Biol* **2008,** *4*, 222.
16. Whiteaker, J. R.; Lin, C.; Kennedy, J.; Hou, L.; Trute, M.; Sokal, I.; Yan, P.; Schoenherr, R. M.; Zhao, L.; Voytovich, U. J.; Kelly-Spratt, K. S.; Krasnoselsky, A.; Gafken, P. R.; Hogan, J. M.; Jones, L. A.; Wang, P.; Amon, L.; Chodosh, L. A.; Nelson, P. S.; McIntosh, M. W.; Kemp, C. J.; Paulovich, A. G., A targeted proteomics-based pipeline for verification of biomarkers in plasma. *Nat Biotechnol* **2011,** *29* (7), 625-34.
17. Gallien, S.; Duriez, E.; Demeure, K.; Domon, B., Selectivity of LC-MS/MS analysis: implication for proteomics experiments. *J Proteomics* **2013,** *81*, 148-58.
18. Gallien, S.; Peterman, S.; Kiyonami, R.; Souady, J.; Duriez, E.; Schoen, A.; Domon, B., Highly multiplexed targeted proteomics using precise control of peptide retention time. *Proteomics* **2012,** *12* (8), 1122-1133.
19. Krokhin, O. V., Sequence-specific retention calculator. Algorithm for peptide retention prediction in ion-pair RP-HPLC: application to 300- and 100-A pore size C18 sorbents. *Anal Chem* **2006,** *78* (22), 7785-95.
20. Escher, C.; Reiter, L.; MacLean, B.; Ossola, R.; Herzog, F.; Chilton, J.; MacCoss, M. J.; Rinner, O., Using iRT, a normalized retention time for more targeted measurement of peptides. *Proteomics* **2012,** *12* (8), 1111-21.
21. Peterson, A. C.; Russell, J. D.; Bailey, D. J.; Westphall, M. S.; Coon, J. J., Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics. *Mol Cell Proteomics* **2012,** *11* (11), 1475-88.
22. Rauniyar, N., Parallel Reaction Monitoring: A Targeted Experiment Performed Using High Resolution and High Mass Accuracy Mass Spectrometry. *Int J Mol Sci* **2015,** *16* (12), 28566-81.
23. Gallien, S.; Kim, S. Y.; Domon, B., Large-Scale Targeted Proteomics Using Internal Standard Triggered-Parallel Reaction Monitoring (IS-PRM). *Molecular & Cellular Proteomics* **2015,** *14* (6), 1630-1644.
24. Ronsein, G. E.; Pamir, N.; von Haller, P. D.; Kim, D. S.; Oda, M. N.; Jarvik, G. P.; Vaisar, T.; Heinecke, J. W., Parallel reaction monitoring (PRM) and selected reaction monitoring (SRM) exhibit comparable linearity, dynamic range and precision for targeted quantitative HDL proteomics. *J Proteomics* **2015,** *113*, 388-99.
25. Nakamura, K.; Hirayama-Kurogi, M.; Ito, S.; Kuno, T.; Yoneyama, T.; Obuchi, W.; Terasaki, T.; Ohtsuki, S., Large-scale multiplex absolute protein quantification of drug-metabolizing enzymes and transporters in human intestine, liver, and kidney microsomes by SWATH-MS: Comparison with MRM/SRM and HR-MRM/PRM. *Proteomics* **2016,** *16* (15-16), 2106-17.
26. Hoffman, M. A.; Fang, B.; Haura, E. B.; Rix, U.; Koomen, J. M., Comparison of Quantitative Mass Spectrometry Platforms for Monitoring Kinase ATP Probe Uptake in Lung Cancer. *J Proteome Res* **2018,** *17* (1), 63-75.

27. Gallien, S.; Bourmaud, A.; Kim, S. Y.; Domon, B., Technical considerations for large-scale parallel reaction monitoring analysis. *J Proteomics* **2014,** *100*, 147-59.
28. Zhu, W.; Smith, J. W.; Huang, C. M., Mass spectrometry-based label-free quantitative proteomics. *J Biomed Biotechnol* **2010,** *2010*, 840518.
29. 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* **2004,** *76* (14), 4193-201.
30. Collins, B. C.; Gillet, L. C.; Rosenberger, G.; Rost, H. L.; Vichalkovski, A.; Gstaiger, M.; Aebersold, R., Quantifying protein interaction dynamics by SWATH mass spectrometry: application to the 14-3-3 system. *Nat Methods* **2013,** *10* (12), 1246-53.
31. Egertson, J. D.; Kuehn, A.; Merrihew, G. E.; Bateman, N. W.; MacLean, B. X.; Ting, Y. S.; Canterbury, J. D.; Marsh, D. M.; Kellmann, M.; Zabrouskov, V.; Wu, C. C.; MacCoss, M. J., Multiplexed MS/MS for improved data-independent acquisition. *Nat Methods* **2013,** *10* (8), 744-6.
32. Levin, Y.; Schwarz, E.; Wang, L.; Leweke, F. M.; Bahn, S., Label-free LC-MS/MS quantitative proteomics for large-scale biomarker discovery in complex samples. *J Sep Sci* **2007,** *30* (14), 2198-203.
33. Shalit, T.; Elinger, D.; Savidor, A.; Gabashvili, A.; Levin, Y., MS1-based label-free proteomics using a quadrupole orbitrap mass spectrometer. *J Proteome Res* **2015,** *14* (4), 1979-86.
34. Cox, J.; Hein, M. Y.; Luber, C. A.; Paron, I.; Nagaraj, N.; Mann, M., Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics* **2014,** *13* (9), 2513-26.
35. Neilson, K. A.; Ali, N. A.; Muralidharan, S.; Mirzaei, M.; Mariani, M.; Assadourian, G.; Lee, A.; van Sluyter, S. C.; Haynes, P. A., Less label, more free: approaches in label-free quantitative mass spectrometry. *Proteomics* **2011,** *11* (4), 535-53.
36. Hsu, J. L.; Huang, S. Y.; Chow, N. H.; Chen, S. H., Stable-isotope dimethyl labeling for quantitative proteomics. *Anal Chem* **2003,** *75* (24), 6843-52.
37. Ross, P. L.; Huang, Y. N.; Marchese, J. N.; Williamson, B.; Parker, K.; Hattan, S.; Khainovski, N.; Pillai, S.; Dey, S.; Daniels, S.; Purkayastha, S.; Juhasz, P.; Martin, S.; Bartlet-Jones, M.; He, F.; Jacobson, A.; Pappin, D. J., Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics* **2004,** *3* (12), 1154-69.
38. Pottiez, G.; Wiederin, J.; Fox, H. S.; Ciborowski, P., Comparison of 4-plex to 8-plex iTRAQ quantitative measurements of proteins in human plasma samples. *J Proteome Res* **2012,** *11* (7), 3774-81.
39. Thompson, A.; Schafer, J.; Kuhn, K.; Kienle, S.; Schwarz, J.; Schmidt, G.; Neumann, T.; Johnstone, R.; Mohammed, A. K.; Hamon, C., Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal Chem* **2003,** *75* (8), 1895-904.
40. McAlister, G. C.; Huttlin, E. L.; Haas, W.; Ting, L.; Jedrychowski, M. P.; Rogers, J. C.; Kuhn, K.; Pike, I.; Grothe, R. A.; Blethrow, J. D.; Gygi, S. P., Increasing the multiplexing capacity of TMTs using reporter ion isotopologues with isobaric masses. *Anal Chem* **2012,** *84* (17), 7469-78.

41. Stepanova, E.; Gygi, S. P.; Paulo, J. A., Filter-Based Protein Digestion (FPD): A Detergent-Free and Scaffold-Based Strategy for TMT Workflows. *J Proteome Res* **2018,** *17* (3), 1227-1234.
42. Frost, D. C.; Greer, T.; Li, L., High-resolution enabled 12-plex DiLeu isobaric tags for quantitative proteomics. *Anal Chem* **2015,** *87* (3), 1646-54.
43. Virreira Winter, S.; Meier, F.; Wichmann, C.; Cox, J.; Mann, M.; Meissner, F., EASI-tag enables accurate multiplexed and interference-free MS2-based proteome quantification. *Nat Methods* **2018,** *15* (7), 527-530.
44. Savitski, M. M.; Mathieson, T.; Zinn, N.; Sweetman, G.; Doce, C.; Becher, I.; Pachl, F.; Kuster, B.; Bantscheff, M., Measuring and managing ratio compression for accurate iTRAQ/TMT quantification. *J Proteome Res* **2013,** *12* (8), 3586-98.
45. Ting, L.; Rad, R.; Gygi, S. P.; Haas, W., MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics. *Nat Methods* **2011,** *8* (11), 937-40.
46. McAlister, G. C.; Nusinow, D. P.; Jedrychowski, M. P.; Wuhr, M.; Huttlin, E. L.; Erickson, B. K.; Rad, R.; Haas, W.; Gygi, S. P., MultiNotch MS3 enables accurate, sensitive, and multiplexed detection of differential expression across cancer cell line proteomes. *Anal Chem* **2014,** *86* (14), 7150-8.
47. Hogrebe, A.; von Stechow, L.; Bekker-Jensen, D. B.; Weinert, B. T.; Kelstrup, C. D.; Olsen, J. V., Benchmarking common quantification strategies for large-scale phosphoproteomics. *Nat Commun* **2018,** *9* (1), 1045.
48. Bantscheff, M.; Lemeer, S.; Savitski, M. M.; Kuster, B., Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present. *Anal Bioanal Chem* **2012,** *404* (4), 939-65.
49. Ong, S. E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D. B.; Steen, H.; Pandey, A.; Mann, M., Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* **2002,** *1* (5), 376-86.
50. 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* **2004,** *76* (17), 4951-9.
51. Hebert, A. S.; Merrill, A. E.; Bailey, D. J.; Still, A. J.; Westphall, M. S.; Strieter, E. R.; Pagliarini, D. J.; Coon, J. J., Neutron-encoded mass signatures for multiplexed proteome quantification. *Nat Methods* **2013,** *10* (4), 332-4.
52. Chen, X. L.; Wei, S. S.; Ji, Y. L.; Guo, X. J.; Yang, F. Q., Quantitative proteomics using SILAC: Principles, applications, and developments. *Proteomics* **2015,** *15* (18), 3175-3192.
53. Hilger, M.; Mann, M., Triple SILAC to determine stimulus specific interactions in the Wnt pathway. *J Proteome Res* **2012,** *11* (2), 982-94.
54. Geiger, T.; Wisniewski, J. R.; Cox, J.; Zanivan, S.; Kruger, M.; Ishihama, Y.; Mann, M., Use of stable isotope labeling by amino acids in cell culture as a spike-in standard in quantitative proteomics. *Nat Protoc* **2011,** *6* (2), 147-157.
55. Geiger, T.; Cox, J.; Ostasiewicz, P.; Wisniewski, J. R.; Mann, M., Super-SILAC mix for quantitative proteomics of human tumor tissue. *Nature Methods* **2010,** *7* (5), 383-U64.
56. Schwanhausser, B.; Gossen, M.; Dittmar, G.; Selbach, M., Global analysis of cellular protein translation by pulsed SILAC. *Proteomics* **2009,** *9* (1), 205-9.

57. Fierro-Monti, I.; Racle, J.; Hernandez, C.; Waridel, P.; Hatzimanikatis, V.; Quadroni, M., A novel pulse-chase SILAC strategy measures changes in protein decay and synthesis rates induced by perturbation of proteostasis with an Hsp90 inhibitor. *PLoS One* **2013,** *8* (11), e80423.
58. Stahl-Zeng, J.; Lange, V.; Ossola, R.; Eckhardt, K.; Krek, W.; Aebersold, R.; Domon, B., High sensitivity detection of plasma proteins by multiple reaction monitoring of N-glycosites. *Mol Cell Proteomics* **2007,** *6* (10), 1809-17.
59. Keshishian, H.; Addona, T.; Burgess, M.; Mani, D. R.; Shi, X.; Kuhn, E.; Sabatine, M. S.; Gerszten, R. E.; Carr, S. A., Quantification of cardiovascular biomarkers in patient plasma by targeted mass spectrometry and stable isotope dilution. *Mol Cell Proteomics* **2009,** *8* (10), 2339-49.
60. Carvajal-Hausdorf, D. E.; Schalper, K. A.; Neumeister, V. M.; Rimm, D. L., Quantitative measurement of cancer tissue biomarkers in the lab and in the clinic. *Lab Invest* **2015,** *95* (4), 385-96.
61. You, J.; Kao, A.; Dillon, R.; Croner, L. J.; Benz, R.; Blume, J. E.; Wilcox, B., A large-scale and robust dynamic MRM study of colorectal cancer biomarkers. *J Proteomics* **2018,** *187*, 80-92.
62. Bos, J. L.; Rehmann, H.; Wittinghofer, A., GEFs and GAPs: critical elements in the control of small G proteins. *Cell* **2007,** *129* (5), 865-77.
63. Cherfils, J.; Zeghouf, M., Regulation of small GTPases by GEFs, GAPs, and GDIs. *Physiol Rev* **2013,** *93* (1), 269-309.
64. Wennerberg, K.; Rossman, K. L.; Der, C. J., The Ras superfamily at a glance. *J Cell Sci* **2005,** *118* (5), 843-846.
65. Colicelli, J., Human RAS superfamily proteins and related GTPases. *Sci STKE* **2004,** *2004* (250), RE13.
66. Yang, Z., Small GTPases: versatile signaling switches in plants. *Plant Cell* **2002,** *14 Suppl*, S375-88.
67. Jaffe, A. B.; Hall, A., Rho GTPases: biochemistry and biology. *Annu Rev Cell Dev Biol* **2005,** *21*, 247-69.
68. Ridley, A. J., Rho GTPases and actin dynamics in membrane protrusions and vesicle trafficking. *Trends Cell Biol* **2006,** *16* (10), 522-9.
69. Bar-Sagi, D.; Hall, A., Ras and Rho GTPases: a family reunion. *Cell* **2000,** *103* (2), 227-38.
70. Raftopoulou, M.; Hall, A., Cell migration: Rho GTPases lead the way. *Dev Biol* **2004,** *265* (1), 23-32.
71. Nielsen, E.; Cheung, A. Y.; Ueda, T., The regulatory RAB and ARF GTPases for vesicular trafficking. *Plant Physiol* **2008,** *147* (4), 1516-26.
72. Hutagalung, A. H.; Novick, P. J., Role of Rab GTPases in membrane traffic and cell physiology. *Physiol Rev* **2011,** *91* (1), 119-49.
73. Sazer, S., The search for the primary function of the Ran GTPase continues. *Trends Cell Biol* **1996,** *6* (3), 81-5.
74. Simanshu, D. K.; Nissley, D. V.; McCormick, F., RAS Proteins and Their Regulators in Human Disease. *Cell* **2017,** *170* (1), 17-33.

75. Recchi, C.; Seabra, M. C., Novel functions for Rab GTPases in multiple aspects of tumour progression. *Biochem Soc Trans* **2012,** *40* (6), 1398-403.
76. Porter, A. P.; Papaioannou, A.; Malliri, A., Deregulation of Rho GTPases in cancer. *Small GTPases* **2016,** *7* (3), 123-38.
77. Bollinger, J. G.; Stergachis, A. B.; Johnson, R. S.; Egertson, J. D.; MacCoss, M. J., Selecting Optimal Peptides for Targeted Proteomic Experiments in Human Plasma Using In Vitro Synthesized Proteins as Analytical Standards. *Methods Mol Biol* **2016,** *1410*, 207-21.
78. Sriram, K.; Insel, P. A., G Protein-Coupled Receptors as Targets for Approved Drugs: How Many Targets and How Many Drugs? *Mol Pharmacol* **2018,** *93* (4), 251-258.
79. Rosenblum, J. S.; Nomanbhoy, T. K.; Kozarich, J. W., Functional interrogation of kinases and other nucleotide-binding proteins. *Febs Lett* **2013,** *587* (13), 1870-1877.
80. Dever, T. E.; Glynias, M. J.; Merrick, W. C., GTP-binding domain: three consensus sequence elements with distinct spacing. *Proc Natl Acad Sci U S A* **1987,** *84* (7), 1814-8.
81. Patricelli, M. P.; Szardenings, A. K.; Liyanage, M.; Nomanbhoy, T. K.; Wu, M.; Weissig, H.; Aban, A.; Chun, D.; Tanner, S.; Kozarich, J. W., Functional interrogation of the kinome using nucleotide acyl phosphates. *Biochemistry* **2007,** *46* (2), 350-8.
82. Qiu, H.; Wang, Y., Probing adenosine nucleotide-binding proteins with an affinity-labeled nucleotide probe and mass spectrometry. *Anal Chem* **2007,** *79* (15), 5547-56.
83. Xiao, Y.; Guo, L.; Jiang, X.; Wang, Y., Proteome-wide discovery and characterizations of nucleotide-binding proteins with affinity-labeled chemical probes. *Anal Chem* **2013,** *85* (6), 3198-206.
84. Xiao, Y.; Ji, D.; Guo, L.; Wang, Y., Comprehensive characterization of (S)GTP-binding proteins by orthogonal quantitative (S)GTP-affinity profiling and (S)GTP/GTP competition assays. *Anal Chem* **2014,** *86* (9), 4550-8.
85. Hunter, J. C.; Gurbani, D.; Ficarro, S. B.; Carrasco, M. A.; Lim, S. M.; Choi, H. G.; Xie, T.; Marto, J. A.; Chen, Z.; Gray, N. S.; Westover, K. D., In situ selectivity profiling and crystal structure of SML-8-73-1, an active site inhibitor of oncogenic K-Ras G12C. *Proc Natl Acad Sci U S A* **2014,** *111* (24), 8895-900.
86. Cai, R.; Huang, M.; Wang, Y., Targeted Quantitative Profiling of GTP-Binding Proteins in Cancer Cells Using Isotope-Coded GTP Probes. *Anal Chem* **2018,** *90* (24), 14339-14346.
87. Montgomery, D. C.; Sorum, A. W.; Meier, J. L., Chemoproteomic profiling of lysine acetyltransferases highlights an expanded landscape of catalytic acetylation. *J Am Chem Soc* **2014,** *136* (24), 8669-76.
88. Kaneda, M.; Masuda, S.; Tomohiro, T.; Hatanaka, Y., A simple and efficient photoaffinity method for proteomics of GTP-binding proteins. *Chembiochem* **2007,** *8* (6), 595-8.
89. George Cisar, E. A.; Nguyen, N.; Rosen, H., A GTP affinity probe for proteomics highlights flexibility in purine nucleotide selectivity. *J Am Chem Soc* **2013,** *135* (12), 4676-9.
90. Zhang, C. C.; Li, R.; Jiang, H.; Lin, S.; Rogalski, J. C.; Liu, K.; Kast, J., Development and application of a quantitative multiplexed small GTPase activity assay using targeted proteomics. *J Proteome Res* **2015,** *14* (2), 967-76.

91. Halvey, P. J.; Ferrone, C. R.; Liebler, D. C., GeLC-MRM quantitation of mutant KRAS oncoprotein in complex biological samples. *J Proteome Res* **2012,** *11* (7), 3908-13.

Chapter 2 A Targeted Quantitative Proteomic Approach Assesses the Reprogramming of Small GTPases during Melanoma Metastasis

2.1 Introduction

Small GTPases of the Ras superfamily are highly conserved in eukaryotes, including more than 100 members that could be divided into six subfamilies, i.e. Ras, Rho, Rab, Sar1/Arf, Ran and others.[1] They can exist in the GTP-bound active state or GDP-bound inactive state, which are modulated by guanine nucleotide exchange factors (GEFs), GTPase-activating proteins (GAPs) and guanine nucleotide dissociation inhibitors (GDIs).[2] Small GTPases serve as master regulators of cellular trafficking and are involved in numerous cell signaling cascades.[1, 3] In addition, emerging evidence has linked aberrant expression of small GTPases with cancer progression, including RHOC and RAB27A in melanoma.[4-5]

Despite the importance of small GTPases in cell signaling and human diseases, very few studies have been conducted to assess quantitatively the small GTPases at the proteome-wide scale. In recent years, multiple-reaction monitoring (MRM)-based targeted proteomic method has emerged as a powerful approach for analyzing proteins and peptides of interest with high specificity and sensitivity.[6-7] We reason that a targeted proteomic method for the measurement of small GTPases may enable mechanistic studies of small GTPase signaling and facilitate the discovery of novel roles of small GTPases in the etiology of human diseases.

In the present study, we developed a facile and effective MRM-based method for high-throughput profiling of small GTPases in cultured human cells and we also applied the method for assessing the roles of small GTPases in melanoma metastasis. We chose to examine the roles of small GTPases in melanoma metastasis because melanoma is one of the most aggressive and treatment-resistant types of human cancers. In this vein, an estimated 91,270 new cases of melanoma and 9320 deaths are expected in the United States in 2018,[8] and the high mortality rate of melanoma is attributed to its high probability to metastasize.[9]

2.2 Materials and Methods

2.1.1 Cell Culture

HCT-116 human colorectal cancer cells, HEK293T human embryonic kidney cells, HL-60 human promyelocytic leukemia cells, Jurkat-T human T lymphocytic leukemia cells, MCF-7 human breast cancer cells, WM-115 and WM-266-4 human melanoma cells were purchased from American Type Culture Collection (ATCC; Manassas, VA). GM00637 human skin fibroblasts were kindly provided by Prof. Gerd P. Pfeifer (the City of Hope). IGR39 and IGR37 human melanoma cells were generous gifts from Prof. Peter H. Duesberg (University of California, Berkeley). WM793 and 1205Lu human melanoma cells were purchased from Wistar Institute. HCT-116, HEK293T, GM00637, MCF-7, WM-115 and WM-266-4 cells were cultured in Dulbecco's Modified Eagle Medium (DMEM; Invitrogen-Gibco, Carlsbad, CA). HL-60, Jurkat-T, IGR39, IGR37, WM793 and 1205Lu cells were cultured in RPMI 1640 Medium (Invitrogen-Gibco). All culture media were supplemented with 10% fetal bovine serum (FBS; Invitrogen-Gibco) and

penicillin/streptomycin (100 IU/mL). Cells were maintained at 37°C in a humidified atmosphere containing 5% CO_2, and the culture medium was changed in every 2 to 3 days as necessary.

The initial passage numbers for melanoma cells used were: WM-115 (p9), WM-266-4 (p6), IGR39 (p4), IGR37 (p7), WM793 (p16), and 1205Lu (p70). All the relevant experiments were conducted within 20 passages from revival of the initial frozen seeds. LookOut Mycoplasma PCR Detection Kit (MP0035, Sigma-Aldrich, MO) for detection of 19 mycoplasma species was used following the manufacturer's instructions. PCRs were performed using HotStart Taq Polymerase. Results were visualized on a 1.2% agarose gel, where mycoplasma-positive samples would show a band at 261 bp, and internal control DNA showed a band at 500 bp. WM-115, WM-266-4, IGR39, IGR37, WM793, and 1205Lu melanoma cell lines were tested by this method to be free of mycoplasma on May 9, 2018. In addition, these six melanoma cell lines were authenticated by ATCC on May 24, 2018 using Short Tandem Repeat (STR) analysis as described in 2012 in ANSI Standard (ASN-0002) Authentication of Human Cell Lines.

For SILAC experiments, $[^{13}C_6,^{15}N_2]$-$_L$-lysine and $[^{13}C_6]$-$_L$-arginine (Cambridge Isotopes Inc., MA), or the corresponding unlabeled lysine and arginine, were added to SILAC DMEM media depleted of $_L$-lysine and $_L$-arginine (Thermo Scientific Pierce, MA) until their final concentrations reached 0.398 and 0.798 mM, respectively, to yield the "heavy" and "light" media. The SILAC RPMI-1640 media were prepared in a similar fashion except that the final concentrations of the added lysine and arginine were 0.274 mM and 1.15 mM, respectively. The SILAC media were again supplemented with 10%

dialyzed FBS (Corning, NY). WM-115 and WM-266-4 cells were cultured in the heavy-DMEM medium, and IGR39, IGR37, WM793 and 1205Lu cells were cultured in the heavy-RPMI medium for at least six cell doublings to ensure complete heavy-isotope incorporation.

2.1.2 Gene Ontology (GO) Analysis and Data Source for Bioinformatic Analyses

Gene Ontology analyses were conducted using the web-based Database for Annotation, Visualization and Integrated Discovery (DAVID, version 6.7; https://david.ncifcrf.gov/).[10] Patient RNAseq data were obtained from The Cancer Genome Atlas (TCGA) via cBioPortal [11]. We used data from 458 melanoma patients in the TCGA-SKCM project for bioinformatic analyses. The Cancer Cell Line Encyclopedia (CCLE) (http://www.broadinstitute.org/ccle/home) were employed for the comprehensive evaluation of mRNA expression for candidate genes among more than 1,000 cell lines representing 37 cancer types [12]. Multi-tumor *RAB38* mRNA expression box plot and scatter plot for melanoma cell lines were retrieved from the CCLE database using cBioPortal. Publicly available transcriptomic profiles with accession numbers GSE7553, GSE7929, GSE8401, GSE22153, GSE44662, GSE46522 and GSE70621 were downloaded from the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) database and analyzed using R (version 3.4.3).

2.1.3 Sample Preparation and LC-MS/MS for Shotgun Proteomic Analysis

For acquiring shotgun proteomic data to establish an MRM library for small GTPases, the individually collected cell pellets from 9 human cell lines were lysed, and the resulting cell lysates were collected by centrifugation and separated by electrophoresis using a 16%

SDS-PAGE gel. The gel bands in the molecular weight range of 15−37 kDa were cut into 7 pieces, reduced with dithiothreitol, alkylated with iodoacetamide, and in-gel digested with trypsin at an enzyme/protein ratio of 1:100.

In shotgun proteomic experiments, peptide samples were subjected to LC-MS/MS analysis on either an LTQ Orbitrap Velos mass spectrometer or a Q Exactive Plus mass spectrometer that was equipped with a nanoelectrospray ionization source and coupled to an EASY-nLC II (Thermo, San Jose, CA), as described previously.[63] A homemade trapping column (150 μm × 50 mm), packed with ReproSil-Pur C18-AQ resin (5 μm in particle size and 120 Å in pore size, Dr. Maisch GmbH HPLC), and an analytical column (75 μm × 120 mm), packed with ReproSil-Pur 120 C18-AQ resin (3 μm in particle size and 120 Å in pore size, Dr. Maisch GmbH HPLC), were used for peptide separation. Peptide samples were initially loaded onto the trapping column with 0.1% formic acid in water at a flow rate of 3.0 μL/min. The peptides were then separated using a 157-min linear gradient of 2−35% acetonitrile in 0.1% formic acid at a flow rate of 230 nL/min. The mass spectrometer was operated in data-dependent acquisition (DDA) mode, where one full MS scan (resolution = 60,000) at an automatic gain control (AGC) target of 10^6 was followed with up to 20 MS/MS scans for the most abundant ions observed in MS. The selected ions were excluded from further analysis for 90 s. Ions with singly or unassigned charge states were not fragmented. The maximum ion accumulation time was 1000 ms for each full MS scan, and 50 ms for each MS/MS scan. The raw data were searched against the IPI human database (version 3.68, 87,061 entries). For database search, up to one trypsin missed

cleavage per peptide was allowed, and MS/MS spectral assignment was performed with less than 1% false discovery rate (FDR).

2.1.4 Retention Time (RT) Prediction for Small GTPase Peptides

To facilitate automated and multiplexed MRM analyses, we generated an interactive MRM spectral library file containing the tandem mass spectra of peptides derived from small GTPases, which were acquired from the above-described shotgun proteomic experiments, using Skyline (version 3.5). Precursor and fragment ions for targeted GTPase peptides were selected and further refined to avoid spectral interference. If applicable, three unique peptides were selected for subsequent targeted analysis in the MRM mode.[13] In addition, we employed the aforementioned iRT algorithm to develop the scheduled MRM method.[14] To calculate the iRT score for each targeted peptide from small GTPases in the MRM spectral library, we selected 10 peptides from the tryptic digestion mixture of bovine serum albumin (BSA) as reference peptides to set up an iRT scale. By manually setting the iRT values of the BSA peptides AEFVEVTK and DAFLGSFLYEYSR as 0 and 100, respectively, a linear regression equation was obtained by plotting the iRT scores against their empirically measured retention times (RTs) derived from the LC-MS/MS analyses. Subsequently, the iRT values for the other eight standard BSA peptides were determined using the above-mentioned linear equation along with their empirically measured RTs in the shotgun proteomic experiments.

The BSA peptide mixture was then spiked into the small GTPase peptide mixtures for the LC-MS/MS analyses on an LTQ Orbitrap Velos mass spectrometer equipped with an Easy-nLC II system with a 150-min linear gradient of 2–35% buffer B (0.1% formic acid

in acetonitrile) in buffer A (0.1% formic acid), or a Q Exactive Plus mass spectrometer equipped with an Easy-nLC 1000 system with a 172-min linear gradient of 8–35% buffer B (80% acetonitrile in water with 0.1 % formic acid) in buffer A (0.1% formic acid). RTs were extracted for all BSA standard peptides as well as targeted peptides using the Skyline MS1 filtering workflow. The iRT values for all small GTPase peptides were calculated based on linear regression of iRT and experimentally measured RT of peptides with previously determined iRT scores.

2.1.5 Sample Preparation and Scheduled LC-MRM Analysis

To rigorously examine the differential expression of small GTPases in paired primary/metastatic melanoma cells, we conducted forward- and reverse-SILAC experiments, where the light and heavy lysates of paired primary melanoma cells (WM-115, IGR39 or WM793) were mixed with the heavy and light lysates from metastatic melanoma cells (WM-266-4, IGR37 or 1205Lu), respectively, at a 1:1 ratio (by mass). The mixed cell lysates were then loaded onto a 10% SDS-PAGE gel and separated by electrophoresis. The gel bands corresponding to the molecular weight range of 15–37 kDa were cut, reduced with 20 mM dithiothreitol, alkylated with 55 mM iodoacetamide, and in-gel digested with trypsin at an enzyme/protein ratio of 1:100.

All LC-MRM analyses were performed on a TSQ Vantage triple-quadrupole mass spectrometer equipped with an Easy-nLC II system. The peptide mixtures were separated with a 157-min linear gradient of 2–35% acetonitrile in buffer B (0.1% formic acid in acetonitrile) in buffer A (0.1% formic acid) and at a flow rate of 230 nL/min. The spray voltage was 1.8 kV, Q1 and Q3 resolutions were 0.7 Da, and the cycle time was 5 s. The

optimal collisional energy (CE) set for each targeted peptide was calculated using a linear equation specific to the TSQ Vantage instrument and the precursor mass-to-charge ratio (m/z) according to the default setting in Skyline.

The tryptic peptide mixture of BSA was subjected to unscheduled LC-MRM analyses prior to the analyses of targeted peptides of small GTPases in order to calibrate for possible retention time shifts due to changes in chromatographic conditions. The prediction of empirical RT based on the individual iRT scores of the targeted small GTPase peptides was then determined by the linear regression of RTs of BSA standard peptides obtained under the same chromatography conditions. This iRT-based predictor was examined between every six MRM runs by injecting another BSA reference peptide mixture to ensure precise RT prediction.

A total of 432 peptides representing 113 non-redundant or 131 isoform-specific small GTPases were monitored in each scheduled LC-MRM run. For SILAC labeling experiments, the LC-MRM method was further expanded into two runs for monitoring the precursor and fragment ions of peptides harboring the light and heavy forms of lysine and/or arginine. In this regard, three transitions were monitored for each light-/heavy-isotope-coded peptide for LC-MRM quantification. All the targeted transitions can be monitored in two LC runs by using the scheduled MRM mode with a 6-min retention time duration.

All raw files were processed using Skyline (version 3.5) for the generation of extracted-ion chromatograms and peak integration. The targeted peptides were first manually checked to ensure the overlaid chromatographic profiles of multiple fragment ions derived

from the light and heavy forms of the same peptide. The data were then processed to ensure that the distribution of the relative intensities of multiple transitions associated with the same precursor ion correlates with the theoretical distribution derived from the library tandem mass spectra acquired in shotgun proteomic experiments. To ensure reliable peak assignment, dot plot or dot product (dotp) values were calculated by comparison of transition ion intensities based on a linear regression model. In this regard, the dotp value has to exceed 0.80.[15] In addition, the iRT values act an intrinsic property (i.e., hydrophobicity) of a peptide; hence, a substantial deviation of measured RT from that projected from the linear plot of RT over iRT signals a false-positive detection.[14] The sum of peak area from all transitions of light- or heavy-labeled peptides was used for the quantification.

2.1.6 Construction and Transfection of FLAG-tagged Fusion Protein Expression Plasmid

To generate expression plasmid for 3×FLAG-tagged RAB38, the full-length coding sequence of human *RAB38* gene was amplified by PCR using Phusion High-Fidelity DNA Polymerase (NEB). The PCR product was digested with HindIII and EcoRI, and the resulting DNA fragment was inserted into p3×FLAG-CMV-10 (Sigma) to construct the plasmid. The sequence of the constructed plasmid was validated by Sanger sequencing.

For plasmid transfection, WM-115 cells were cultured in six-well plates and transfected at 70% confluency with either the p3×FLAG-CMV-10 empty vector or the FLAG-RAB38 plasmid using TransIT-2020 transfection reagent (Mirus Bio LLC) following the manufacturer's recommended protocol. Transfected cells were incubated for an additional

24 h and analyzed for gene expression of the relative targeted protein using immunoblotting with specific antibodies.

2.1.7 siRNA Transfection

siGENOME non-targeting (NT) siRNA control (D-001210-02-05) and RAB38 SMARTpool (L-010059-00-0005) siRNA were purchased from Dharmacon (Lafayette, CO). The non-targeting siRNA control was designed not to target any known genes in human, mouse or rat, as described by the manufacturer. Briefly, WM-266-4, IGR37 and M14 melanoma cells were cultured in six-well plates and transfected at 50% confluency with 100 nM non-targeting siRNA or RAB38-targeting siRNA using TransIT-X2 transfection reagent (Mirus Bio LLC) following the manufacturer's protocol. Transfected cells were incubated for an additional 48–72 h and analyzed for the expression level of the RAB38 protein using immunoblotting with specific antibodies.

2.1.8 Real-time Quantitative PCR (RT-qPCR)

At 24–72 h following transfection, total RNA was extracted from the cells using the Total RNA Kit I (Omega Bio-Tech), and 1 μg of total RNA was reverse-transcribed to generate cDNA by using M-MLV reverse transcriptase (Promega) and an oligo$(dT)_{16}$ primer. Gene expression levels were evaluated by RT-qPCR using iQ SYBR Green Supermix kit (Bio-Rad) with 100 ng cDNA input in a 20-μL total reaction volume. Expression level of the glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*) gene was used as an internal control for normalization. All primers were used at a final concentration of 500 nM. The relative gene expression was analyzed by the comparative cycle threshold ($2^{-\Delta\Delta Ct}$) method.

2.1.9 Immunoblotting

Total protein was extracted from cell pellet using ice-cold CelLytic M cell lysis reagent (Sigma-Aldrich, MO) containing 1% (v/v) protease inhibitor cocktail (Sigma-Aldrich, MO). After cell lysis, the protein concentration was determined by the Quick Start™ Bradford Protein Assay (Bio-Rad, CA). Approximately 10–50 μg whole cell lysates, mixed with 4×Laemmli SDS loading buffer, were electrophoresed in 10% SDS-PAGE gels and transferred to nitrocellulose membranes. The membranes were incubated with primary antibodies against human RAB12 (Thermo Fisher; rabbit polyclonal, 1:2,000), RAB27A (Abcam; rabbit polyclonal, 1:5,000), RAB31 (4D12, Santa Cruz; rabbit polyclonal, 1:2,000), RAB32 (Thermo Fisher; rabbit polyclonal, 1:2,000), RAB38 (A-8, Santa Cruz; mouse polyclonal, 1:2,000), MITF (D-9, Santa Cruz; mouse polyclonal, 1:5,000), or β-actin (Thermo Fisher; rabbit polyclonal, 1:10,000), followed by incubation with peroxidase-labeled donkey anti-rabbit secondary antibody (Thermo Fisher; 1:10,000) or mouse m-IgGκ BP-HRP (Santa Cruz; 1:10,000). Amersham ECL Prime Western Blot Detecting Reagent (GE Healthcare, CA) was used to visualize the protein bands.

2.1.10 Migration and Invasion Assays

For transwell migration assay, cells ($0.5-1\times10^5$) were placed in the upper chamber of transwell inserts (Corning, NY) with serum-free DMEM medium. DMEM medium containing 10% FBS was added to the lower chamber as chemoattractants and the cells were incubated at 37°C for 24 h. After removal of unmigrated cells, the cells attached to the reverse side of the membrane were stained with 0.5% crystal violet, and 5 randomly selected fields were counted under an inverted microscope in each experiment. The

invasion assay was conducted under the same conditions except that the transwell membranes were pre-coated with Matrigel (Corning, NY).

2.1.11 Gelatin Zymography Assay

At 24 h following plasmid transfection or 72 h following siRNA transfection, the culture medium was removed, and the cells were washed twice with, and reconstituted in, serum-free DMEM medium. After a 24-h incubation, conditioned medium (CM) was collected by centrifugation to remove cell debris. The collected CM was further concentrated using Microcon centrifugal filter units with a molecular weight cutoff of 30 kDa (EMD Millipore, CA) and the Quick Start™ Bradford Protein Assay was used to determine the total protein concentration. Subsequently, 5–10 μg total CM proteins were separated using 7.5% SDS-PAGE gels containing 0.1% gelatin. After electrophoresis, the gels were incubated with zymography washing buffer (2.5% Triton X-100, 50 mM Tris-HCl, pH 7.5) at room temperature for 1 h to remove excess SDS and renature the matrix metalloproteinases (MMPs). The gels were then incubated at 37°C for 24 h in zymography developing buffer (1.0% Triton X-100, 50 mM Tris-HCl, pH 7.5) to induce gelatin digestion by the renatured enzymes. The gels were subsequently stained with 0.5% Coomassie blue G-250 and destained until clear bands were visible against the dark background, indicative of proteolytic activities of MMPs.

2.1.12 Chromatin Immunoprecipitation (ChIP) and RT-qPCR

For ChIP, approximately 1×10^7 WM-115 and WM-266-4 cells were harvested and fixed in PBS with 1% formaldehyde at room temperature for 10 min. After cross-linking, the cell pellets were resuspended in 1 mL of lysis buffer (50 mM Tris-HCl, pH 8.0, 150

mM NaCl, 5 mM EDTA, 1% Triton X-100, and 0.1% sodium deoxycholate) and sonicated to obtain DNA fragments of 300–500 bp in length. Anti-MITF antibody (ab12039, Abcam) or normal IgG (2729S, Cell Signaling Technology) was used to precipitate the chromatin. The precipitated DNA was purified using the QIAquick PCR purification Kit (Qiagen, MD) and used for RT-qPCR analysis. The primers used in RT-qPCR coupled to ChIP were: *RAB38* forward, GCCACAAACTTGTGAGGTGT; *RAB38* reverse, CTTCAGACCTGTGGTCAACG; *TBC1D16* forward, GGCCACATACAAAGGGATCG; *TBC1D16* reverse, CTCGCGGAGGCAATCTGA.

2.1.13 Bisulfite Sequencing

Approximately 5×10^3 cells collected from six melanoma cell lines: WM-115, WM-266-4, IGR39, IGR37, WM793 and 1205Lu, respectively, were lysed and treated with bisulfite using the EZ DNA Methylation-Direct Kit (Zymo Research, CA). The resulting DNA was subsequently amplified using Zymo*Taq* DNA Polymerase (Zymo Research, CA). RAB38 primers were designed using the MethPrimer 2.0 online tool (http://www.urogene.org/methprimer2/) to amplify the 179 bp fragment of the promoter region of the bisulfite converted-RAB38 gene (Chr11: 87,908,686–87,908,864, UCSC Genome Browser Human Feb. 2009 Assembly, GRCh37/hg19). The outer PCR was set up using the following primers: *RAB38* forward primer, 5′-GGTTAGGGTTATAGGTGAAAATAGT-3′, *RAB38* reverse primer, 5′-AACTCCTCCCCTAAAAATTAATCC-3′. The reaction mixture was heated to 95°C for 10 min followed by 40 cycles, denaturing at 95°C for 30 sec, annealing at 55°C for 45 sec and elongating at 72°C for 1 min, followed by a final elongation step at 72°C for 7 min.

The successful amplification and the right size of the amplicon were validated using 2% agarose gel. Amplicons were purified using the E.Z.N.A. Gel Extraction Kit (Omega Bio-tek, GA). The purified PCR products were cloned into the pGEM®-T vector (Promega Corporation, WI) and the ligation products were selected by blue/white colony screening. Ten white colonies selected for each cell line were grown in liquid Lysogeny broth (LB) media overnight and the plasmids were extracted using GeneJET Plasmid Miniprep kit (Thermo Fishier Scientific, MA). The plasmids were then subjected to Sanger sequencing and the methylation status of each individual CpG dinucleotides in the amplicon was subsequently determined. Sequencing results were analyzed by the BiQ Analyzer software (http://biq-analyzer.bioinf.mpi-inf.mpg.de/) to generate the lollipop-representation.

2.1.14 5-Aza-2′-deoxycytidine (5-Aza) Treatment.

WM-115 and IGR39 cells were seeded at a density of 2×10^4 cells/mL and treated at 24 h later with 5 μM 5-Aza (Sigma-Aldrich, MO). Dimethyl sulfoxide (DMSO) was used as the vehicle control. The cells were replenished with freshly prepared DMSO/5-Aza in complete growth medium in every 24 h for up to 96 h of treatment (four pulses). Total RNAs were isolated for cDNA conversion and RT-qPCR analyses.

2.3 Results

2.3.1 Development of a High-throughput LC-MRM Assay for Targeted Quantitative Profiling of Small GTPases in Cultured Human Cancer Cells

We set out to develop a high-throughput, multiplexed MRM-based targeted proteomics method for interrogating small GTPases in the entire human proteome. In this context, Halvey *et al.* [16] developed an LC-MRM method for quantifying wild-type and mutant K-

RAS proteins in cultured cancer cells and pancreatic cyst fluids after enrichment of low-molecular weight (20–25 kDa) proteins using sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). In addition, Zhang *et al.* [17] described the use of MRM in combination with SDS-PAGE-based enrichment of proteins in the molecular weight range of 15–25 kDa to measure simultaneously the activities of 12 small GTPases after affinity enrichment using the GTPase-binding domains of four effector proteins. Building upon these previous studies, we developed an SDS-PAGE fractionation coupled with LC-MRM workflow for targeted quantification of small GTPases at the entire proteome scale.

A high-throughput LC-MRM workflow for the proteome-wide interrogation of small GTPases requires the collection of tandem mass spectra and chromatographic retention time of unique (or signature) peptides derived from the targeted small GTPases. Because the expression of small GTPases differs among different cell lines, we established an MRM spectral library based on the data collected from shotgun proteomic analyses of tryptic digestion mixtures of low-molecular weight proteins (15–37 kDa) from the lysates of 9 human cell lines of different tissue origins. These included GM00637 (skin), HCT-116 (colon), HEK293T (kidney), HL-60 (peripheral blood), Jurkat T (peripheral blood), K562 (bone marrow), MCF-7 (breast), WM-115 (skin), and WM-266-4 (skin). To this end, we fractionated the whole-cell protein lysate using SDS-PAGE, excised the gel bands in the molecular weight region of 15–37 kDa, reduced the cysteine residues in proteins with dithiothreitol and alkylated them with iodoacetamide. The proteins were then digested in gel with trypsin and the resulting peptide mixtures subjected to LC-MS/MS analysis in the DDA mode. The identified proteins (>5000) were then filtered using the DAVID

bioinformatic tool with the Gene Ontology (GO) term of "small GTPase".[10] The tandem mass spectra of all peptides from small GTPases along with their retention time information were subsequently imported into Skyline (version 3.6)[18] to establish the MRM spectral library.

To achieve reliable MRM-based quantification, we selected an average of three peptides that are unique for each small GTPase, and when necessary, to its specific isoform(s). To maximize selectivity and sensitivity for the MRM measurements, we chose the transitions corresponding to the formation of the three most abundant y-ions based on the MS/MS acquired from shotgun proteomic analyses (Figure 2.1).[19] The complete list of small GTPases of the Ras superfamily was organized according to the individual subfamilies (Ras, Rho, Rab, Sar1/Arf, Ran, and others). The current version of the MRM spectral library encompassed 432 distinct peptides representing 113 non-redundant small GTPases encoded by unique genes. To our knowledge, this is the first targeted proteomic method developed for profiling comprehensively the Ras superfamily of small GTPases.

To increase the throughput of the assay, we employed scheduled LC-MRM with the use of normalized retention time (iRT) [14]. The iRT is a dimensionless score for a peptide derived from its empirical retention time observed in shotgun proteomic analysis and the retention times for a set of standard peptides analyzed under the same LC conditions. In scheduled LC-MRM analysis, the mass spectrometer could be scheduled to collect subsets of transitions in predefined retention time windows according to the chromatographic setup, where the retention times for the targeted peptides were predicted from their iRT values in the library and from the actual retention times observed for the standard peptides.

In doing so, we established a robust and high-throughput MRM-based targeted proteomic workflow for the Ras superfamily of small GTPases, where the 432 unique peptides from small GTPases could be monitored in a single LC-MRM run.

2.3.2 Scheduled LC-MRM Analysis Revealed Differential Expression of Small GTPases in Paired Primary/Metastatic Melanoma Cells

Paired cell lines derived from the same cancer patients are powerful resources for investigating the mechanisms of cancer progression. Here we employed three pairs of primary/metastatic melanoma cell lines for the targeted analyses of small GTPases: The "WM" pair consists of WM-115 and WM-266-4, which were derived from the primary tumor site and the right thigh skin metastatic site of the same melanoma patient, respectively;[20] the "IGR" pair is comprised of IGR39 and IGR37, which were respectively derived from the primary tumor site and the groin metastatic site of another individual;[21] in the "WMLu" pair, WM793 was initiated from a superficial spreading melanoma, and 1205Lu was derived from a lung metastasis of WM793 cells after subcutaneous injection into the tail vein of an immune-deficient mouse.[22]

We employed stable isotope-labeling by amino acid in cell culture (SILAC),[23] in conjunction with the above-described SDS-PAGE fractionation and LC-MRM analysis, for assessing the differential expression of small GTPases in the three pairs of matched primary/metastatic melanoma cells (Figure 2.1). To this end, we first modified the Skyline MRM library by incorporating the corresponding transitions for the "heavy" forms of precursor and fragment ions. By using this approach, we were able to quantify approximately 100 small GTPases in each of the three paired melanoma cell lines. Among

the 101 small GTPases quantified for the WM pair (Figure 2.2), 14 and 10 were substantially up- and down-regulated (by at least 1.5-fold), respectively, in the metastatic (WM-266-4) relative to the primary (WM-115) melanoma cells (Figure 2.3). In addition, among the 93 small GTPases quantified for the IGR pair, 20 and 12 were considerably up- and down-regulated, respectively, in the metastatic (IGR37) relative to the primary (IGR39) melanoma cells (Figure 2.3). Of the 93 small GTPases quantified for the third pair, 9 and 24 were up- and down-regulated by at least 1.5-fold in the metastatic (1205Lu) compared to the primary (WM793) melanoma cells, respectively (Figure 2.3).

We also explored the similarities and differences in the expression profiles of small GTPases by hierarchical clustering analysis of the quantitative proteomics data. It turned out that the features in differential expression of small GTPases, induced by metastatic transformation, were more similar for the WM and IGR pairs than between either of the two pairs and the WMLu pair (Figure 2.3). This might be attributed in part to the differences in transcriptional and/or epigenetic regulations in the three pairs of melanoma cell lines (*vide infra*) and the fact that the metastatic lines in the first two pairs were derived from melanoma patients, whereas that of the last pair was obtained from experimental metastasis in mouse, as noted above.

It is worth noting that, by utilizing the iRT algorithms, we were able to accurately predict the actual retention times for the targeted peptides with the use of a 6-min retention time window. As depicted in Figure 2.4, the R^2 values were 0.992 and 0.996 for LC-MRM measurements of peptide samples obtained from the WM-115/WM266-4 and

IGR39/IGR37 paired cell lines, respectively. Moreover, this method also displayed excellent reproducibility between different replicates.

For comparison, we also analyzed the peptide samples from the WM-115/WM-266-4 cells using shotgun proteomic approach on an LTQ Orbitrap Velos mass spectrometer. The results showed that the LC-MRM method outperformed the shotgun proteomic method in reproducibility and sensitivity, the latter of which is reflected by the pronouncedly larger numbers of small GTPases quantified by the former method (Figure 2.4). The excellent reproducibility of the MRM-based quantification is manifested by the observation that 101 small GTPase proteins could be reproducibly quantified in all three sets of SILAC labeling experiments. In contrast, among the 59 small GTPases detected by the shotgun proteomic method, 9 and 4 were exclusively detected in the forward- and reverse-SILAC experiments, respectively (Figure 2.4). Taken together, the established targeted proteomic workflow provided excellent sensitivity and reproducibility, and it allowed for robust and high-throughput quantifications of small GTPases in melanoma cells.

2.3.3 Targeted Proteomics Revealed the Up-regulation of RAB27A and RAB38 in WM-266-4 and IGR37 Metastatic Melanoma Cells

One goal of the present study was to uncover small GTPases that drive and/or suppress melanoma metastasis. Hence, we expected to confirm the differential expression of some previously reported drivers and/or suppressors for melanoma metastasis. In this vein, RAB27A was shown to promote melanoma metastasis through the regulation of the MET network.[24] Indeed, our LC-MRM data revealed significant up-regulations of RAB27A in the WM-266-4 and IGR37 metastatic melanoma cell lines relative to the corresponding

primary melanoma cells, though similar observation was not made for the WMLu pair (Figure 2.5).

Our LC-MRM quantification data showed that 14, 20 and 12 small GTPases were differentially expressed by at least two-fold in the metastatic (i.e. WM-266-4, IGR-37, and 1205Lu) compared to the corresponding paired primary (i.e. WM-115, IGR-39, and WM793) melanoma cells. Among other differentially expressed small GTPases, RAB38 was expressed at much lower levels in two out of the three primary melanoma cell lines (WM-115 and IGR39) than the corresponding metastatic lines (WM-266-4 and IGR37), as determined from LC-MRM analyses and confirmed by Western blot analyses (Figure 2.6). RAB38 was, however, not detectable in the WM793 or 1205Lu cells by LC-MRM or Western blot analyses (Figure 2.6). We also validated, by using Western blot analyses, the LC-MRM quantification results for several other small GTPases, including RAB12, RAB31, and RAB32 (data not shown**)**. The highly consistent results obtained from LC-MRM and Western blot analyses underscored the robustness of the LC-MRM method in assessing quantitatively the differential expression of small GTPase proteins.

2.3.4 Potential Roles of RAB38 in Melanoma Progression

We next asked whether RAB38 expression level modulates prognosis in melanoma patients. We performed Kaplan-Meier survival analysis in melanoma patient cohort in the Cancer Genome Atlas (TCGA) database, and the results showed that poorer patient survival was significantly correlated with higher levels of mRNA expression of *RAB38* gene (hazard ratio, HR = 1.323; 95% confidence interval, 95% CI = 1.009−1.736; Logrank p = 0.0402) (Figure 2.7). Furthermore, pan-cancer analysis of TCGA data using cBio

Cancer Genomics Portal (cBioPortal: http://www.cbioportal.org/)[25] revealed that the mRNA expression levels of *RAB38* gene were highly up-regulated in two types of melanoma (skin cutaneous melanoma, SKCM; uveal melanoma, UVM) compared to other types of cancers (Figure 2.7**)**.

We also queried public databases for the expression levels of *RAB38* gene in other melanoma cell lines. First, analysis of the NCI-60 Human Tumor Cell Lines Database (https://dtp.cancer.gov/discovery_development/nci-60/) revealed the marked up-regulation of *RAB38* gene in various metastatic melanoma cell lines (Figure 2.7). Likewise, the mRNA expression levels of *RAB38* gene were up-regulated in the majority of 61 metastatic melanoma cell lines in the Cancer Cell Line Encyclopedia (CCLE) Database (http://www.broadinstitute.org/ccle/home) (Figure 2.7**)**. Moreover, we utilized publicly accessible transcriptomic profiles in the Gene Expression Omnibus (GEO) database to analyze the mRNA expression of small GTPases in melanoma cells. In particular, we assessed previously published data about the differential gene expression between the highly metastatic human melanoma cell lines derived from an animal metastasis model and the poorly metastatic parental lines (accession number: GSE7929).[26] We found that *RAB38* mRNA levels were again significantly up-regulated in the highly metastatic melanoma cells relative to the poorly metastatic counterparts (Figure 2.7**)**. Taken together, the above results suggested RAB38 as a potential driver for melanoma metastasis.

2.3.5 RAB38 Promotes Invasion of Melanoma Cells through Up-regulation of Matrix Metalloproteinases (MMPs)

We next investigated, by employing transwell migration and invasion assay, whether the invasive phenotypes of melanoma cells could be modulated by the expression levels of *RAB38* gene. Our results showed that ectopic overexpression of RAB38 protein in the WM-115 primary melanoma cells to a similar level as that in the metastatic WM-266-4 cells resulted in a significant increase in the number of invaded cells (Figure 2.8). Reciprocal experiment with the metastatic WM-266-4 cells showed that the siRNA-mediated knockdown of RAB38 led to a significant decline in cell invasion, which is accompanied with a slight diminution of cell migration (Figure 2.8). In this vein, the knockdown efficiency of *RAB38* gene by siRNA was confirmed by both real-time quantitative PCR (RT-qPCR) and immunoblot analysis. Collectively, we demonstrated that RAB38 promotes melanoma invasion *in vitro*.

We also examined the roles of matrix metalloproteinases (MMPs) in RAB38-mediated alterations in invasiveness of melanoma cells. Degradation of extracellular matrix (ECM) proteins by MMPs, a family of zinc- and calcium-dependent proteolytic enzymes, constitutes a crucial initiating step in tumor invasion. Among the 23 members of the human MMP family, MMP2 (gelatinase A) and MMP9 (gelatinase B) are responsible for remodeling the ECM environment and facilitating cancer metastasis.[27] Hence, we explored how the expression levels of RAB38 alter the mRNA expression and enzymatic activities of MMP2 and MMP9.

Consistent with the RT-qPCR results showing the diminished mRNA expression of *MMP2* and *MMP9* genes, gelatin zymography assay showed that the RNAi knockdown of RAB38 led to markedly diminished activities of both the pro-enzyme and active forms of MMP2 and MMP9 in the metastatic WM-266-4 cells (Figure 2.9). This result supports the role of RAB38 in modulating the mRNA expression and activities of MMP2 and MMP9, thereby altering the invasive potential of melanoma cells. In a reciprocal experiment, overexpression of RAB38 induced slight, yet significant increases in the enzymatic activities of MMP2 and MMP9 in WM-115 primary melanoma cells, which were in accordance with the heightened mRNA expression of these two genes as revealed by RT-qPCR analysis (Figure 2.9). Together, these results demonstrated that RAB38 regulates the expression levels and activities of MMP2 and MMP9 in melanoma cells.

Having revealed the regulatory roles of RAB38 in the secretion of MMPs in WM-115 and WM-266-4 melanoma cell lines, we extended our studies to IGR37 and M14 metastatic melanoma cells. In this context, M14 cells were also chosen for the study because these cells displayed pronounced expression of RAB38 (Figure 2.10). Consistent with our hypothesis, RT-qPCR experiments revealed down-regulations of MMP2 and MMP9 after RNAi knockdown of RAB38 in M14 cells (Figure 2.10). In addition, in IGR37 cells, we only observed diminished mRNA levels of MMP9, but not MMP2, after RNAi knockdown of RAB38 (Figure 2.10). Similar as what we observed for WM-266-4 cells, gelatin zymography assay results showed that RAB38 knockdown led to significantly decreased activities of MMP2 and MMP9 in M14 and IGR37 cells (Figure 2.10), lending further evidence to support that RAB38 regulates MMP2 and MMP9 activities in a range of

metastatic melanoma cell lines. The observation of a decreased level of secreted MMP2 protein from IGR37 cells, but not the mRNA expression of the *MMP2* gene in these cells, upon genetic depletion of RAB38 suggests that RAB38 modulates the level of secreted MMP2 through a post-transcriptional mechanism.

2.3.6 Epigenetic Reactivation of RAB38 in Metastatic Melanoma Cells

We next examined the mechanisms through which *RAB38* gene was overexpressed in metastatic over primary melanoma cells. We first asked whether elevated RAB38 expression is accompanied with previously reported genetic alterations in melanoma, including mutations in *BRAF*, *NRAS* and *TP53* genes. It turned out that, in the TCGA SKCM patient cohort, the expression levels of *RAB38* gene did not exhibit any significant correlations with frequently observed mutations in *BRAF, NRAS,* or *TP53* gene.

Numerous studies have underscored the significant roles of epigenetic and transcriptional regulations of oncogenes during cancer progression;[28] hence, we next assessed whether these mechanisms contribute to elevated expression of *RAB38* gene in metastatic melanoma cells. Microphthalmia-associated transcription factor (MITF) is the master regulator of melanocyte development, function, and survival through modulating many genes involved in differentiation and cell cycle progression.[29] In this vein, earlier ChIP-Seq experiments conducted in 501Mel human melanoma cells identified *MITF* loci immediately upstream of the promoters of several Rab GTPase genes, including, among others, *RAB27A* and *RAB38*.[30] Therefore, we next asked whether elevated expression of *RAB38* in the metastatic WM-266-4 and IGR37 cells are due to heightened transcriptional regulation mediated by MITF. We indeed observed the up-regulation of MITF at both the

mRNA and protein levels in the two metastatic lines of melanoma cells (i.e. WM-266-4 and IGR37) relative to the corresponding primary melanoma cells (i.e. WM-115 and IGR39), though the mRNA expression of *MITF* was not detectable in 1205Lu cells.

To explore further the possible functional linkage between *RAB38* expression and *MITF* regulation, we analyzed publicly available data for different cell line and patient cohorts. In cohort GSE7929, both *MITF* and *RAB38* were highly up-regulated at the mRNA levels in the highly metastatic derivatives of A375 human melanoma cells in comparison with the poorly metastatic parental lines. In addition, interrogation of the gene expression data of 120 melanoma cell lines (120Mel) and the Cancer Cell Line Encyclopedia (CCLE) Database (http://www.broadinstitute.org/ccle/home) revealed that the mRNA expression levels of *RAB38* and *MITF* were positively correlated.

We also extended the bioinformatic analyses by examining *RAB38*/*MITF* expressions in different patient cohorts including the TCGA-SCKM cohort and two other patient cohorts (GSE7553 and GSE8401) retrieved from the GEO database. We again observed a clear positive correlation between *MITF* and *RAB38* in both primary and metastatic melanoma tissues (Figure 2.11). Notably, *RAB38* expression was significantly up-regulated in the metastatic melanoma tissues relative to primary melanoma tissues from patients displaying *MITF*-high signature, namely for the patient population stratified with higher levels of *MITF* expression (Figure 2.11); however, an opposite trend was observed for patients exhibiting *MITF*-low signature, suggesting that the upregulation of RAB38 in metastatic melanoma is likely driven by MITF. Moreover, pathway analysis of the *RAB38*

gene co-expression signature in the TCGA-SKCM data showed that *MITF* is highly enriched and functionally involved with *RAB38* (Figure 2.11).

We also extended the analyses of the TCGA data to two known driver genes for melanoma metastasis, i.e. *RAB27A* and *TBC1D16*, both of which are enriched in the *RAB38* gene co-expression signature and regulated by MITF (Figure 2.12). Similar to *RAB38*, in two additional patient cohorts (GSE7553 and GSE8401), we observed significantly higher levels of expression of *RAB27A* and *TBC1D16* in the metastatic melanoma tissue samples carrying *MITF*-high signature (i.e. with high levels of MITF expression), but not in those with *MITF*-low signature (Figure 2.12).

To further substantiate the direct regulation of RAB38 by MITF, we performed chromatin immunoprecipitation (ChIP) followed by quantitative PCR (ChIP-qPCR) analysis to assess the occupancy of MITF protein in the promoter regions of *RAB38* and *TBC1D16* genes. In this respect, the 47-kDa isoform of TBC1D16 was observed to be regulated by MITF through binding to its remote promoter region.[31-32] Indeed, our ChIP-qPCR results revealed higher levels of enrichment of MITF to the promoter elements of both *RAB38* and *TBC1D16* genes in WM-266-4 cells relative to WM-115 cells, suggesting that RAB38 is directly regulated by MITF in the metastatic WM-266-4 melanoma cells (Figure 2.12).

Having assessed the MITF-mediated transcriptional regulation of RAB38, we next asked whether the expression of RAB38 in these melanoma cells are epigenetically modulated. Analyses of the previously published methylation microarray data for IGR39/IGR37 (accession number: GSE46522) and WM-115/WM-266-4 (accession

number: GSE70621) cells showed diminished levels of cytosine methylation at several CpG sites in the promoter region of *RAB38* gene in the metastatic over primary melanoma cells (Figure 2.13). These results suggest that promoter hypomethylation and the ensuing epigenetic reactivation may elicit increased levels of *RAB38* expression in the two metastatic lines (i.e. IGR37 and WM-266-4). To further substantiate this finding, we assessed the methylation status at 10 CpG sites in the promoter region of *RAB38* gene in the three pairs of primary/metastatic melanoma cell lines by employing bisulfite sequencing. Strikingly, our results revealed that these 10 CpG sites were entirely unmethylated (0.0% methylation) in WM-266-4 and IGR37 cells, whereas the overall methylation levels at these sites were 98.0% and 41.0% in WM-115 and IGR39 cells, respectively (Figure 2.13). In contrast, these CpG sites are hypermethylated in both the WM793 (99.0%) and its matched metastatic melanoma line (i.e. 1205Lu, 93.0%) (Figure 2.13). These results support that epigenetic reactivation contributes to elevated expression of *RAB38* in the metastatic lines of the WM and IGR pairs of melanoma cells, whereas epigenetic silencing led to lack of detectable levels of RAB38 protein in the primary or metastatic melanoma lines of the WMLu pair. To further validate that *RAB38* expression is regulated by CpG methylation, we treated WM-115 and IGR39 cells with a DNA demethylating reagent, 5-aza-2′-deoxycytidine (5-Aza), for 96 h, and assessed the mRNA levels of *RAB38* by RT-qPCR. Indeed, our results showed that expression level of *RAB38* was significantly increased upon 5-Aza treatment (Figure 2.13).

We also assessed whether *RAB38* hypomethylation occurs in metastatic melanoma patients of a previously reported melanoma cohort (accession number: GSE44662). It

turned out that the promoter methylation of *RAB38* gene was significantly lower in metastatic than primary melanoma tissues (Figure 2.14). In addition, analysis of the TCGA-SKCM cohort revealed that the mRNA expressions of *MITF* and *RAB38* genes were inversely correlated with their promoter methylation levels, and promoter hypomethylation of the *RAB38* gene was correlated with poor patient survival (Figure 2.14).

Together, our above results furnished evidence to support a model where loss of cytosine methylation in the promoter region of *RAB38* gene leads to its epigenetic reactivation, which involves augmented binding of MITF transcription factor to the promoter region.

2.4 Discussion

Small GTPases of the Ras superfamily are master regulators of cellular trafficking. Here, we developed a novel targeted quantitative proteomic method for human small GTPase proteome with an unprecedented level of coverage. Our MRM-based targeted proteomic method enabled a powerful and high-throughput discovery of small GTPases that become aberrantly expressed during metastatic transformation of melanoma.

Our quantitative proteomic data, along with the results obtained from cell-based assays and from bioinformatic analyses of publicly available data, support the role of RAB38 in promoting melanoma metastasis. Thus, RAB38 joins other members of the small GTPase family that regulate melanoma metastasis, including RAB27A,[5] RND3,[33] and ARF6.[34]

RAB38 displays a unique tissue-specific expression pattern, with the highest levels being observed in the lung and skin.[35] Together with RAB27A and RAB32, RAB38 has a well-established function in regulating the melanosome biogenesis and maturation.[36] It was

also found to be important for pigmentation in chocolate mice by regulating the trafficking of tyrosinase-related protein 1 (TYRP1).[37] Furthermore, previous studies unveiled the role of RAB38 in mesenchymal subtypes and malignant progression of glioma, where elevated expression of RAB38 confers poor prognosis in glioma patients.[38] However, no reports have yet elucidated the mechanistic relationship between RAB38 and the invasive properties of any type of tumor. Here, we unveiled a previously unrecognized role of RAB38 in regulating melanoma metastasis. Furthermore, our results support that RAB38 promotes melanoma progression by regulating the secretion and activities of MMP2 and MMP9, which are essential for metastatic transformation of tumor cells.

To the best our knowledge, this is the first report to link RAB38 with matrix metalloproteinase pathways. Several small GTPases were previously reported to be involved in the regulation of the MMP pathways through their roles in trafficking. For instance, RAB37 was previously identified as a metastasis suppressor in lung adenocarcinoma by influencing the metallopeptidase inhibitor 1 (TIMP1)-MMP9 pathway.[39] In particular, RAB37 was found to suppress metastasis through regulating the exocytotic trafficking of TIMP1, thereby inactivating MMP9 signaling and suppressing invasion. Moreover, RAB2A and RAB27B were shown to promote breast cancer invasion by stimulating endocytic trafficking of membrane type 1 (MT1)-MMP and MMP2, respectively.[40-41] Thus, we reason that RAB38 may play a novel role in the endocytic or exocytotic trafficking of MMP enzymes and/or their regulators, and future studies are warranted for illustrating the exact mechanisms through which RAB38 regulates MMPs.

We also explored the potential upstream mechanisms of RAB38 regulation. In this connection, our results revealed a strong correlation between the mRNA expressions of *RAB38* and *MITF*. Furthermore, we observed a complete loss of cytosine methylation, which is accompanied with elevated enrichment of MITF transcription factor, in the promoter of *RAB38* gene in WM-266-4 cells relative to WM-115 cells, supporting that epigenetic reactivation contributes to the elevated expression of *RAB38* gene in metastatic melanoma cells. The complete loss of promoter methylation was also observed for the metastatic IGR37 cells, but not for the metastatic 1205Lu cells or the paired WM793 primary melanoma cells. These findings are consistent with the relative levels of RAB38 proteins in the three paired melanoma cell lines. In this context, it is worth noting that Mueller *et al*.[42] observed, from Western blot analysis, higher levels of RAB38 in 2 human melanocyte samples than 3 primary and 3 metastatic melanoma tissues, though these cells and tissues were not derived from the same patients. These observations are in line with the notion that melanoma is a highly heterogeneous type of cancer.[43] Further studies are therefore needed to reveal the mechanisms underlying the metastatic transformation for the WM793/1205Lu paired cell lines. Nevertheless, the interrogation of TCGA and other patient cohort data uncovered a significant correlation between the elevated mRNA expression level of *RAB38* gene, or its promoter hypomethylation, and poor prognosis in melanoma patients. Moreover, a strong correlation between the expression levels of *MITF* and *RAB38* genes was observed in a large number of melanoma cell lines and tumor tissues. Thus, this epigenetic and transcriptional mechanism might be at play for a substantial subset of melanoma patients.

Apart from RAB38, our targeted proteomic approach led to the discovery of other small GTPases that may function in melanoma metastasis. For instance, the consistent down-regulation of RAB12 in the WM-266-4 and 1205Lu metastatic melanoma cells relative to their primary melanoma counterparts suggests that this protein may serve as a suppressor for melanoma metastasis. RAB12 was found to regulate the constitutive degradation of transferrin receptor,[44] and elevated levels of transferrin receptors were previously observed in melanoma cells metastasized to brain.[45] Thus, RAB12 may suppress melanoma metastasis through elevated accumulation of transferrin receptors. Furthermore, we found that RAB31 was consistently down-regulated in all three metastatic melanoma lines relative to the corresponding primary lines. Grismayer *et al*.[46] demonstrated that the increased levels of RAB31 led to a switch of invasive to proliferative phenotype in breast cancer cells. It will be important to explore, in the future, the role of RAB31 in the metastatic transformation of other types of cancer, including melanoma.

Small GTPases, like other types of GTP-binding proteins, can shuffle between the GTP-bound active states and the GDP-bound inactive states, which are regulated by GEFs, GAPs and GDIs.[2] The conformational alterations of small GTPases between these two states can modulate their binding towards different downstream effector proteins.[2] In this vein, a limitation of our targeted proteomic approach is its inability in profiling the activities of small GTPases. This limitation can be overcome by further multiplexing the assay with affinity-based techniques. As discussed above, by combining affinity enrichment with gel-based fractionation, Zhang *et al*.[17] developed an MRM-based assay to profile the activities of 12 small GTPases; however, the throughput of this assay was

relatively low. In addition, proteome-wide enrichment of active small GTPases using binding domains of their effector proteins is very challenging due to the tremendous structural diversity of effectors and the lack of knowledge about the effectors for some small GTPases. On the other hand, enrichment of small GTPases and other GTP-binding proteins with the use of acyl nucleotide affinity probes,[47] together with LC-MRM analysis, may constitute an alternative approach for high-throughput profiling of activities of small GTPases. Such an approach is currently being explored in our laboratory.

In conclusion, we developed successfully a novel MRM-based targeted quantitative proteomic method for the comprehensive profiling of small GTPases. By using this method, we assessed the differential expression of small GTPases in paired primary/metastatic melanoma cell lines. The method, when combined with bioinformatic analysis of publicly available data and cell-based assays, constitutes an integrated and effective approach to discover small GTPase that serve as drivers or suppressors for melanoma metastasis. We found that RAB38 promotes melanoma metastasis *in vitro* through the regulation of matrix metalloproteinases, and the increased expression of RAB38 in metastatic melanoma cells arises from diminished promoter methylation and heightened binding of the MITF transcription factor. It can be envisaged that the targeted proteomic method can also be employed for studying small GTPase signaling (e.g. for discovering small GTPase substrates for GEFs and GAPs) and for investigating the implications of small GTPases in other aspects of cancer biology or cancer therapy (e.g. in therapeutic resistance).

Our finding that epigenetic reactivation of *RAB38* gene stimulates melanoma metastasis suggests that the expression level of RAB38, in conjunction with the expression level of MITF, may serve as a biomarker for the prognosis of melanoma patients. In addition, targeting epigenetic modulation of RAB38 and/or its interactions with other proteins may serve as the basis for the therapeutic interventions of metastatic melanoma. In the latter respect, small-molecule inhibitors were previously reported for suppressing the interactions between small GTPases and their effector or GEF proteins.[48-49]

1. Takai, Y.; Sasaki, T.; Matozaki, T., Small GTP-binding proteins. *Physiol Rev* **2001,** *81*, 153-208.
2. Bos, J. L.; Rehmann, H.; Wittinghofer, A., GEFs and GAPs: critical elements in the control of small G proteins. *Cell* **2007,** *129*, 865-77.
3. Stenmark, H., Rab GTPases as coordinators of vesicle traffic. *Nat Rev Mol Cell Bio* **2009,** *10*, 513-525.
4. Clark, E. A.; Golub, T. R.; Lander, E. S.; Hynes, R. O., Genomic analysis of metastasis reveals an essential role for RhoC. *Nature* **2000,** *406*, 532-5.
5. Peinado, H.; Aleckovic, M.; Lavotshkin, S.; Matei, I.; Costa-Silva, B.; Moreno-Bueno, G.; Hergueta-Redondo, M.; Williams, C.; Garcia-Santos, G.; Ghajar, C.; Nitadori-Hoshino, A.; Hoffman, C.; Badal, K.; Garcia, B. A.; Callahan, M. K.; Yuan, J.; Martins, V. R.; Skog, J.; Kaplan, R. N.; Brady, M. S.; Wolchok, J. D.; Chapman, P. B.; Kang, Y.; Bromberg, J.; Lyden, D., Melanoma exosomes educate bone marrow progenitor cells toward a pro-metastatic phenotype through MET. *Nat Med* **2012,** *18*, 883-91.
6. Lange, V.; Picotti, P.; Domon, B.; Aebersold, R., Selected reaction monitoring for quantitative proteomics: a tutorial. *Mol Syst Biol* **2008,** *4*.
7. Picotti, P.; Aebersold, R., Selected reaction monitoring-based proteomics: workflows, potential, pitfalls and future directions. *Nat Methods* **2012,** *9*, 555-66.
8. Society, A. C. *Cancer Facts and Figures-2018*; American Cancer Society: Atlanta, GA, 2018; pp 1-52.
9. Lo, J. A.; Fisher, D. E., The melanoma revolution: from UV carcinogenesis to a new era in therapeutics. *Science* **2014,** *346*, 945-9.
10. Huang, D. W.; Sherman, B. T.; Lempicki, R. A., Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **2009,** *4*, 44-57.
11. Cerami, E.; Gao, J. J.; Dogrusoz, U.; Gross, B. E.; Sumer, S. O.; Aksoy, B. A.; Jacobsen, A.; Byrne, C. J.; Heuer, M. L.; Larsson, E.; Antipin, Y.; Reva, B.; Goldberg, A. P.; Sander, C.; Schultz, N., The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. *Cancer Discovery* **2012,** *2*, 401-404.
12. Barretina, J.; Caponigro, G.; Stransky, N.; Venkatesan, K.; Margolin, A. A.; Kim, S.; Wilson, C. J.; Lehar, J.; Kryukov, G. V.; Sonkin, D.; Reddy, A.; Liu, M.; Murray, L.; Berger, M. F.; Monahan, J. E.; Morais, P.; Meltzer, J.; Korejwa, A.; Jane-Valbuena, J.; Mapa, F. A.; Thibault, J.; Bric-Furlong, E.; Raman, P.; Shipway, A.; Engels, I. H.; Cheng, J.; Yu, G. K.; Yu, J.; Aspesi, P., Jr.; de Silva, M.; Jagtap, K.; Jones, M. D.; Wang, L.; Hatton, C.; Palescandolo, E.; Gupta, S.; Mahan, S.; Sougnez, C.; Onofrio, R. C.; Liefeld, T.; MacConaill, L.; Winckler, W.; Reich, M.; Li, N.; Mesirov, J. P.; Gabriel, S. B.; Getz, G.; Ardlie, K.; Chan, V.; Myer, V. E.; Weber, B. L.; Porter, J.; Warmuth, M.; Finan, P.; Harris, J. L.; Meyerson, M.; Golub, T. R.; Morrissey, M. P.; Sellers, W. R.; Schlegel, R.; Garraway, L. A., The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **2012,** *483*, 603-7.
13. Yocum, A. K.; Chinnaiyan, A. M., Current affairs in quantitative targeted proteomics: multiple reaction monitoring-mass spectrometry. *Brief Funct Genomic Proteomic* **2009,** *8*, 145-57.

14. Escher, C.; Reiter, L.; MacLean, B.; Ossola, R.; Herzog, F.; Chilton, J.; MacCoss, M. J.; Rinner, O., Using iRT, a normalized retention time for more targeted measurement of peptides. *Proteomics* **2012,** *12*, 1111-1121.
15. Kawahara, R.; Bollinger, J. G.; Rivera, C.; Ribeiro, A. C.; Brandao, T. B.; Paes Leme, A. F.; MacCoss, M. J., A targeted proteomic strategy for the measurement of oral cancer candidate biomarkers in human saliva. *Proteomics* **2016,** *16*, 159-73.
16. Halvey, P. J.; Ferrone, C. R.; Liebler, D. C., GeLC-MRM quantitation of mutant KRAS oncoprotein in complex biological samples. *J Proteome Res* **2012,** *11*, 3908-13.
17. Zhang, C. C.; Li, R.; Jiang, H.; Lin, S.; Rogalski, J. C.; Liu, K.; Kast, J., Development and application of a quantitative multiplexed small GTPase activity assay using targeted proteomics. *J Proteome Res* **2015,** *14*, 967-76.
18. MacLean, B.; Tomazela, D. M.; Shulman, N.; Chambers, M.; Finney, G. L.; Frewen, B.; Kern, R.; Tabb, D. L.; Liebler, D. C.; MacCoss, M. J., Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* **2010,** *26*, 966-8.
19. Liebler, D. C.; Zimmerman, L. J., Targeted quantitation of proteins by mass spectrometry. *Biochemistry* **2013,** *52*, 3797-806.
20. Balaban, G.; Herlyn, M.; Guerry, D. t.; Bartolo, R.; Koprowski, H.; Clark, W. H.; Nowell, P. C., Cytogenetics of human malignant melanoma and premalignant lesions. *Cancer Genet Cytogenet* **1984,** *11*, 429-39.
21. Aubert, C.; Rouge, F.; Galindo, J. R., Tumorigenicity of Human-Malignant Melanocytes in Nude-Mice in Relation to Their Differentiation Invitro. *J Natl Cancer Inst* **1980,** *64*, 1029-1040.
22. Herlyn, D.; Iliopoulos, D.; Jensen, P. J.; Parmiter, A.; Baird, J.; Hotta, H.; Adachi, K.; Ross, A. H.; Jambrosic, J.; Koprowski, H.; et al., In vitro properties of human melanoma cells metastatic in nude mice. *Cancer Res* **1990,** *50*, 2296-302.
23. Ong, S. E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D. B.; Steen, H.; Pandey, A.; Mann, M., Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* **2002,** *1*, 376-86.
24. Peinado, H.; Aleckovic, M.; Lavotshkin, S.; Matei, I.; Costa-Silva, B.; Moreno-Bueno, G.; Hergueta-Redondo, M.; Williams, C.; Garcia-Santos, G.; Ghajar, C. M.; Nitadori-Hoshino, A.; Hoffman, C.; Badal, K.; Garcia, B. A.; Callahan, M. K.; Yuan, J. D.; Martins, V. R.; Skog, J.; Kaplan, R. N.; Brady, M. S.; Wolchok, J. D.; Chapman, P. B.; Kang, Y. B.; Bromberg, J.; Lyden, D., Melanoma exosomes educate bone marrow progenitor cells toward a pro-metastatic phenotype through MET. *Nature Medicine* **2012,** *18*, 883-+.
25. Cancer Genome Atlas Research, N.; Weinstein, J. N.; Collisson, E. A.; Mills, G. B.; Shaw, K. R.; Ozenberger, B. A.; Ellrott, K.; Shmulevich, I.; Sander, C.; Stuart, J. M., The Cancer Genome Atlas Pan-Cancer analysis project. *Nat Genet* **2013,** *45*, 1113-20.
26. Xu, L.; Shen, S. S.; Hoshida, Y.; Subramanian, A.; Ross, K.; Brunet, J. P.; Wagner, S. N.; Ramaswamy, S.; Mesirov, J. P.; Hynes, R. O., Gene expression changes in an animal melanoma model correlate with aggressiveness of human melanoma metastases. *Mol Cancer Res* **2008,** *6*, 760-769.

27. Roomi, M. W.; Monterrey, J. C.; Kalinovsky, T.; Rath, M.; Niedzwiecki, A., Patterns of MMP-2 and MMP-9 expression in human cancer cell lines. *Oncol Rep* **2009,** *21*, 1323-1333.
28. Wouters, J.; Vizoso, M.; Martinez-Cardus, A.; Carmona, F. J.; Govaere, O.; Laguna, T.; Joseph, J.; Dynoodt, P.; Aura, C.; Foth, M.; Cloots, R.; van den Hurk, K.; Balint, B.; Murphy, I. G.; McDermott, E. W.; Sheahan, K.; Jirstrom, K.; Nodin, B.; Mallya-Udupi, G.; van den Oord, J. J.; Gallagher, W. M.; Esteller, M., Comprehensive DNA methylation study identifies novel progression-related and prognostic markers for cutaneous melanoma. *BMC Med* **2017,** *15*, 101.
29. Levy, C.; Khaled, M.; Fisher, D. E., MITF: master regulator of melanocyte development and melanoma oncogene. *Trends Mol Med* **2006,** *12*, 406-14.
30. Strub, T.; Giuliano, S.; Ye, T.; Bonet, C.; Keime, C.; Kobi, D.; Le Gras, S.; Cormont, M.; Ballotti, R.; Bertolotto, C.; Davidson, I., Essential role of microphthalmia transcription factor for DNA replication, mitosis and genomic stability in melanoma. *Oncogene* **2011,** *30*, 2319-32.
31. Gade, P.; Kalvakolanu, D. V., Chromatin Immunoprecipitation Assay as a Tool for Analyzing Transcription Factor Activity. *Methods Mol Biol* **2012,** *809*, 85-104.
32. Vizoso, M.; Ferreira, H. J.; Lopez-Serra, P.; Carmona, F. J.; Martinez-Cardus, A.; Girotti, M. R.; Villanueva, A.; Guil, S.; Moutinho, C.; Liz, J.; Portela, A.; Heyn, H.; Moran, S.; Vidal, A.; Martinez-Iniesta, M.; Manzano, J. L.; Fernandez-Figueras, M. T.; Elez, E.; Munoz-Couselo, E.; Botella-Estrada, R.; Berrocal, A.; Ponten, F.; Oord, J.; Gallagher, W. M.; Frederick, D. T.; Flaherty, K. T.; McDermott, U.; Lorigan, P.; Marais, R.; Esteller, M., Epigenetic activation of a cryptic TBC1D16 transcript enhances melanoma progression by targeting EGFR. *Nat Med* **2015,** *21*, 741-50.
33. Klein, R. M.; Higgins, P. J., A switch in RND3-RHOA signaling is critical for melanoma cell invasion following mutant-BRAF inhibition. *Mol Cancer* **2011,** *10*.
34. Grossmann, A. H.; Yoo, J. H.; Clancy, J.; Sorensen, L. K.; Sedgwick, A.; Tong, Z.; Ostanin, K.; Rogers, A.; Grossmann, K. F.; Tripp, S. R.; Thomas, K. R.; D'Souza-Schorey, C.; Odelberg, S. J.; Li, D. Y., The small GTPase ARF6 stimulates beta-catenin transcriptional activity during WNT5A-mediated melanoma invasion and metastasis. *Sci Signal* **2013,** *6*, ra14.
35. Osanai, K.; Oikawa, R.; Higuchi, J.; Kobayashi, M.; Tsuchihara, K.; Iguchi, M.; Jongsu, H.; Toga, H.; Voelker, D. R., A Mutation in Rab38 Small GTPase Causes Abnormal Lung Surfactant Homeostasis and Aberrant Alveolar Structure in Mice. *Am J Pathol* **2008,** *173*, 1265-1274.
36. Hume, A. N.; Collinson, L. M.; Rapak, A.; Gomes, A. Q.; Hopkins, C. R.; Seabra, M. C., Rab27a regulates the peripheral distribution of melanosomes in melanocytes. *J Cell Biol* **2001,** *152*, 795-808.
37. Loftus, S. K.; Larson, D. M.; Baxter, L. L.; Antonellis, A.; Chen, Y.; Wu, X.; Jiang, Y.; Bittner, M.; Hammer, J. A., 3rd; Pavan, W. J., Mutation of melanosome protein RAB38 in chocolate mice. *Proc Natl Acad Sci U S A* **2002,** *99*, 4471-6.
38. Wang, H. J.; Jiang, C. L., RAB38 confers a poor prognosis, associated with malignant progression and subtype preference in glioma. *Oncol Rep* **2013,** *30*, 2350-2356.

39. Tsai, C. H.; Cheng, H. C.; Wang, Y. S.; Lin, P.; Jen, J.; Kuo, I. Y.; Chang, Y. H.; Liao, P. C.; Chen, R. H.; Yuan, W. C.; Hsu, H. S.; Yang, M. H.; Hsu, M. T.; Wu, C. Y.; Wang, Y. C., Small GTPase Rab37 targets tissue inhibitor of metalloproteinase 1 for exocytosis and thus suppresses tumour metastasis. *Nat Commun* **2014,** *5*, 4804.
40. Kajiho, H.; Kajiho, Y.; Frittoli, E.; Confalonieri, S.; Bertalot, G.; Viale, G.; Di Fiore, P. P.; Oldani, A.; Garre, M.; Beznoussenko, G. V.; Palamidessi, A.; Vecchi, M.; Chavrier, P.; Perez, F.; Scita, G., RAB2A controls MT1-MMP endocytic and E-cadherin polarized Golgi trafficking to promote invasive breast cancer programs. *Embo Reports* **2016,** *17*, 1061-1080.
41. Hendrix, A.; Maynard, D.; Pauwels, P.; Braems, G.; Denys, H.; Van den Broecke, R.; Lambert, J.; Van Belle, S.; Cocquyt, V.; Gespach, C.; Bracke, M.; Seabra, M. C.; Gahl, W. A.; De Wever, O.; Westbroek, W., Effect of the secretory small GTPase Rab27B on breast cancer growth, invasion, and metastasis. *J Natl Cancer Inst* **2010,** *102*, 866-80.
42. Mueller, D. W.; Rehli, M.; Bosserhoff, A. K., miRNA expression profiling in melanocytes and melanoma cell lines reveals miRNAs associated with formation and progression of malignant melanoma. *J Invest Dermatol* **2009,** *129*, 1740-51.
43. Merlino, G.; Herlyn, M.; Fisher, D. E.; Bastian, B. C.; Flaherty, K. T.; Davies, M. A.; Wargo, J. A.; Curiel-Lewandrowski, C.; Weber, M. J.; Leachman, S. A.; Soengas, M. S.; McMahon, M.; Harbour, J. W.; Swetter, S. M.; Aplin, A. E.; Atkins, M. B.; Bosenberg, M. W.; Dummer, R.; Gershenwald, J. E.; Halpern, A. C.; Herlyn, D.; Karakousis, G. C.; Kirkwood, J. M.; Krauthammer, M.; Lo, R. S.; Long, G. V.; McArthur, G.; Ribas, A.; Schuchter, L.; Sosman, J. A.; Smalley, K. S.; Steeg, P.; Thomas, N. E.; Tsao, H.; Tueting, T.; Weeraratna, A.; Xu, G.; Lomax, R.; Martin, A.; Silverstein, S.; Turnham, T.; Ronai, Z. A., The state of melanoma: challenges and opportunities. *Pigment Cell Melanoma Res* **2016,** *29*, 404-16.
44. Matsui, T.; Itoh, T.; Fukuda, M., Small GTPase Rab12 regulates constitutive degradation of transferrin receptor. *Traffic* **2011,** *12*, 1432-43.
45. Nicolson, G. L.; Nakajima, M.; Herrmann, J. L.; Menter, D. G.; Cavanaugh, P. G.; Park, J. S.; Marchetti, D., Malignant-Melanoma Metastasis to Brain - Role of Degradative Enzymes and Responses to Paracrine Growth-Factors. *J Neurooncol* **1994,** *18*, 139-149.
46. Grismayer, B.; Solch, S.; Seubert, B.; Kirchner, T.; Schafer, S.; Baretton, G.; Schmitt, M.; Luther, T.; Kruger, A.; Kotzsch, M.; Magdolen, V., Rab31 expression levels modulate tumor-relevant characteristics of breast cancer cells. *Mol Cancer* **2012,** *11*.
47. Xiao, Y.; Ji, D.; Guo, L.; Wang, Y., Comprehensive characterization of (S)GTP-binding proteins by orthogonal quantitative (S)GTP-affinity profiling and (S)GTP/GTP competition assays. *Anal Chem* **2014,** *86*, 4550-8.
48. Gao, Y.; Dickerson, J. B.; Guo, F.; Zheng, J.; Zheng, Y., Rational design and characterization of a Rac GTPase-specific small molecule inhibitor. *Proc. Natl. Acad. Sci. USA* **2004,** *101*, 7618-23.
49. Ostrem, J. M.; Peters, U.; Sos, M. L.; Wells, J. A.; Shokat, K. M., K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions. *Nature* **2013,** *503*, 548-51.

Figure 2.1 Schematic diagram showing a targeted proteomic strategy for high-throughput quantitative profiling of small GTPases in paired primary/metastatic melanoma cells.

(A) A representative MS/MS obtained from data-dependent acquisition supporting the reliable identification of the peptide LLALGDSGVGK from RAB27B; **(B)** LC-MRM spectra for the same peptide from targeted analysis with light- and heavy-labeled lysine on the C terminus, respectively. The distribution of the peak intensities was consistent with the theoretical distribution found in the MS/MS from the MRM spectral library; **(C)** Schematic diagram showing the targeted proteomic workflow, relying on metabolic labeling with SILAC, SDS-PAGE fractionation, and LC-MRM analysis, for quantifying the differential expressions of small GTPases in WM-115 (primary) and WM-266-4 (metastatic) melanoma cells.

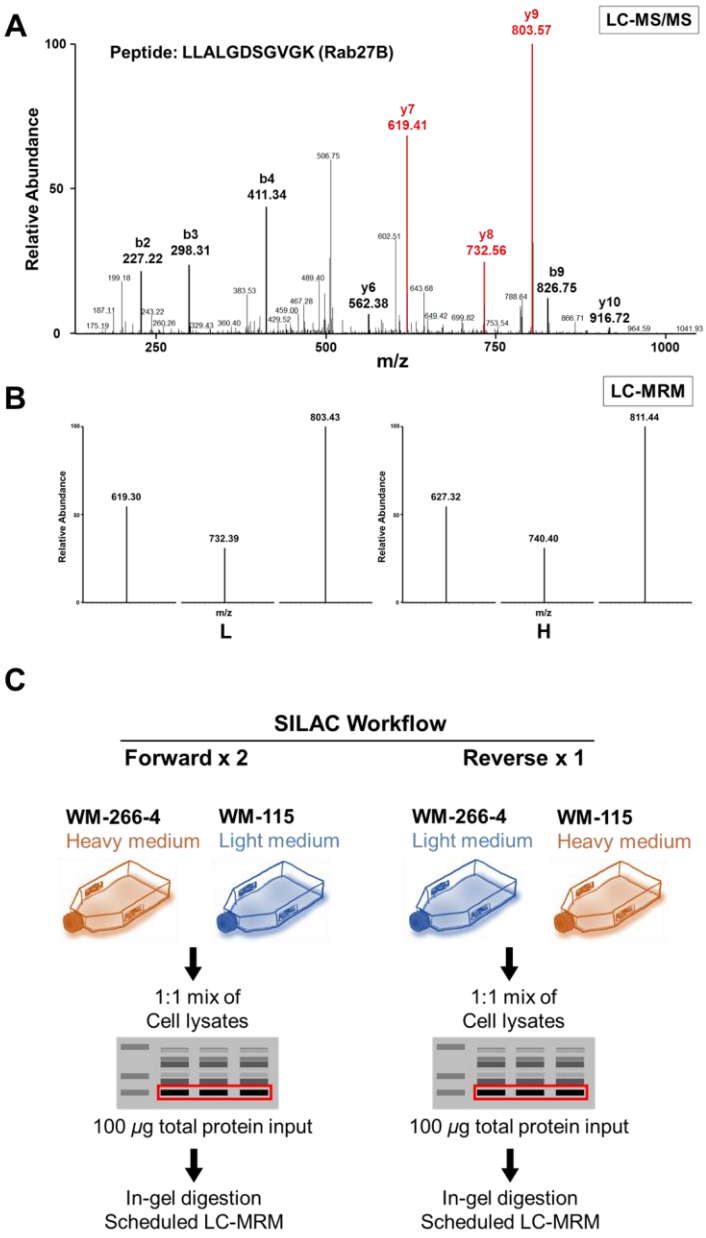

A
LC-MS/MS
Peptide: LLALGDSGVGK (Rab27B)
Relative Abundance
m/z
y9
803.57
y7
619.41
b4
411.34
b2
227.22
b3
298.31
y8
732.56
b9
826.75
y6
562.38
y10
916.72
B
LC-MRM
803.43
619.30
732.39
811.44
627.32
740.40
L
H
C
SILAC Workflow
Forward x 2
Reverse x 1
WM-266-4
Heavy medium
WM-115
Light medium
WM-266-4
Light medium
WM-115
Heavy medium
1:1 mix of
Cell lysates
1:1 mix of
Cell lysates
100 μg total protein input
100 μg total protein input
In-gel digestion
Scheduled LC-MRM
In-gel digestion
Scheduled LC-MRM

Figure 2.2 Heatmap showing the differential expression of small GTPases in paired WM-115 and WM-266-4 melanoma cells.

Shown are the Log_2R(WM-266-4/WM-115) values obtained from scheduled LC-MRM analyses of samples from two forward- and one reverse-SILAC labeling experiments. The red and blue bars designate those small GTPases that are up-and down-regulated, respectively, in the WM-266-4 metastatic melanoma cells as compared with in the WM-115 primary melanoma cells, as indicated by the scale bar.

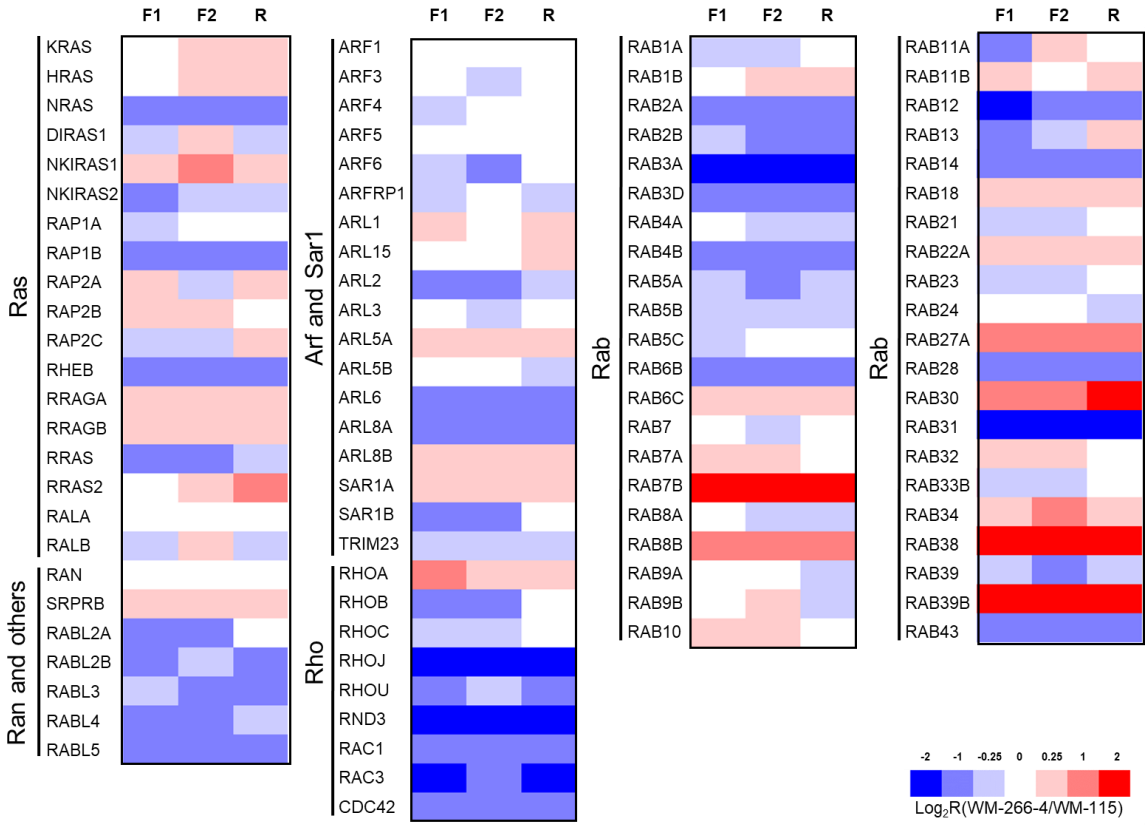


Figure 2.3 LC-MRM-based targeted quantitative proteomic assay revealed differential expression of small GTPases during melanoma metastasis.

(A−C) Bar graphs showing substantially up-regulated (>1.5-fold) and down-regulated (>1.5-fold) small GTPases quantified from LC-MRM experiments for: **(A)** WM-115/WM-266-4 paired melanoma cell lines; **(B)** IGR39/IGR37 paired melanoma cell lines; and **(C)** WM793/1205Lu paired melanoma cell lines; Error bars in (A)–(C) represent means ± standard deviation of results from three independent SILAC labeling experiments (two forward- and one reverse-SILAC).

(D) Hierarchical clustering of commonly quantified small GTPases among the WM, IGR and WMLu paired melanoma cell lines using the Z-score values for R(metastatic/primary) for individual proteins.

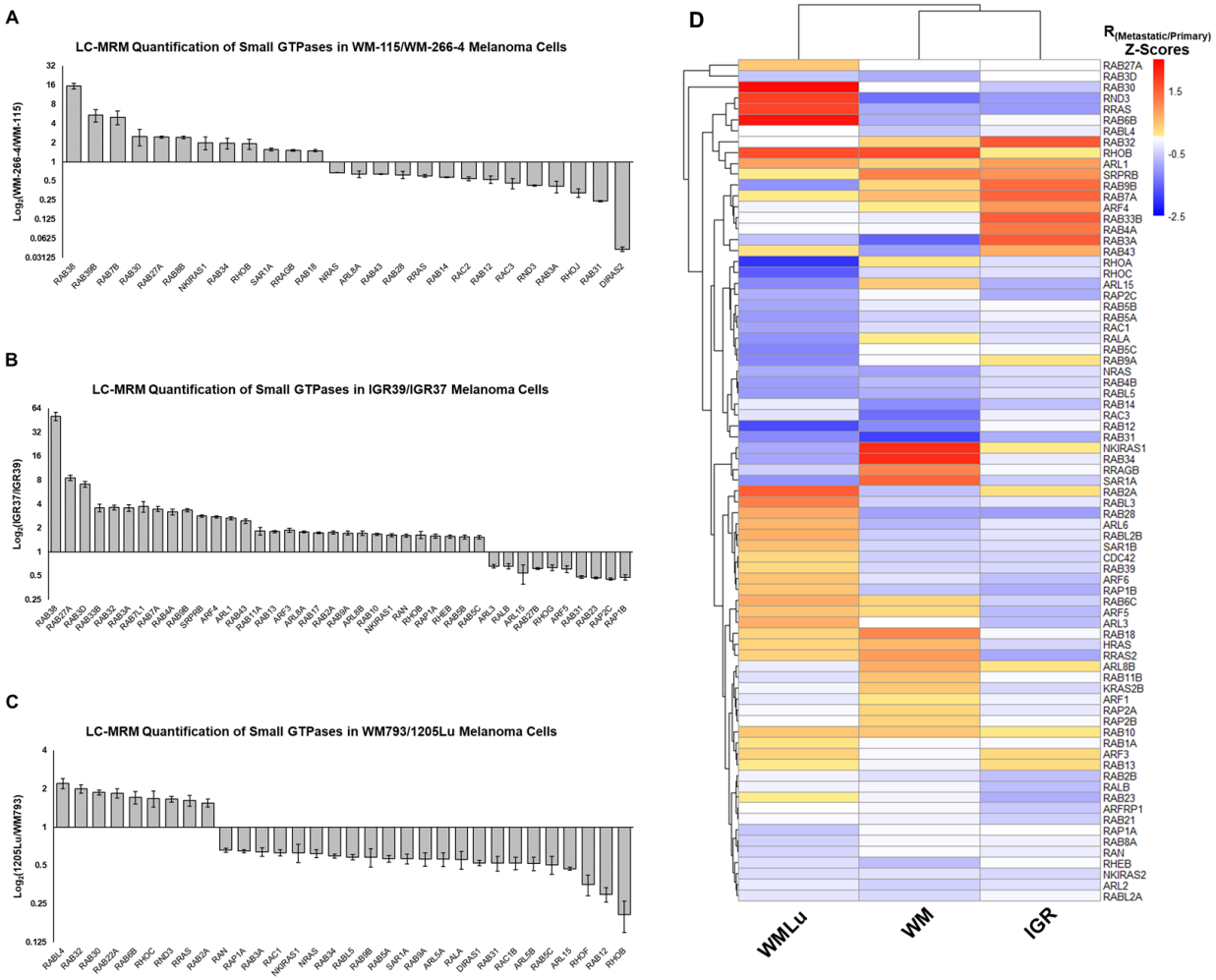


Figure 2.4 The scheduled LC-MRM platform provides accurate retention time (RT) prediction of targeted peptides, high sensitivity and excellent reproducibility.

(A) The correlation between iRT and measured RT values in LC-MRM experiments on a TSQ Vantage triple-quadruple mass spectrometer (157-min linear gradient) with a very high correlation coefficient (R^2 = 0.992); **(G)** Venn diagram showing the overlap of quantified small GTPases from WM-115 and WM-266-4 SILAC cell lysates obtained from LC-MRM analysis and shotgun proteomics experiments; **(H)** Venn diagram depicting the overlap of quantified small GTPases from WM-115 and WM-266-4 SILAC cell lysates the forward (F) and reverse (R) labeling experiments obtained from shotgun proteomics experiments.

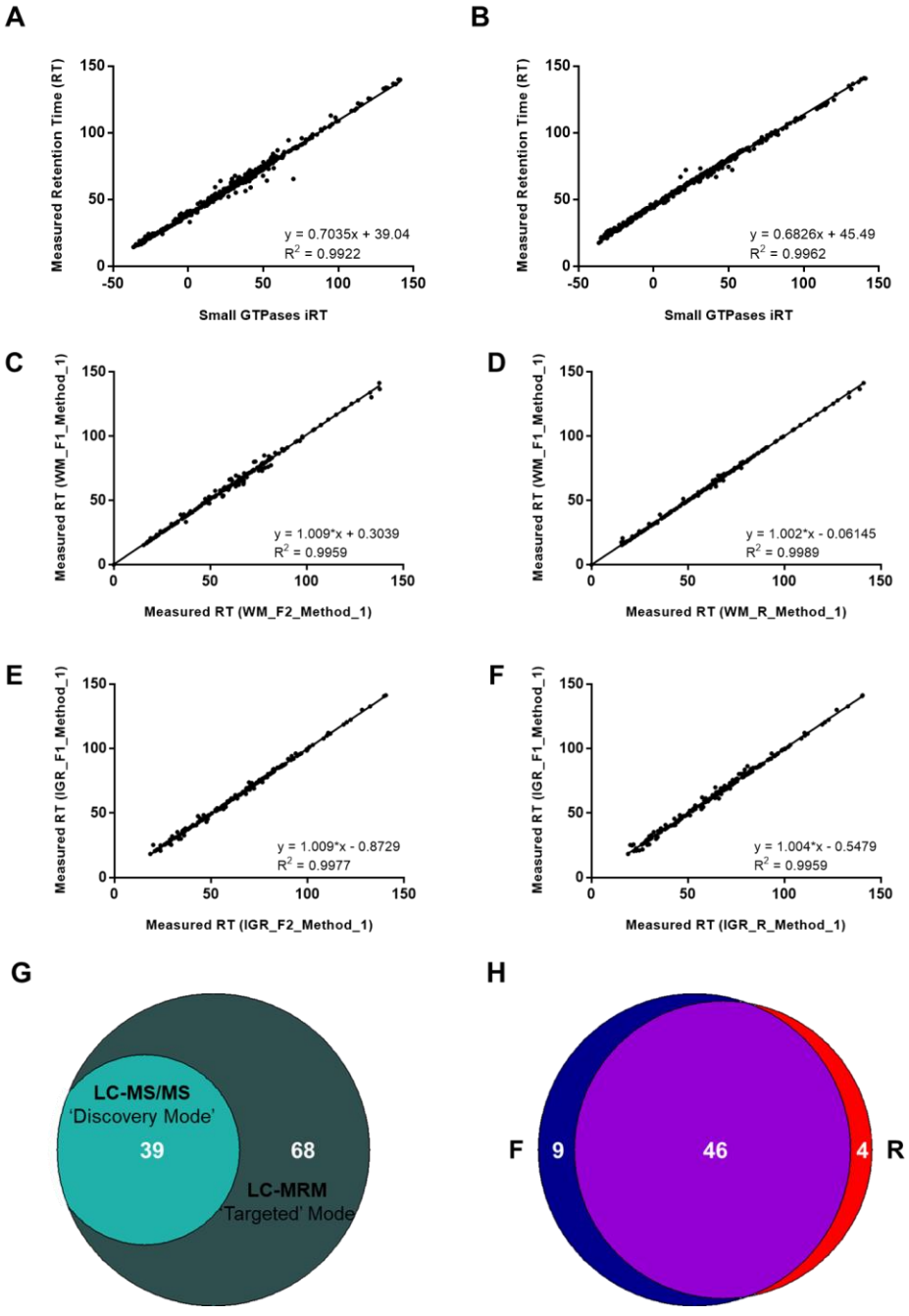

A
Measured Retention Time (RT)
y = 0.7035x + 39.04
R2 = 0.9922
Small GTPases iRT
B
Measured Retention Time (RT)
y = 0.6826x + 45.49
R2 = 0.9962
Small GTPases iRT
C
Measured RT (WM_F1_Method_1)
y = 1.009*x + 0.3039
R2 = 0.9959
Measured RT (WM_F2_Method_1)
D
Measured RT (WM_F1_Method_1)
y = 1.002*x - 0.06145
R2 = 0.9989
Measured RT (WM_R_Method_1)
E
Measured RT (IGR_F1_Method_1)
y = 1.009*x - 0.8729
R2 = 0.9977
Measured RT (IGR_F2_Method_1)
F
Measured RT (IGR_F1_Method_1)
y = 1.004*x - 0.5479
R2 = 0.9959
Measured RT (IGR_R_Method_1)
G
LC-MS/MS
'Discovery Mode'
39
68
LC-MRM
'Targeted' Mode
H
F
9
46
4
R

Figure 2.5 LC-MRM and Western blots revealed consistently higher levels of expression of RAB27A proteins in metastatic melanoma cell lines (WM-266-4 and IGR37) than the paired primary melanoma cell lines (WM-115 and ICR39).

(A) Extracted MRM traces for three transitions (y_8, y_7, and y_6) monitored for a unique tryptic peptide from RAB27A, i.e. TSVLYQYTDGK, with light (blue) and heavy (red) labels in forward- and reverse-SILAC experiments for both WM-115/WM-266-4 and IGR39/IGR37 paired melanoma cells; **(B)** Western blot analysis confirmed the elevated expression of RAB27A in WM-266-4 and IGR37 cells; **(C)** Quantification results for RAB27A from LC-MRM and Western blot analyses. The values represent the mean and standard deviation of results obtained from three independent experiments.

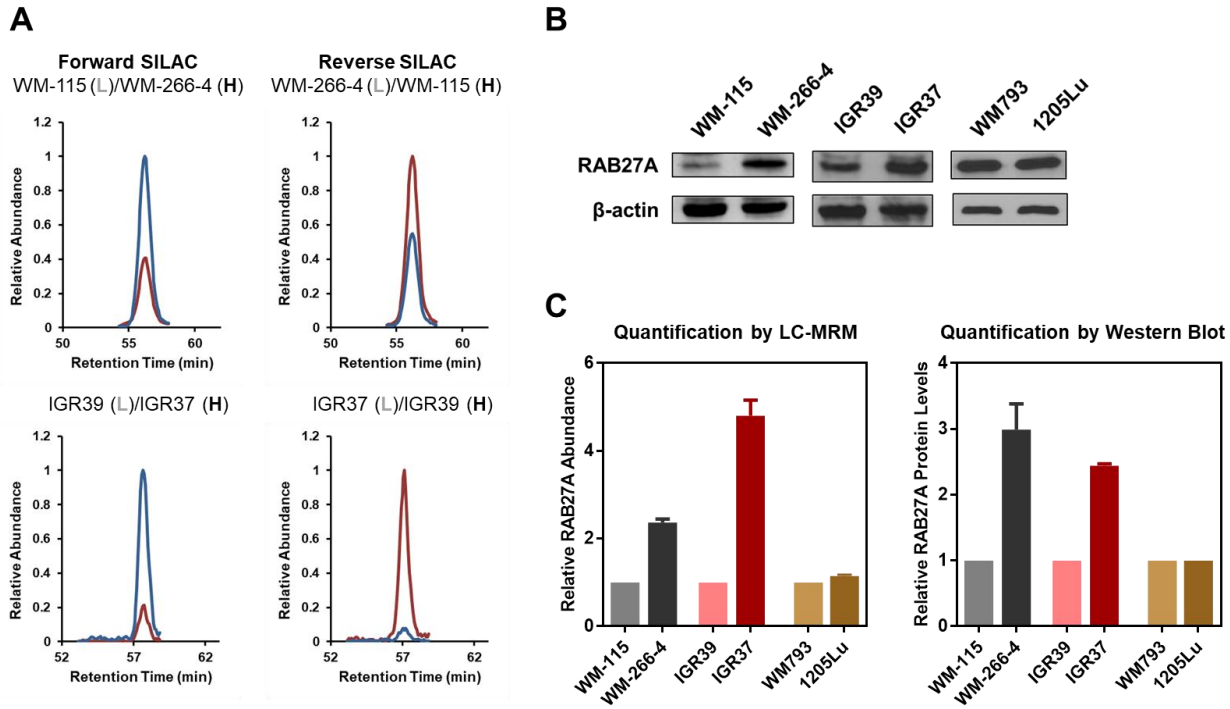


Figure 2.6 LC-MRM and Western blots revealed consistently higher levels of expression of RAB38 proteins in metastatic melanoma cell lines (WM-266-4 and IGR37) than the paired primary melanoma cell lines (WM-115 and ICR39).

(A) Extracted MRM traces for three transitions (y_9, y_8, and y_7) monitored for a unique tryptic peptide from RAB38, i.e. LLVIGDLGVGK, with light (blue) and heavy (red) labels in forward and reverse SILAC experiments for both WM-115/WM-266-4 and IGR39/IGR37 paired melanoma cells; **(B)** Western blot analysis confirmed the elevated expression of RAB38 in WM-266-4 and IGR37 cells; **(C)** Quantification results for RAB38 from LC-MRM and Western blot analyses. The values represent the mean and standard deviation of results obtained from three independent experiments.

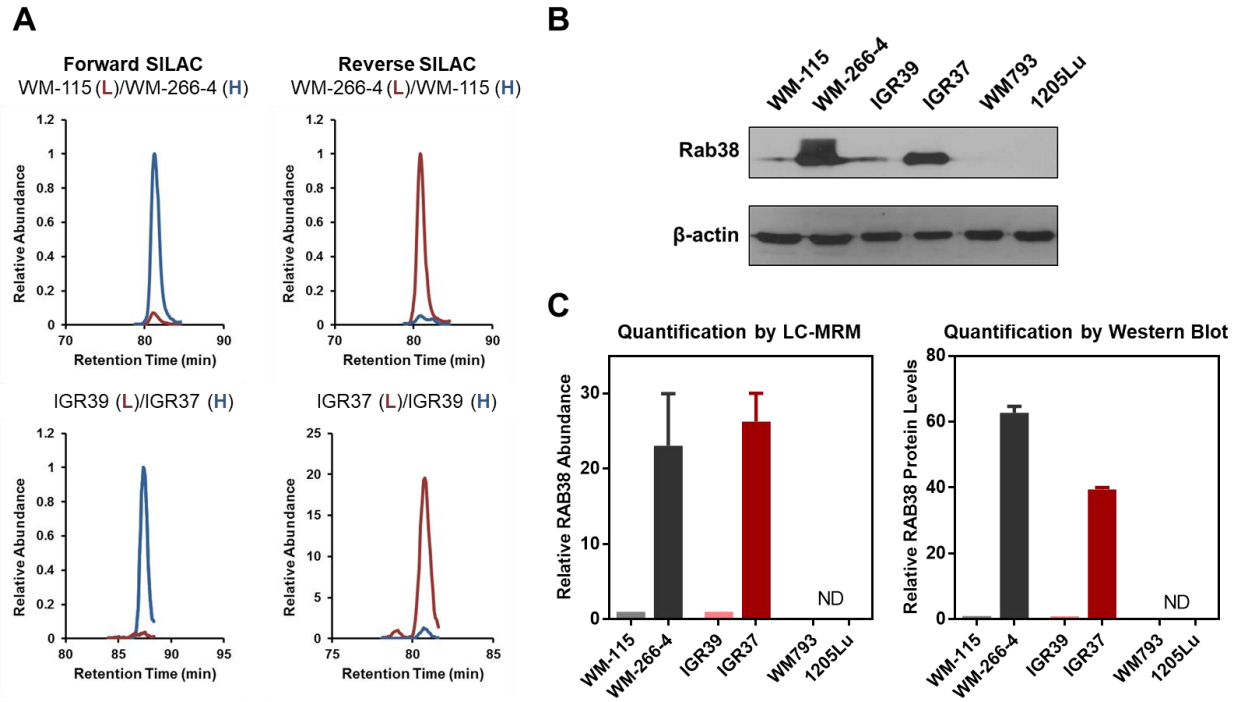


Figure 2.7 Bioinformatic analyses revealed RAB38 as a potential driver for melanoma progression.

(A) Kaplan-Meier plot of overall patient survival stratified by median *RAB38* mRNA expression in the TCGA-SKCM cohort. Log-rank test *p*-value is displayed; **(B)** Box plot showing enriched *RAB38* mRNA expressions in the TCGA-SKCM and the TCGA-UVM cohorts; **(C)** Scatter plot showing up-regulated *RAB38* mRNA expression in various metastatic melanoma cell lines (highlighted as black dots) in the NCI-60 Human Tumor Cell Lines Database; **(D)** Scatter plot showing up-regulated *RAB38* mRNA expressions in melanoma cell lines in the Cancer Cell Line Encyclopedia (CCLE) Database; **(E)** Scattered plot showing up-regulated *RAB38* mRNA expressions in various metastatic melanoma cell lines (highlighted as dark grey dots) in the CCLE Database; **(F)** *RAB38* mRNA levels were significantly up-regulated in the highly metastatic derivatives of A375 cells cell lines compared to the poorly metastatic A375 parental cells (GEO data series: GSE7929). The *p* values were calculated by using an unpaired two-tailed Student's *t* test.

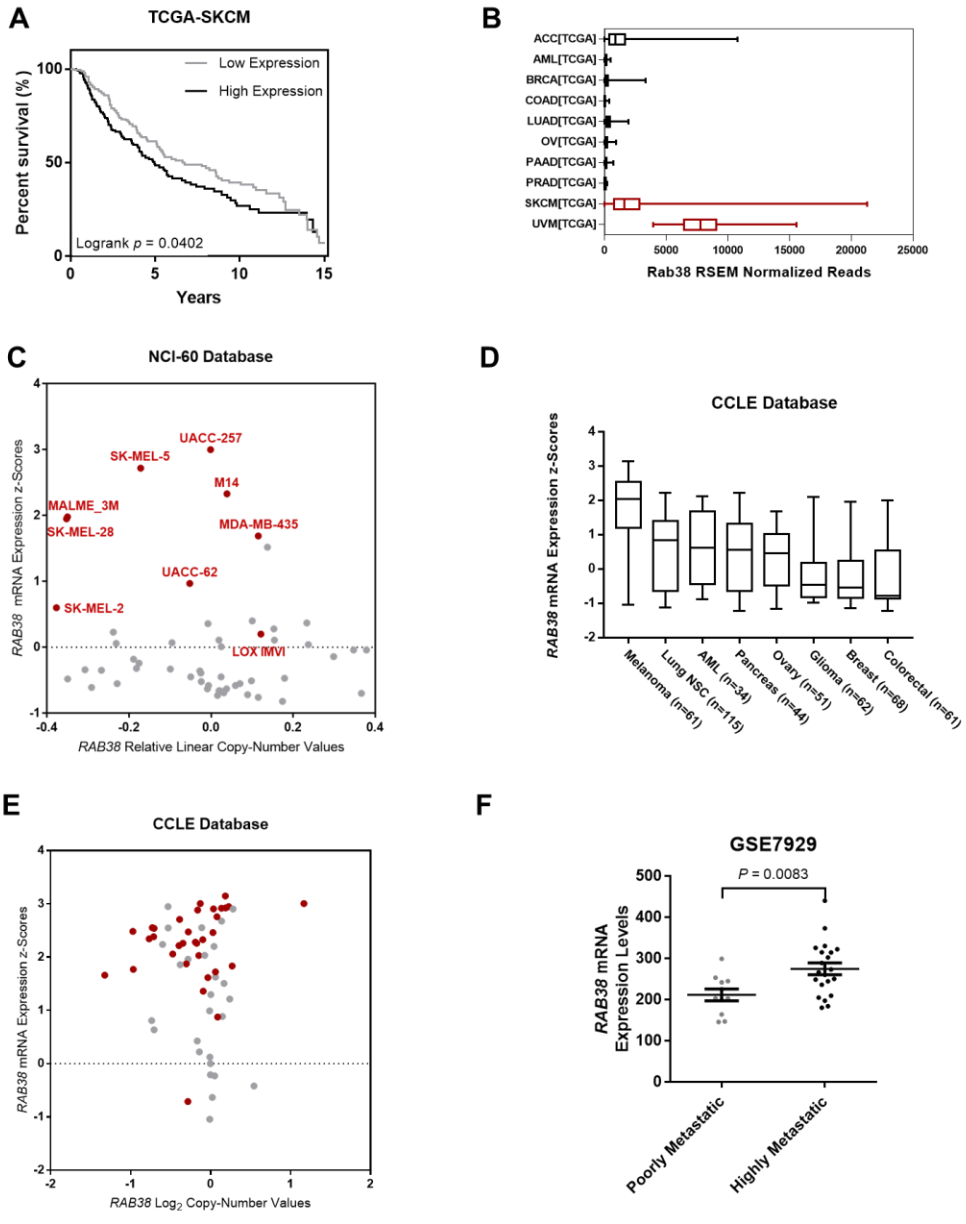

A
TCGA-SKCM
Low Expression
High Expression
Percent survival (%)
Logrank p = 0.0402
Years
B
ACC[TCGA]
AML[TCGA]
BRCA[TCGA]
COAD[TCGA]
LUAD[TCGA]
OV[TCGA]
PAAD[TCGA]
PRAD[TCGA]
SKCM[TCGA]
UVM[TCGA]
Rab38 RSEM Normalized Reads
C
NCI-60 Database
UACC-257
SK-MEL-5
M14
MALME_3M
SK-MEL-28
MDA-MB-435
UACC-62
SK-MEL-2
LOX IMVI
RAB38 mRNA Expression z-Scores
RAB38 Relative Linear Copy-Number Values
D
CCLE Database
RAB38 mRNA Expression z-Scores
Melanoma (n=61)
Lung NSC (n=115)
AML (n=34)
Pancreas (n=44)
Ovary (n=51)
Glioma (n=62)
Breast (n=68)
Colorectal (n=61)
E
CCLE Database
RAB38 mRNA Expression z-Scores
RAB38 Log2 Copy-Number Values
F
GSE7929
P = 0.0083
RAB38 mRNA Expression Levels
Poorly Metastatic
Highly Metastatic

Figure 2.8 RAB38 enhanced melanoma metastasis *in vitro*.

(A) *In vitro* migration and invasion assays showed similar migration rates but significantly increased invasion rates in FLAG-RAB38-expressing WM-115 cells as compared to empty vector control. Migration and invasion capabilities were measured by using transwell migration and Matrigel-based invasion assays, respectively; **(B)** Quantification results for *in vitro* migration and invasion assay shown in panel (A). **(C)** *In vitro* migration and invasion assays showed similar migration rates but significantly decreased invasion rates for WM-266-4 cells with siRAB38 knockdown as compared to non-targeting siRNA control; **(D)** Quantification results for *in vitro* migration and invasion assay shown in panel (C). Error bars represent mean ± standard deviation (n = 3). The *p* values were calculated by using an unpaired, two-tailed Student's *t* test.

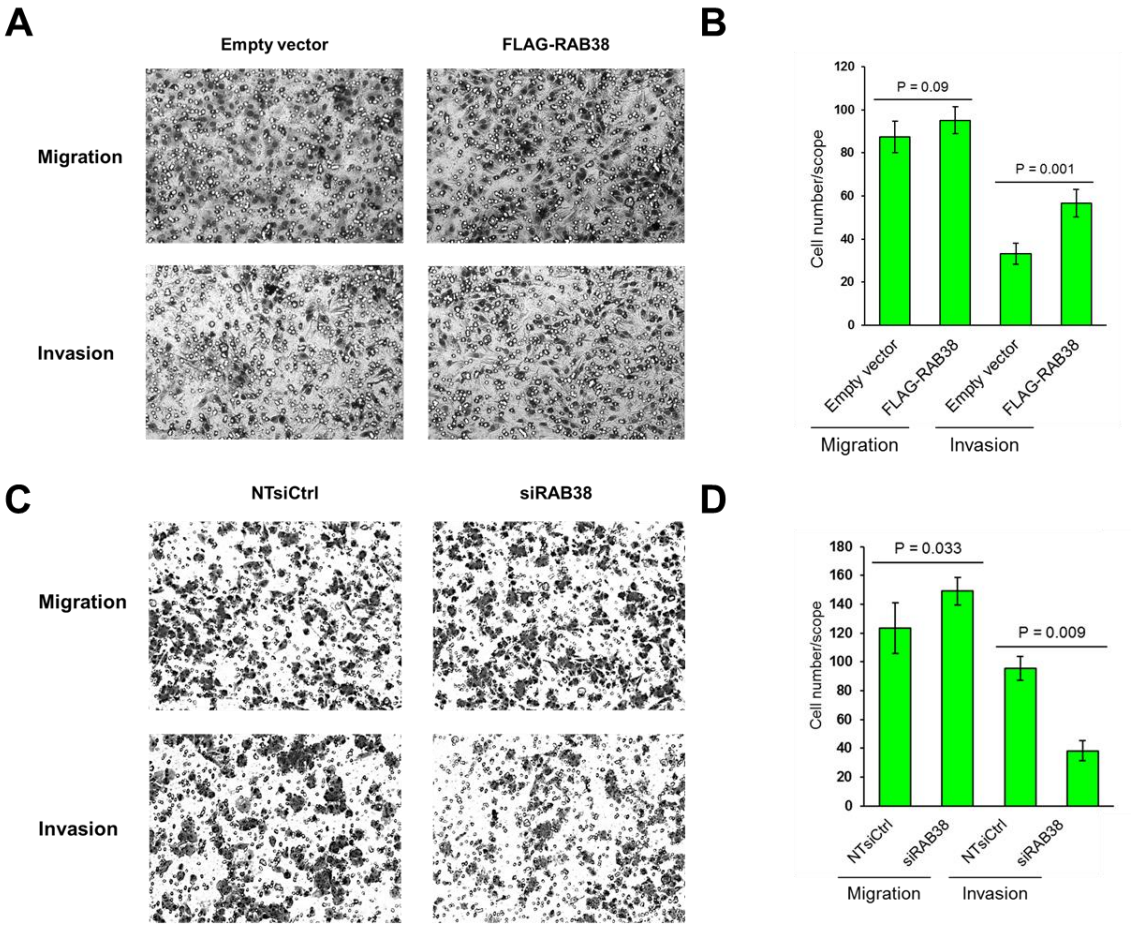


Figure 2.9 RAB38 regulates melanoma metastasis by mediating the expression levels and activities of MMP2 and MMP9.

(A) RT-qPCR assays showed decreased expression levels of MMP2 and MMP9 in WM-266-4 cells with siRAB38 knockdown as compared to non-targeting siRNA control. Error bars represent mean ± standard error of the mean (SEM) (n = 3); **(B)** Gelatin zymography assays revealed diminished enzymatic activities of MMP2 and MMP9 in WM-266-4 cells with siRAB38 knockdown as compared to non-targeting siRNA control; **(C)** Quantification results for gelatin zymography assays shown in (B). Error bars represent mean ± standard deviation (n = 3). **(D)** Real-time quantitative PCR (RT-qPCR) assays showed increased expression levels of MMP2 and MMP9 in FLAG-RAB38-expressing WM-115 cells as compared to empty vector control. Error bars represent mean ± SEM (n = 3); **(E)** Gelatin zymography assays revealed elevated enzymatic activities of MMP2 and MMP9 in FLAG-RAB38-expressing WM-115 cells as compared to empty vector control; **(F)** Quantification results for gelatin zymography assays shown in (E). Error bars represent mean ± standard deviation (n = 3). The *p* values for all figures are as follows: "∗", $0.01 < p < 0.05$; "∗∗", $0.001 < p < 0.01$; "∗∗∗", $p < 0.001$. The *p* values were calculated by using a paired, two-tailed Student's *t* test.

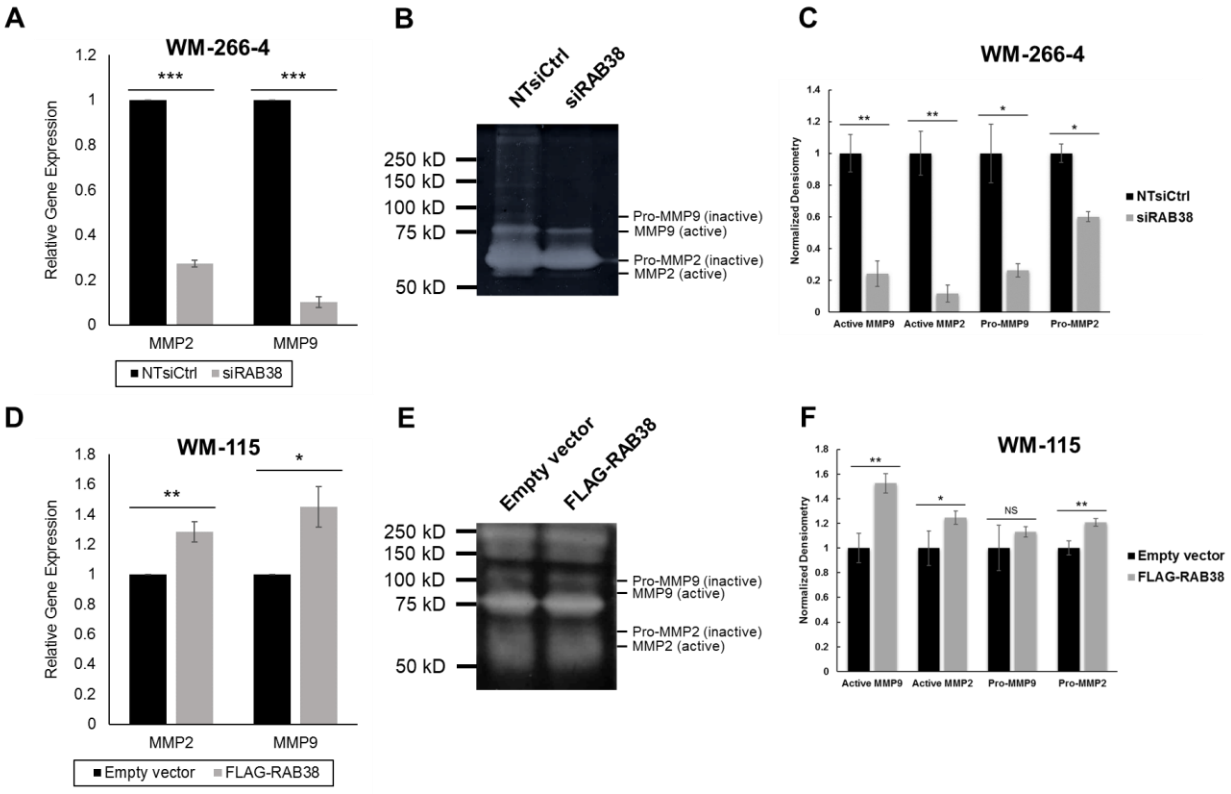


Figure 2.10 RAB38 mediates both the expression levels and activities of MMP2 and MMP9 in metastatic M14 and IGR37 cells.

(A) RT-qPCR analysis showed pronouncedly higher expression levels of *RAB38* in M14 cells as compared to WM793 and 1205Lu cells. Error bars represent mean ± SEM (n = 3); **(B)** RT-qPCR analysis showed decreased expression levels of *MMP2* and *MMP9* in M14 cells with siRAB38 knockdown as compared to non-targeting siRNA control. Error bars represent mean ± SEM (n = 3); **(C)** RT-qPCR analysis showed decreased expression level of *MMP9*, but not *MMP2*, in IGR37 cells with siRAB38 knockdown as compared to non-targeting siRNA control. Error bars represent mean ± SEM (n = 3); The *p* values in (A) and (B) were calculated by using unpaired two-tailed Student's *t* test: "∗", $0.01 < p < 0.05$; "∗∗", $0.001 < p < 0.01$; "∗∗∗", $p < 0.001$; **(D)** Gelatin zymography assays revealed diminished enzymatic activities of MMP2 and MMP9 in M14 cells with siRAB38 knockdown as compared to non-targeting siRNA control; **(E)** Gelatin zymography assays revealed diminished enzymatic activities of MMP2 and MMP9 in IGR37 cells with siRAB38 knockdown as compared to non-targeting siRNA control.

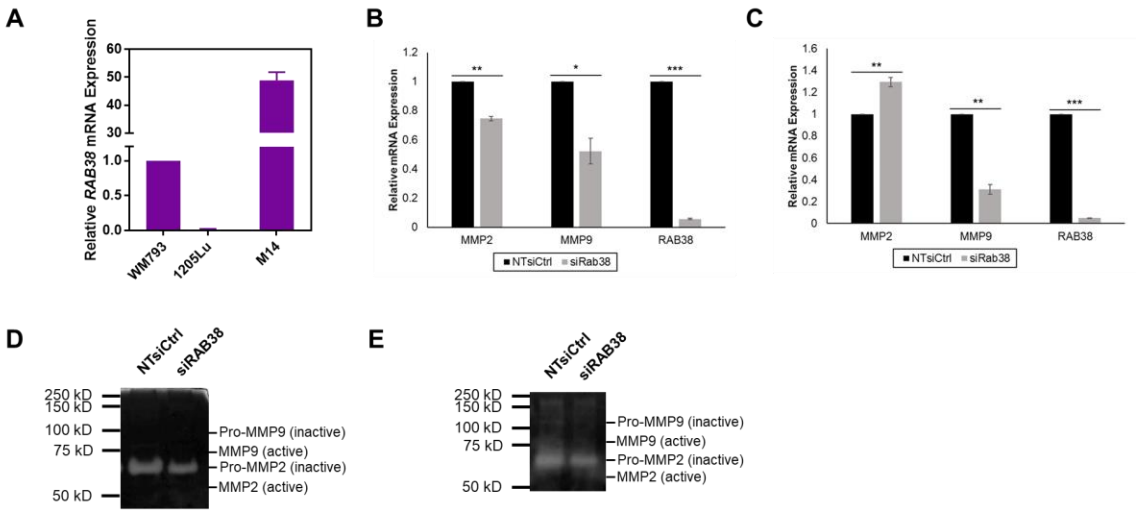


Figure 2.11 RAB38 expressions in large melanoma cell line or patient cohorts were highly correlated with the melanoma lineage-specific transcription factor MITF.

(A−D) Heatmaps showing the correlations between *RAB38* and *MITF* mRNA expressions in the: **(A)** Primary melanoma tissues in the GSE7553 cohort; **(B)** Metastatic melanoma tissues in the GSE7553 cohort; **(C)** Primary melanoma tissues in the GSE8401 cohort; **(D)** Metastatic melanoma tissues in the GSE8401 cohort;

(E−H) Box plots showing the *RAB38* mRNA expressions in the: **(E)** Metastatic patient tissues carrying *MITF*-high signature in the GSE7553 cohort; **(F)** Metastatic patient tissues carrying *MITF*-low signature in the GSE7553 cohort; **(G)** Metastatic patient tissues carrying *MITF*-high signature in the GSE8401 cohort; **(H)** Metastatic patient tissues carrying *MITF*-high signature in the GSE8401 cohort.

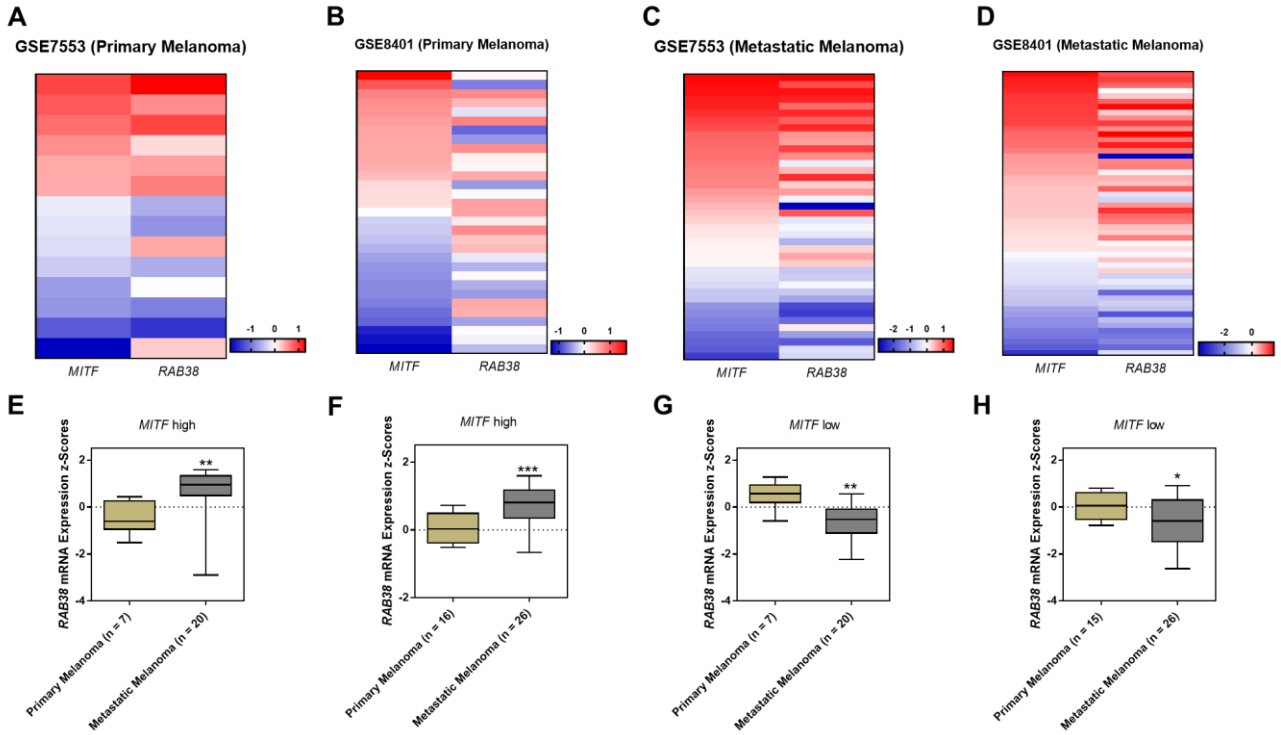
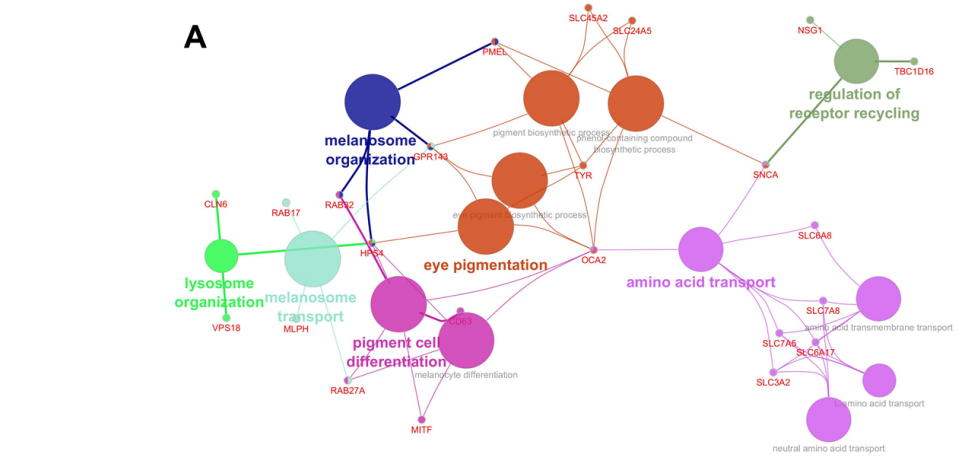
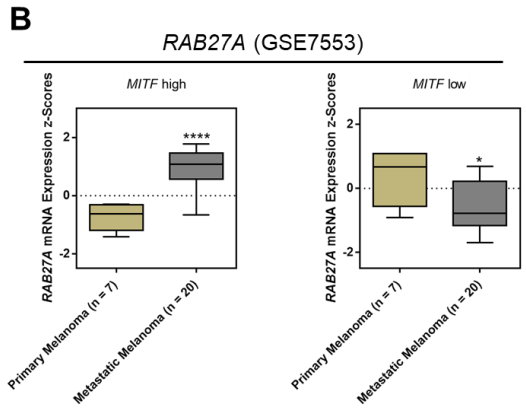
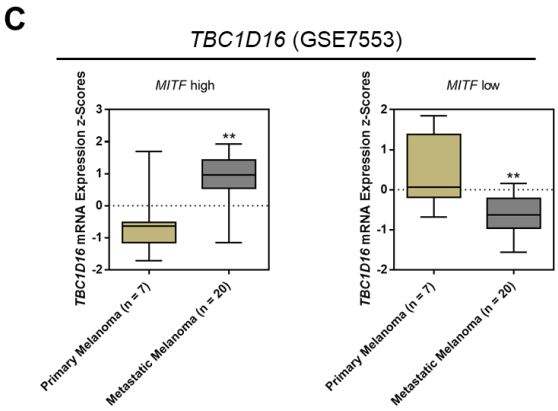


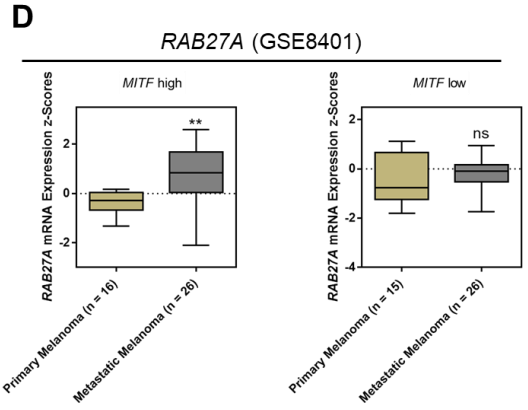
Figure 2.12 Higher levels of expression of *RAB38* in metastatic melanoma compared to primary melanoma were observed in melanoma patient with *MITF* high signature.

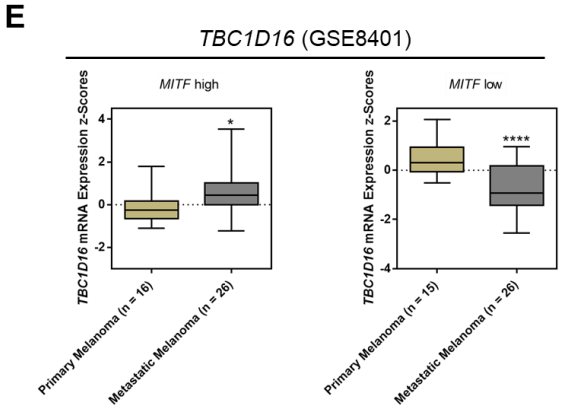
(A) Pathway analysis showed enrichment of *MITF*, *TBC1D16* and *RAB27A* in the *RAB38* co-expression signature; **(B)** Box plot showed *RAB27A* mRNA expressions in the metastatic patient tissues carrying *MITF*-high or *MITF*-low signature in the GSE7553 cohort; **(C)** Box plot showed *TBC1D16* mRNA expressions in the metastatic patient tissues carrying *MITF*-high or *MITF*-low signature in the GSE7553 cohort; **(D)** Box plot showed *RAB27A* mRNA expressions in the metastatic patient tissues carrying *MITF*-high or *MITF*-low signature in the GSE8401 cohort; **(E)** Box plot showed *TBC1D16* mRNA expressions in the metastatic patient tissues carrying *MITF*-high or *MITF*-low signature in the GSE8401 cohort; **(F)** ChIP-PCR assays showed enrichment of the association of *MITF* with the promoters of *TBC1D16* and *RAB38* genes in WM-266-4 cells relative to WM-115 cells. Data were normalized to control IgG ChIP. Error bars represent means ± standard deviations (n = 3).


A
SLC45A2
SLC24A5
PMEL
NSG1
TBC1D16
regulation of
receptor recycling
melanosome
organization
pigment biosynthetic process
phenol-containing compound
biosynthetic process
GPR143
TYR
SNCA
CLN6
RAB17
RAB32
eye pigment biosynthetic process
HPS4
eye pigmentation
OCA2
SLC6A8
amino acid transport
lysosome
organization
melanosome
transport
VPS18
MLPH
SLC7A8
amino acid transmembrane transport
SLC7A6
SLC6A17
pigment cell
differentiation
melanocyte differentiation
RAB27A
SLC3A2
amino acid transport
MITF
neutral amino acid transport


B
RAB27A (GSE7553)
MITF high
MITF low
RAB27A mRNA Expression z-Scores
****
*
Primary Melanoma (n = 7)
Metastatic Melanoma (n = 20)


C
TBC1D16 (GSE7553)
MITF high
MITF low
TBC1D16 mRNA Expression z-Scores
**
**
Primary Melanoma (n = 7)
Metastatic Melanoma (n = 20)


D
RAB27A (GSE8401)
MITF high
MITF low
RAB27A mRNA Expression z-Scores
**
ns
Primary Melanoma (n = 16)
Metastatic Melanoma (n = 26)
Primary Melanoma (n = 15)
Metastatic Melanoma (n = 26)


E
TBC1D16 (GSE8401)
MITF high
MITF low
TBC1D16 mRNA Expression z-Scores
*
****
Primary Melanoma (n = 16)
Metastatic Melanoma (n = 26)
Primary Melanoma (n = 15)
Metastatic Melanoma (n = 26)

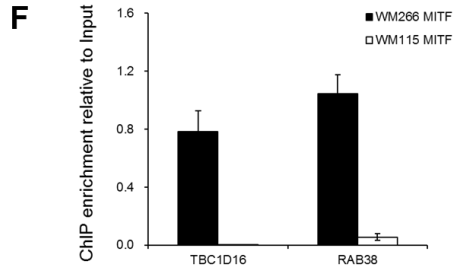

F
WM266 MITF
WM115 MITF
ChIP enrichment relative to Input
1.6
1.2
0.8
0.4
0.0
TBC1D16
RAB38

Figure 2.13 Investigation of *RAB38* methylation status in melanoma cell lines.

(A) Bisulfite sequencing demonstrated the methylation status of CpG sites in the promoter region of *RAB38* gene in the three paired primary/metastatic cell lines, where high levels of methylation were observed for the WM-115, IGR39, WM793 and 1205Lu, but not for the WM-266-4 and IGR37 melanoma cell lines. CpG sites in the promoter region of *RAB38* gene are indicated by short vertical bars, and exons are designated with black rectangles on the top. The arrow indicates the transcription start site (TSS). Each horizontal line represents one separate clone that was sequenced, and open and filled circles represent unmethylated and methylated CpG sites, respectively; **(B)** Increased mRNA expression levels of *MITF*, *TBC1D16*-47kDa and *RAB38* after 5-aza-2′-deoxycytidine (5-Aza) treatment (96 h) in WM-115 cells; **(C)** Increased mRNA expression levels of *MITF*, *TBC1D16*-47kDa and *RAB38* after 5-Aza treatment (96 h) in IGR39 cells; **(D)** Box plots representing DNA methylation in 4 primary melanoma and 33 metastatic melanoma samples (accession number: GSE44662). Metastatic melanomas contained lower *RAB38* promoter methylation. The error bars in panels (B) and (C) represent mean ± SEM. The *p* values were calculated by using an unpaired two-tailed Student's *t* test: "ns", not significant; "∗", $0.01 < p < 0.05$; "∗∗", $0.001 < p < 0.01$; "∗∗∗", $0.0001 < p < 0.001$; "∗∗∗∗", $p < 0.0001$.

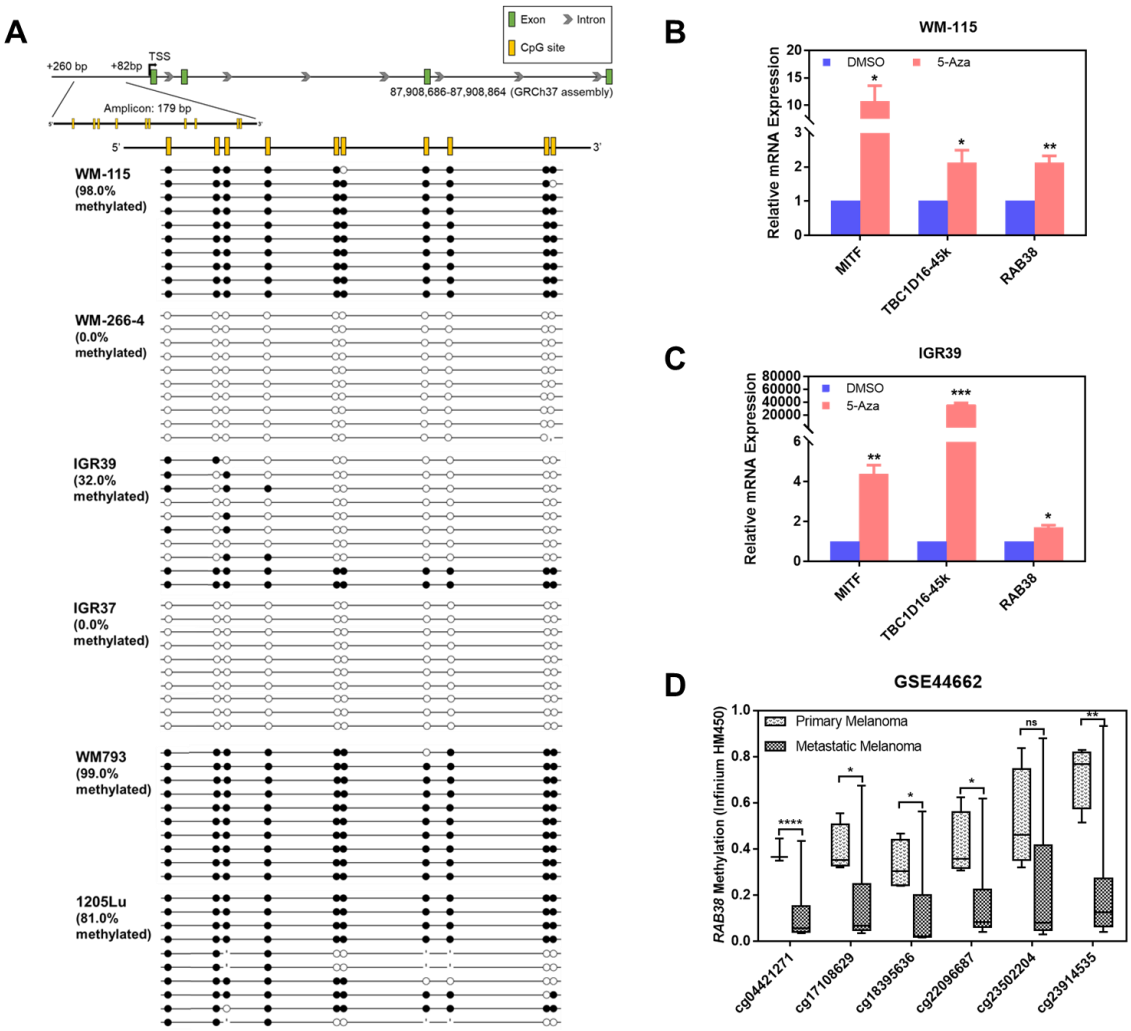
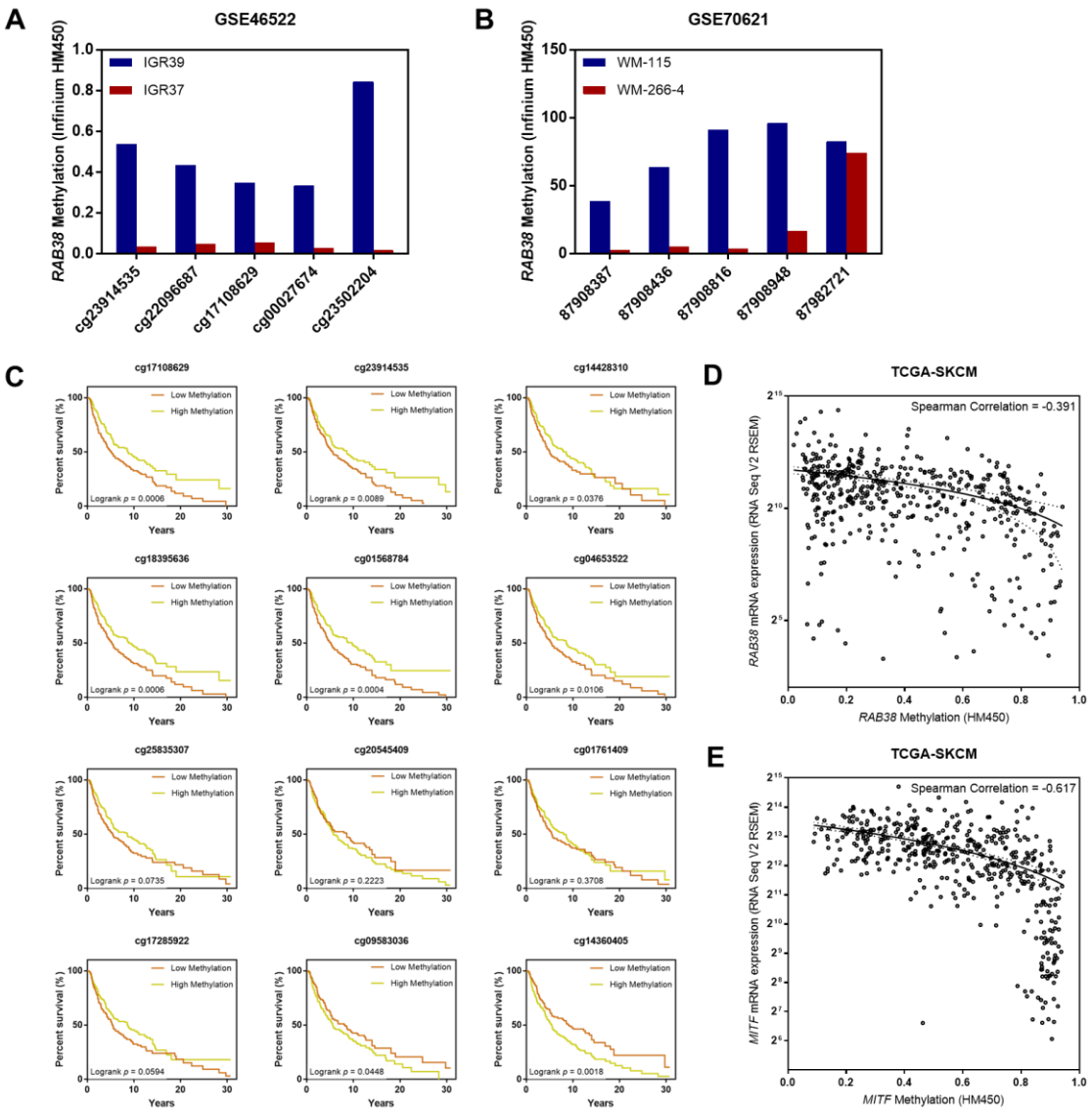

A
Exon
Intron
CpG site
+260 bp
+82bp
TSS
87,908,686-87,908,864 (GRCh37 assembly)
Amplicon: 179 bp
5'
3'
WM-115
(98.0%
methylated)
WM-266-4
(0.0%
methylated)
IGR39
(32.0%
methylated)
IGR37
(0.0%
methylated)
WM793
(99.0%
methylated)
1205Lu
(81.0%
methylated)
B
WM-115
DMSO
5-Aza
Relative mRNA Expression
MITF
TBC1D16-45k
RAB38
C
IGR39
DMSO
5-Aza
Relative mRNA Expression
MITF
TBC1D16-45k
RAB38
D
GSE44662
Primary Melanoma
Metastatic Melanoma
RAB38 Methylation (Infiniium HM450)
ns
cg04421271
cg17108629
cg18395636
cg22096687
cg23502204
cg23914535

Figure 2.14 Occurrence and prognostic values of *RAB38* promoter hypomethylation.

(A) Methylation microarray data showed lower methylation levels at multiple CpG sites of the *RAB38* promoter region in IGR37 cell line relative to IGR39 cell line in the GSE46522 cohort; **(B)** Methylation microarray data showed lower methylation levels at multiple CpG sites of the *RAB38* promoter region in WM-266-4 cell line relative to WM-115 cell line in the GSE70621 cohort; **(C)** Kaplan-Meier curves showing that the presence of *RAB38* hypomethylation at multiple CpG sites in melanoma patients (n = 464) is significantly associated with shorter overall survival in the TCGA-SKCM cohort; **(D)** Scatter plot showing the correlations of *RAB38* mRNA expressions with *RAB38* promoter CpG methylation in the TCGA-SKCM cohort; **(E)** Scatter plot showing the correlations of *MITF* mRNA expressions with *MITF* promoter CpG methylation in the TCGA-SKCM cohort.


A
GSE46522
RAB38 Methylation (Infinium HM450)
IGR39
IGR37
cg23914535
cg22096687
cg17108629
cg00027674
cg23502204
B
GSE70621
RAB38 Methylation (Infinium HM450)
WM-115
WM-266-4
87908387
87908436
87908816
87908948
87982721
C
cg17108629
cg23914535
cg14428310
cg18395636
cg01568784
cg04653522
cg25835307
cg20545409
cg01761409
cg17285922
cg09583036
cg14360405
Low Methylation
High Methylation
Percent survival (%)
Years
D
TCGA-SKCM
Spearman Correlation = -0.391
RAB38 mRNA expression (RNA Seq V2 RSEM)
RAB38 Methylation (HM450)
E
TCGA-SKCM
Spearman Correlation = -0.617
MITF mRNA expression (RNA Seq V2 RSEM)
MITF Methylation (HM450)

Chapter 3 Roles of Small GTPases in Acquired Tamoxifen Resistance in MCF-7 Cells Revealed by Targeted, Quantitative Proteomic Analysis

3.1 Introduction

Breast cancer is the most prevalent cancer among women worldwide. According to the American Cancer Society, 268,000 new breast cancer cases and 41,400 deaths were estimated in 2018 in the United States.[1] As a highly heterogeneous disease, breast cancer can be categorized into three major subtypes: estrogen receptor α-positive (ER-positive), human epidermal growth factor receptor 2 (HER2)-amplified (HER2-positive) and triple-negative breast cancer (TNBC). Among them, the ER-positive subtype remains the most prevalent and diverse, accounting for approximately 80% of diagnosed cases of breast cancer.[2] In this respect, antiestrogen drugs including tamoxifen and fulvestrant are frequently used, and tamoxifen remains the standard front-line endocrine therapy complementary to surgery. However, approximately half of patients who receive tamoxifen as the first-line therapy for recurrent diseases do not respond to the treatment because of intrinsic resistance. For those patients who initially respond to the drug, development of resistance is a major cause of treatment failure.[3]

Small GTPases of the Ras family are crucial regulators of intracellular trafficking and can mediate a wide range of biological events.[4] Several small GTPases play important roles in breast cancer progression, including RAB2A (tumorigenesis),[5] RAB27B (migration and invasion)[6], RAB31 (proliferation and metastasis),[7] and RND1 (tumorigenesis and invasion).[8] Defective endocytic pathways arising from down-regulation of small GTPases RAB5, RAC1, and RHOA were found in squamous-cell carcinomas (SCCs) carrying

resistance to chemotherapeutic drugs such as cisplatin.[9] Cellular pathways that involve small GTPase signaling, including the Ras–Raf-1–MAPK pathway, Rac1–PAK1 pathway, and Cdc42-mediated redox pathway, have also been shown to mediate responses to tamoxifen in breast cancer cells.[10-12] Therefore, we reason that a systematic interrogation of small GTPases involved in modulating tamoxifen resistance may offer a better understanding of the mechanisms of resistance in tamoxifen therapy.

Mass spectrometry-based proteomic methods have been widely employed for studying drug resistance and for discovering novel biomarkers and therapeutic targets in breast cancer.[13-15] Building upon the previously reported method of gel fractionation followed by liquid chromatography–tandem mass spectrometry (LC-MS/MS),[16, 17] we recently developed a targeted quantitative proteomic approach, relying on stable isotope labeling by amino acids in cell culture (SILAC), gel fractionation, and LC-MS/MS in scheduled multiple-reaction monitoring (MRM) mode, for high-throughput profiling of small GTPases.[18] In the present study, we employed this method to comprehensively investigate alterations in the expression of small GTPases during the development of tamoxifen resistance in a pair of matched wild-type (WT)/tamoxifen-resistant (TamR) breast cancer cell lines, i.e. MCF-7/WT and MCF-7/TamR. The quantitative proteomic data and cell-based assays with the use of two ER-positive cell lines (i.e. MCF-7 and T47D), together with bioinformatic analysis of publicly available data, led to the discovery of a novel role for RAB31 in modulating acquired tamoxifen resistance.

3.2 Experimental Section

3.2.1 Cell Culture

MCF-7 cells were purchased from ATCC (#HTB-22). The tamoxifen-resistant variant of MCF-7 cells (MCF-7/TamR) was generously provided by Dr. Guandi Wang (Xavier University). The T47D cells were kindly provided by Dr. Ameae Walker (University of California, Riverside). The cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS; Invitrogen-Gibco) and penicillin/streptomycin (100 IU/mL) at 37°C in an atmosphere with 5% CO_2. The MCF-7/TamR cells were continuously cultured in the above-described medium containing 0.10 μM (Z)-4-hydroxytamoxifen (4-OHT) (Sigma-Aldrich) for at least six months to allow them to develop resistance to the drug with an IC_{50} of ~5 μM 4-OHT.[13]

For other experiments, 4-OHT was dissolved in ethanol at a concentration of 10 mM and stored at -20° C. For SILAC experiments, [$^{13}C_6$,$^{15}N_2$]-L-lysine and [$^{13}C_6$]-L-arginine (Cambridge Isotopes Inc.), or their unlabeled counterparts, were added to SILAC DMEM medium depleted of L-lysine and L-arginine (Thermo Scientific Pierce) until their final concentrations reached 0.398 and 0.798 mM, respectively, to yield "heavy" and "light" media. Cells were cultured in the "heavy" SILAC DMEM medium for at least 6 cell doublings to ensure complete incorporation of heavy isotope-labeled amino acids.

3.2.2 Sample Preparation and Scheduled LC-MRM Analysis

To assess the differential expression of small GTPases in wild-type (WT) and tamoxifen-resistant (TamR) MCF-7 cells, we conducted SILAC-based quantitative proteomic experiments with forward and reverse labeling strategies. Briefly, we combined

lysates of light-labeled WT cells and heavy-labeled TamR cells at a 1:1 ratio in the forward labeling experiments. The reverse labeling experiment was conducted in the opposite way. The mixed cell lysates (100 μg in total) were loaded onto a 10% SDS-PAGE gel and separated by electrophoresis. The gel bands corresponding to the molecular weight range of 15–37 kDa were cut, reduced with 20 mM dithiothreitol, alkylated with 55 mM iodoacetamide, and digested in-gel with trypsin at an enzyme/protein ratio of 1:100. After tryptic digestion, the peptide mixtures were desalted and subjected to LC-MRM analyses. All LC-MRM experiments were performed on a TSQ vantage triple-quadruple mass spectrometer (Thermo Scientific) coupled with an EASY-nLC II system (Thermo Scientific). The samples were automatically loaded onto a 4-cm trapping column (150 μm i.d.) packed with ReproSil-Pur 120 C18-AQ resin (5 μm in particle size and 120 Å in pore size, Dr. Maisch GmbH HPLC) at 3 μL/min. The trapping column was coupled to a 20-cm fused silica analytical column (75 μm i.d.) packed with ReproSil-Pur 120 C18-AQ resin (3 μm in particle size and 120 Å in pore size, Dr. Maisch GmbH HPLC). The peptide mixtures were then separated using a 157-min linear gradient of 2−35% acetonitrile in 0.1% formic acid at a flow rate of 230 nL/min. The spray voltage was set as 1.8 kV. Ions were isolated in both Q1 and Q3 using 0.7 fwhm resolution, for which the cycle time was set as 5 s. The optimal collisional energy (CE) for each targeted peptide was calculated using a linear equation specific to the TSQ Vantage instrument and the precursor mass-to-charge ratio (*m/z*) according to the default setting in Skyline.[19]

To enable high-throughput quantitative analysis, we applied a previously developed scheduled LC-MRM method[18], in which the mass spectrometer was programmed to

acquire the MS/MS of the precursor ions for a limited number of peptides in each 6-min retention time (RT) window. The MRM data for all targeted peptides were manually inspected to ensure correct peak picking. In this regard, the dot-product (dotp) value has to exceed 0.80.[20] In addition, the iRT value represents an intrinsic property (i.e., hydrophobicity) of a peptide; hence, a substantial deviation of measured RT from that projected from the linear plot of RT over iRT signals a false-positive detection.

3.2.3 Data Sources for Bioinformatic Analyses

Patient RNAseq data were obtained from The Cancer Genome Atlas (TCGA) via cBioPortal (http://www.cbioportal.org/).[21] The complete clinical data files were downloaded from the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/geo/). The Cancer Cell Line Encyclopedia (CCLE) (http://www.broadinstitute.org/ccle/home) were interrogated for the comprehensive evaluation of mRNA expression for candidate genes among more than 100 breast cancer cell lines.[22] Publicly available transcriptomic profiles with accession numbers GSE3494, GSE4922, GSE6434, GSE24460, GSE26495 and GSE42568 were downloaded from the GEO database and analyzed using R (version 3.4.3).

3.2.4 Patient Survival Analysis

Kaplan–Meier survival curves were generated using an online database Kaplan-Meier plotter (kmplotter.com) for breast cancer.[23] Data were analyzed using the JetSet best probe set to analyze gene expression and relapse-free survival (RFS). Briefly, gene names were entered into the database to obtain Kaplan–Meier survival plot where the hazard ratio (HR), 95% confidence intervals (95% CI) and logrank *p* values were calculated and displayed.

3.2.5 Plasmid Construction

pLKO.1-shRAB31 and pLKO.1-scramble plasmids were constructed by inserting a short hairpin double-stranded oligonucleotide targeting RAB31 or a scrambled sequence into the pLKO.1 lentiviral vector. Human RAB31 cDNA was amplified from an in-house cDNA library and subcloned into the p3×Flag-CMV10 vector.

3.2.6 Generation of Stable Knockdown Cell Lines

The lentiviral vectors pLKO.1-shRAB31 and pLKO.1-scramble were generated as described above. Recombinant lentiviruses were produced by co-transfection of HEK293T cells with the pLKO.1-scramble or pLKO.1-shRAB31 shRNA plasmids, envelope plasmid pLTR-G (Addgene #17532) and packaging plasmid pCMV-dR8.2 dvpr (Addgene #8455). Lentivirus-containing supernatant was harvested and filtered through 0.45-μm pore size filters at 48 h post-transfection. Infection of MCF-7 or T47D cells with recombinant lentivirus was conducted in the presence of 5 μg/ml polybrene. After removal of virus, the cells were selected in 1 μg/mL puromycin-containing medium for 3 days to eliminate uninfected cells. After selection, the cells were maintained in medium containing 1 μg/mL puromycin and used for subsequent experiments.

3.2.7 Cell Proliferation Assay

The proliferation of cells under 4-OHT treatment was evaluated using a cell counting kit-8 (CCK-8; Dojindo Laboratories). Briefly, MCF-7/WT, T47D and MCF-7/TamR cells were seeded in a 96-well flat-bottomed microplate (3000 cells/well) in complete growth medium (100 μL/well) for 24 h. The cells were then incubated with or without various concentrations of 4-OHT for 5 days in the dose-dependent experiments, or with 1 μM 4-

OHT in the time-dependent experiments, and ethanol was used as the vehicle control. At the end of treatments, 10 μL of the CCK-8 dye was added to each well and the cells were incubated at 37℃ for 4 h prior to using the Synergy™ H1 Hybrid Multi-Mode Microplate Reader (BioTek Instruments) for measuring the absorbance at 450 nm.

3.2.8 Colony Formation Assay

MCF-7/WT or MCF-7/TamR cells were cultured in complete growth media. The cells were seeded at a density of 2000 cells/well in 2 mL of medium in six-well plates and allowed to adhere overnight. The next day, the cells were treated with 1 μM 4-OHT, and an equal volume of ethanol was used as a vehicle control. The cells were then allowed to grow until colonies reached > 50 cells per colony for the control group (approximately 10 to 14 days). Colonies were then fixed with glutaraldehyde for 30 min, stained with crystal violet (0.1% in 20% methanol) for 30 min, and washed. Colony numbers were determined manually. The experiments were conducted in triplicate, and the data represent means ± standard errors of the means.

3.3 Results and Discussion

3.3.1 Application of a High-throughput LC-MRM Assay for Studying Acquired Tamoxifen Resistance

We set out to explore the alterations of small GTPases during the development of tamoxifen resistance in ER-positive breast cancer. To this end, we employed our recently developed scheduled MRM-based targeted proteomic method[18] to assess, in high-throughput, the reprogramming of the small GTPase proteome during development of tamoxifen resistance. The method involved metabolic labeling of MCF-7 cells and

isogenic cells that are resistant to tamoxifen using SILAC, SDS-PAGE for the enrichment of proteins in the molecular weight range of 15-37 kDa (the molecular weights for ~95% of small GTPases fall in this range), in-gel tryptic digestion, and LC-MS/MS analysis of the resulting tryptic peptides in the scheduled MRM mode (Figure 3.1).

To obtain reliable quantification results, we conducted SILAC experiments in triplicate, with two sets of forward labeling and one set of reverse labeling. The method facilitated the quantification of a total of 96 small GTPases, among which 13 and 10 proteins were significantly down- and up-regulated (with >1.5-fold change), respectively, in the drug-resistant MCF-7 cells relative to in the parental line (Figure 3.1). In this vein, we chose a cutoff of 1.5-fold change on the basis of the average relative standard deviation (RSD = 14%) for all the quantified small GTPases. The method facilitated the coverage of approximately 65% of the human small GTPase proteome in two LC-MRM runs.

We found that several RAB small GTPases were down-regulated (e.g., RAB27B, RAB30, RAB31, and RAB32), whereas several others were up-regulated (e.g., RAB7A, RAB18, and RAB6B), in tamoxifen-resistant cells (Figure 3.1). The extracted-ion chromatograms (XICs) for several differentially expressed small GTPases were displayed in Figure 3.2, including ARL3 (ARF subfamily), RHOF (RHO subfamily), RAB30 (RAB subfamily), and RRAS2 (RAS subfamily). Notably, we observed a ~2-fold up-regulation of RRAS2 protein in the tamoxifen-resistant MCF-7 cells. In this vein, RRAS2 is known to promote primary tumorigenesis and late steps of metastasis in breast cancer cells, and it can also contribute to increased resistance to tamoxifen.[24-26] Hence, our method validated

the differential expression of a small GTPase that was previously shown to be involved with tamoxifen resistance.

There have been no literature precedents about the functions of ARL3, RHOF, and RRAS in drug resistance in breast cancer. Nevertheless, ARL3 was suggested to be transcriptionally regulated by ER-related mechanisms.[27] In addition, RRAS inhibits the proliferation, migration and cell cycle progression of cultured breast cancer cells.[28] RHOF plays an important role in controlling the formation of filopodia, which may contribute to proliferation, invasion and formation of micrometastases of cancer cells. RHOF, however, enhances the resistance of pancreatic cancer to gemcitabine through regulation of the epithelial-to-mesenchymal transition.[29] The potential roles of ARL3, RRAS, and RHOF in modulating tamoxifen resistance warrant future investigation.

To assess the performance of our MRM-based method, we also analyzed the same samples by employing LC-MS/MS in the data-dependent acquisition (DDA) mode. As shown in Figure 3.3, DDA analyses only led to the identification of 51 and 45 small GTPases in the two forward-SILAC samples (F1 and F2, respectively), and 44 small GTPases in the reverse-SILAC sample (R). In stark contrast, the LC-MRM approach led to substantially higher coverage of the small GTPase proteome (Figure 3.3). As noted, the quantification was based on three independent LC-MRM experiments, which included two forward- and one reverse-SILAC labeling experiments, and the small GTPases reported were reproducibly quantified in all three replicates, with the mean RSD being 14%. Therefore, these results demonstrated that the scheduled LC-MRM method outperformed the shotgun proteomic approach in terms of reproducibility and sensitivity. Additionally,

we observed an excellent linear fit (R^2 = 0.9238) for the $\log_2$-transformed SILAC ratios of all the quantified small GTPases obtained from one forward- and one reverse-SILAC labeling experiment (Figure 3.3), which again underscored the excellent reproducibility of the method. With respect to retention time (RT) scheduling, all of the 10 standard peptides derived from BSA exhibited an excellent linear fit (R^2 = 0.9996) between the observed RTs and the iRT values in the library (Figure 3.3). Highly reliable and reproducible RT prediction was also reflected by the superb linearity for the RTs observed for the small GTPase peptides in different replicates (Figure 3.3).

To achieve confident identification of targeted peptides, we manually processed the LC-MRM data to ensure the coelution of the MRM transitions with the dot-product (dotp) value being >0.8, as calculated by Skyline.[20] The dotp of a peptide was based on the correlation for the ratio of the MRM peak intensities observed versus those in the library MS/MS acquired from shotgun proteomic experiments. Shown in Figure 3.4 are the representative XICs displaying nearly identical retention times for different MRM transitions monitored for each of the three distinct peptides of RAB31 with high dotp values (>0.95). Relative quantification was achieved by integrating the areas of the peaks found in the XICs acquired from the LC-MRM analyses (Figure 3.4 shows the MRM traces acquired from the forward- and reverse-SILAC labeling experiments). The down-regulation of RAB31 protein in the tamoxifen-resistant line was further confirmed by Western-blot analysis (Figure 3.4). In addition to RAB31, we also validated the differential expression of RAB27A and RAB27B proteins by Western-blot analysis (Figure 3.5),

supporting that the LC-MRM method is capable of accurately profiling the differential expression of small GTPases.

3.3.2 Prognostic Values of RAB31 in Breast Cancer Revealed by Bioinformatic Analyses

As discussed above, our LC-MRM results confirmed a known protein target that promotes acquired tamoxifen resistance: RRAS2. We next sought to further examine the roles of other differentially expressed small GTPases in breast cancer by performing Kaplan–Meier survival analysis for significantly up- and down-regulated proteins shown in Figure 3.1. Particularly, for down-regulated proteins, lower mRNA expression levels of *ARL3*, *RHOF* and *RAB30* were significantly associated with poor relapse-free survival (RFS) (Figure 3.6). In contrast, higher *RRAS2* mRNA levels are significantly correlated with poor RFS (Figure 3.6). For patients who received endocrine therapy (tamoxifen only), increased *RAB31* expression predicted better outcomes, whereas for those without endocrine therapy, RAB31 did not serve as an effective indicator for RFS (Figure 3.7). Notably, among those genes that exhibited significantly altered protein abundance in tamoxifen resistance, only the mRNA expression of *RAB31* is significantly correlated with tamoxifen efficacy in ER-positive breast cancer patients, but not that of *ARL3*, *RAB30*, *RRAS2* or *RHOF* (Figures 3.6 and 3.7). We also assessed the prognostic values of RAB31 stratified by ER status and found only a significant correlation between *RAB31* expression and survival in ER-positive, but not ER-negative breast cancer patients (Figure 3.7).

We also interrogated several Gene Expression Omnibus (GEO) data sets for Kaplan–Meier survival analysis and differential mRNA expression analysis: GSE3494 (Karolinska Institute, n = 236),[30] GSE4922 (Genome Institute of Singapore, n = 347),[31] GSE6434

(Baylor College of Medicine, n = 24),[32] GSE42568 (Dublin City University, n = 115),[33] GSE24460 (National Cancer Institute, n = 52),[34] and GSE26495 (Emory University).[35] The GSE6434, GSE26495 and GSE24460 data sets revealed significantly decreased mRNA levels of *RAB31* in docetaxel-, tamoxifen- and doxorubicin-resistant MCF-7 cell lines, respectively (Figure 3.8). In this respect, docetaxel and doxorubicin are other small-molecule drugs commonly used in chemotherapy against breast cancer. Therefore, these results suggest that RAB31 down-regulation could be a common event in drug-resistant MCF-7 cells. For survival analysis, lower *RAB31* expression was significantly associated with poorer RFS and overall survival (OS) in the GSE42568 cohort (Figure 3.8). Meanwhile, the same trend holds for the disease-specific survival (DSS) in the GSE3494 data set (Figure 3.8). Notably, in the GSE4922 cohort, RAB31 is significantly down-regulated in the most metastatic grade III tumors as compared with in grade I tumors (Figure 3.8), supporting a potential role for RAB31 in breast cancer progression.[36] Taken together, these results suggest that RAB31 may serve as a potential biomarker for predicting acquired tamoxifen resistance in breast cancer patients.

3.3.3 Correlation of RAB31 Expression with ER Status and Breast Cancer Subtypes

To elucidate the association between *RAB31* expression and breast cancer subtypes, we next interrogated a large and comprehensive breast cancer patient cohort: The Cancer Genome Atlas Breast Invasive Carcinoma (TCGA-BRCA). Our results showed that *RAB31* mRNA expression levels were significantly correlated with the mRNA expression levels of the *ESR1* gene, which encodes ERα, in the TCGA-BRCA cohort (Figure 3.9). In addition, the same correlations could be observed for the GSE4922 and GSE23988 data

sets (Figure 3.9). Aside from ER status, we extended the expression analysis to elucidate the association between *RAB31* mRNA levels and different molecular subtypes of breast cancer: luminal A, luminal B, HER2-enriched, basal-like, and normal-like.[37] As shown in Figure 3.9, the mRNA expression levels of *RAB31* were significantly lower in the basal subtype and the HER2-enriched subtype than in the luminal-A subtype. Clinical data demonstrate reduced responses to endocrine therapy in tumors with HER2 amplification. Suppressed expression of RAB31 might therefore be correlated with reduced ER expression and the HER2-overexpressing resistant phenotypes.

3.3.4 RAB31 Knockdown Rendering Elevated Tamoxifen Resistance in MCF-7/WT and T47D Cells

The above results showed that diminished *RAB31* expression confers poor prognosis for ER-positive breast cancer patients, especially for those who received tamoxifen as a first-line therapy. Hence, we reason that RAB31 may be associated with tamoxifen efficacy and therefore may play a role in tamoxifen resistance. To further examine the role of RAB31 in tamoxifen resistance, we used lentiviral transduced shRNA to enable stable knockdown of *RAB31* gene expression in MCF-7 cells. Western-blot analysis showed that two separate shRNA sequences gave rise to 60–70% depletions of RAB31 protein compared with that from the control shRNA sequence (shScramble) in MCF-7 cells (Figure 3.10). We next performed cell proliferation assay to assess the effects of RAB31 knockdown on tamoxifen response in MCF-7 cells. We found that, upon treatment with 1 μM (Z)-4-hydroxytamoxifen (4-OHT), an active metabolite of tamoxifen, genetic depletion of *RAB31* led to significantly higher cell viability at 72 and 96 h, but not at 24 or

48 h (Figure 3.10). Moreover, RAB31 knockdown led to higher resistance to 4-OHT treatment in a broad range of doses (Figure 3.10). Collectively, the results support that loss of RAB31 could modulate tamoxifen response in MCF-7 cells.

To further substantiate the above findings, we assessed how shRNA-mediated depletion of RAB31 in T47D cells alters the sensitivity of these cells toward tamoxifen. Consistent with our hypothesis, diminished expression of RAB31 in both MCF-7 and T47D cells led to increased proliferation rates and elevated resistance toward tamoxifen (Figure 3.11). In addition, RAB31 knockdown gave rise to augmented resistance to tamoxifen in T47D cells (Figure 3.11). Taken together, the above results support the role of RAB31 in tamoxifen efficacy in two ER-positive breast cancer cell lines, i.e. MCF-7 and T47D.

3.3.5 Sensitization of MCF-7/TamR Cells to 4-OHT Treatment by RAB31 Overexpression

We next examined how a gain of function of RAB31 modulates tamoxifen sensitivity by ectopically overexpressing the *RAB31* gene in tamoxifen-resistant MCF-7/TamR cells. The cell proliferation assay results indicated that RAB31-overexpressing MCF-7/TamR cells exhibited elevated sensitivity towards 4-OHT treatment as compared to the empty vector group in both dose-dependent and time-dependent manner (Figure 3.12). In agreement with the cell proliferation assay results, colony formation assays showed that ectopic expression of RAB31 conferred significantly increased tamoxifen sensitivity for the MCF-7/TamR cells (Figure 3.12). In summary, our results substantiated the role of RAB31 in suppressing acquired tamoxifen resistance in MCF-7/TamR cells.

3.4 Conclusions

In the present study, we applied a targeted quantitative proteomic method to assess the differential expression of small GTPases in paired wild-type and tamoxifen-resistant MCF-7 cells. The method provided high-throughput, accurate and reproducible quantifications of the relative expression levels of >90 small GTPases in the paired cell lines. Among them, 10 and 13 small GTPases were up- and down-regulated by at least 1.5-fold in the drug-resistant cells compared with in the parental MCF-7 cells. In this context, it is worth noting that our current targeted proteomic approach relies on SILAC labeling, which is not applicable to clinical specimens (e.g. biological fluids and tumor tissue samples) from breast cancer patients that manifest *de novo* or acquired tamoxifen resistance. Nevertheless, the targeted proteomic method can be adapted for handling clinical specimens with the use of heavy isotope-labeled synthetic peptides, and such an approach is currently being explored in our laboratory.

Combined with bioinformatic analyses, we identified RAB31 as a novel predictive marker for acquired tamoxifen resistance. Through the use of two ER-positive cell lines (i.e. MCF-7 and T47D), we uncovered, for the first time, a role of RAB31 in modulating tamoxifen sensitivity. The functions of RAB31 were also explored in other cancer types.[38] RAB31 can localize to endocytic compartments and functions in the post-Golgi, endocytic or exocytic trafficking of the epidermal growth factor receptor (EGFR) in A431 and HeLa cells.[39] Silencing of RAB31 inhibited the endocytic trafficking of the ligand-bound EGFR to late endosomes and its subsequent degradation. It is also worth noting that increased levels of receptor tyrosine kinases including EGFR and HER2 can directly alter the cellular

response to tamoxifen.[40] Therefore, we reason that down-regulation of RAB31 in MCF-7/TamR cells may perturb EGFR trafficking, thereby influencing drug resistance.

In addition to receptor tyrosine kinase signaling, ER-associated proteins are essential players in tamoxifen resistance. Some suggest elevated ESR1 mRNA levels in tamoxifen-resistant ER-positive breast cancer cells.[25] However, several lines of evidence demonstrated suppressed yet functional ER-regulated signaling pathway in MCF-7/TamR cells, as reflected by down-regulated *ESR1* expression.[13, 41-43] Notably, *RAB31* is among the 11 genes that are robustly overexpressed in ER-positive breast carcinoma samples,[44] which is in line with the observation that the promoter region of the *RAB31* gene harbors an ER-responsive element.[45] On the basis of these findings, we reason that down-regulated RAB31 in tamoxifen-resistant MCF-7 cells may arise from ER-modulated molecular adaptations.

Taken together, we showed, for the first time, that targeted quantitative proteomics, in combination with bioinformatics, provided novel insights into the roles of small GTPases in acquired tamoxifen resistance in breast cancer. It can be envisaged that the method can also be employed for understanding the therapeutic resistance of other types of cancer.

1. Siegel, R. L.; Miller, K. D.; Jemal, A., Cancer statistics, 2018. *CA Cancer J Clin* **2018,** *68* (1), 7-30.
2. Dai, X.; Chen, A.; Bai, Z., Integrative investigation on breast cancer in ER, PR and HER2-defined subgroups using mRNA and miRNA expression profiling. *Sci Rep* **2014,** *4*, 6566.
3. Meijer, D.; van Agthoven, T.; Bosma, P. T.; Nooter, K.; Dorssers, L. C., Functional screen for genes responsible for tamoxifen resistance in human breast cancer cells. *Mol Cancer Res* **2006,** *4* (6), 379-86.
4. Takai, Y.; Sasaki, T.; Matozaki, T., Small GTP-binding proteins. *Physiol Rev* **2001,** *81* (1), 153-208.
5. Luo, M. L.; Gong, C.; Chen, C. H.; Hu, H.; Huang, P.; Zheng, M.; Yao, Y.; Wei, S.; Wulf, G.; Lieberman, J.; Zhou, X. Z.; Song, E.; Lu, K. P., The Rab2A GTPase promotes breast cancer stem cells and tumorigenesis via Erk signaling activation. *Cell Rep* **2015,** *11* (1), 111-24.
6. Hendrix, A.; Maynard, D.; Pauwels, P.; Braems, G.; Denys, H.; Van den Broecke, R.; Lambert, J.; Van Belle, S.; Cocquyt, V.; Gespach, C.; Bracke, M.; Seabra, M. C.; Gahl, W. A.; De Wever, O.; Westbroek, W., Effect of the secretory small GTPase Rab27B on breast cancer growth, invasion, and metastasis. *J Natl Cancer Inst* **2010,** *102* (12), 866-80.
7. Grismayer, B.; Solch, S.; Seubert, B.; Kirchner, T.; Schafer, S.; Baretton, G.; Schmitt, M.; Luther, T.; Kruger, A.; Kotzsch, M.; Magdolen, V., Rab31 expression levels modulate tumor-relevant characteristics of breast cancer cells. *Mol Cancer* **2012,** *11*, 62.
8. Okada, T.; Sinha, S.; Esposito, I.; Schiavon, G.; Lopez-Lago, M. A.; Su, W.; Pratilas, C. A.; Abele, C.; Hernandez, J. M.; Ohara, M.; Okada, M.; Viale, A.; Heguy, A.; Socci, N. D.; Sapino, A.; Seshan, V. E.; Long, S.; Inghirami, G.; Rosen, N.; Giancotti, F. G., The Rho GTPase Rnd1 suppresses mammary tumorigenesis and EMT by restraining Ras-MAPK signalling. *Nat Cell Biol* **2015,** *17* (1), 81-94.
9. Shen, D. W.; Su, A.; Liang, X. J.; Pai-Panandiker, A.; Gottesman, M. M., Reduced expression of small GTPases and hypermethylation of the folate binding protein gene in cisplatin-resistant cells. *Br J Cancer* **2004,** *91* (2), 270-6.
10. McGlynn, L. M.; Kirkegaard, T.; Edwards, J.; Tovey, S.; Cameron, D.; Twelves, C.; Bartlett, J. M.; Cooke, T. G., Ras/Raf-1/MAPK pathway mediates response to tamoxifen but not chemotherapy in breast cancer patients. *Clin Cancer Res* **2009,** *15* (4), 1487-95.
11. Gonzalez, N.; Cardama, G. A.; Comin, M. J.; Segatori, V. I.; Pifano, M.; Alonso, D. F.; Gomez, D. E.; Menna, P. L., Pharmacological inhibition of Rac1-PAK1 axis restores tamoxifen sensitivity in human resistant breast cancer cells. *Cell Signal* **2017,** *30*, 154-161.
12. Chen, H. Y.; Yang, Y. M.; Stevens, B. M.; Noble, M., Inhibition of redox/Fyn/c-Cbl pathway function by Cdc42 controls tumour initiation capacity and tamoxifen sensitivity in basal-like breast cancer cells. *EMBO Mol Med* **2013,** *5* (5), 723-36.
13. Zhou, C.; Zhong, Q.; Rhodes, L. V.; Townley, I.; Bratton, M. R.; Zhang, Q.; Martin, E. C.; Elliott, S.; Collins-Burow, B. M.; Burow, M. E.; Wang, G., Proteomic analysis of acquired tamoxifen resistance in MCF-7 cells reveals expression signatures associated with enhanced migration. *Breast Cancer Res* **2012,** *14* (2), R45.

14. Hengel, S. M.; Murray, E.; Langdon, S.; Hayward, L.; O'Donoghue, J.; Panchaud, A.; Hupp, T.; Goodlett, D. R., Data-independent proteomic screen identifies novel tamoxifen agonist that mediates drug resistance. *J Proteome Res* **2011,** *10* (10), 4567-78.
15. Umar, A.; Kang, H.; Timmermans, A. M.; Look, M. P.; Meijer-van Gelder, M. E.; den Bakker, M. A.; Jaitly, N.; Martens, J. W.; Luider, T. M.; Foekens, J. A.; Pasa-Tolic, L., Identification of a putative protein profile associated with tamoxifen therapy resistance in breast cancer. *Mol Cell Proteomics* **2009,** *8* (6), 1278-94.
16. Halvey, P. J.; Ferrone, C. R.; Liebler, D. C., GeLC-MRM quantitation of mutant KRAS oncoprotein in complex biological samples. *J Proteome Res* **2012,** *11* (7), 3908-13.
17. Zhang, C. C.; Li, R.; Jiang, H.; Lin, S.; Rogalski, J. C.; Liu, K.; Kast, J., Development and application of a quantitative multiplexed small GTPase activity assay using targeted proteomics. *J Proteome Res* **2015,** *14* (2), 967-76.
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* **2018,** *78* (18), 5431-5445.
19. MacLean, B.; Tomazela, D. M.; Shulman, N.; Chambers, M.; Finney, G. L.; Frewen, B.; Kern, R.; Tabb, D. L.; Liebler, D. C.; MacCoss, M. J., Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* **2010,** *26* (7), 966-8.
20. Kawahara, R.; Bollinger, J. G.; Rivera, C.; Ribeiro, A. C.; Brandao, T. B.; Paes Leme, A. F.; MacCoss, M. J., A targeted proteomic strategy for the measurement of oral cancer candidate biomarkers in human saliva. *Proteomics* **2016,** *16* (1), 159-73.
21. Cerami, E.; Gao, J. J.; Dogrusoz, U.; Gross, B. E.; Sumer, S. O.; Aksoy, B. A.; Jacobsen, A.; Byrne, C. J.; Heuer, M. L.; Larsson, E.; Antipin, Y.; Reva, B.; Goldberg, A. P.; Sander, C.; Schultz, N., The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data. *Cancer Discovery* **2012,** *2* (5), 401-404.
22. Barretina, J.; Caponigro, G.; Stransky, N.; Venkatesan, K.; Margolin, A. A.; Kim, S.; Wilson, C. J.; Lehar, J.; Kryukov, G. V.; Sonkin, D.; Reddy, A.; Liu, M.; Murray, L.; Berger, M. F.; Monahan, J. E.; Morais, P.; Meltzer, J.; Korejwa, A.; Jane-Valbuena, J.; Mapa, F. A.; Thibault, J.; Bric-Furlong, E.; Raman, P.; Shipway, A.; Engels, I. H.; Cheng, J.; Yu, G. K.; Yu, J.; Aspesi, P., Jr.; de Silva, M.; Jagtap, K.; Jones, M. D.; Wang, L.; Hatton, C.; Palescandolo, E.; Gupta, S.; Mahan, S.; Sougnez, C.; Onofrio, R. C.; Liefeld, T.; MacConaill, L.; Winckler, W.; Reich, M.; Li, N.; Mesirov, J. P.; Gabriel, S. B.; Getz, G.; Ardlie, K.; Chan, V.; Myer, V. E.; Weber, B. L.; Porter, J.; Warmuth, M.; Finan, P.; Harris, J. L.; Meyerson, M.; Golub, T. R.; Morrissey, M. P.; Sellers, W. R.; Schlegel, R.; Garraway, L. A., The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **2012,** *483* (7391), 603-7.
23. Gyorffy, B.; Lanczky, A.; Eklund, A. C.; Denkert, C.; Budczies, J.; Li, Q.; Szallasi, Z., An online survival analysis tool to rapidly assess the effect of 22,277 genes on breast cancer prognosis using microarray data of 1,809 patients. *Breast Cancer Res Treat* **2010,** *123* (3), 725-31.
24. Rokavec, M.; Schroth, W.; Amaral, S. M.; Fritz, P.; Antoniadou, L.; Glavac, D.; Simon, W.; Schwab, M.; Eichelbaum, M.; Brauch, H., A polymorphism in the TC21

promoter associates with an unfavorable tamoxifen treatment outcome in breast cancer. *Cancer Res* **2008,** *68* (23), 9799-808.
25. Mendes-Pereira, A. M.; Sims, D.; Dexter, T.; Fenwick, K.; Assiotis, I.; Kozarewa, I.; Mitsopoulos, C.; Hakas, J.; Zvelebil, M.; Lord, C. J.; Ashworth, A., Genome-wide functional screen identifies a compendium of genes affecting sensitivity to tamoxifen. *Proc Natl Acad Sci U S A* **2012,** *109* (8), 2730-5.
26. Larive, R. M.; Moriggi, G.; Menacho-Marquez, M.; Canamero, M.; de Alava, E.; Alarcon, B.; Dosil, M.; Bustelo, X. R., Contribution of the R-Ras2 GTP-binding protein to primary breast tumorigenesis and late-stage metastatic disease. *Nat Commun* **2014,** *5*, 3881.
27. Zhou, X.; Shi, T.; Li, B.; Zhang, Y.; Shen, X.; Li, H.; Hong, G.; Liu, C.; Guo, Z., Genes dysregulated to different extent or oppositely in estrogen receptor-positive and estrogen receptor-negative breast cancers. *PLoS One* **2013,** *8* (7), e70017.
28. Song, J.; Zheng, B.; Bu, X.; Fei, Y.; Shi, S., Negative association of R-Ras activation and breast cancer development. *Oncol Rep* **2014,** *31* (6), 2776-84.
29. Yang, R. M.; Zhan, M.; Xu, S. W.; Long, M. M.; Yang, L. H.; Chen, W.; Huang, S.; Liu, Q.; Zhou, J.; Zhu, J.; Wang, J., miR-3656 expression enhances the chemosensitivity of pancreatic cancer to gemcitabine through modulation of the RHOF/EMT axis. *Cell Death Dis* **2017,** *8* (10), e3129.
30. Miller, L. D.; Smeds, J.; George, J.; Vega, V. B.; Vergara, L.; Ploner, A.; Pawitan, Y.; Hall, P.; Klaar, S.; Liu, E. T.; Bergh, J., An expression signature for p53 status in human breast cancer predicts mutation status, transcriptional effects, and patient survival. *Proc Natl Acad Sci U S A* **2005,** *102* (38), 13550-5.
31. Ivshina, A. V.; George, J.; Senko, O.; Mow, B.; Putti, T. C.; Smeds, J.; Lindahl, T.; Pawitan, Y.; Hall, P.; Nordgren, H.; Wong, J. E.; Liu, E. T.; Bergh, J.; Kuznetsov, V. A.; Miller, L. D., Genetic reclassification of histologic grade delineates new clinical subtypes of breast cancer. *Cancer Res* **2006,** *66* (21), 10292-301.
32. Chang, J. C.; Wooten, E. C.; Tsimelzon, A.; Hilsenbeck, S. G.; Gutierrez, M. C.; Tham, Y. L.; Kalidas, M.; Elledge, R.; Mohsin, S.; Osborne, C. K.; Chamness, G. C.; Allred, D. C.; Lewis, M. T.; Wong, H.; O'Connell, P., Patterns of resistance and incomplete response to docetaxel by gene expression profiling in breast cancer patients. *J Clin Oncol* **2005,** *23* (6), 1169-77.
33. Clarke, C.; Madden, S. F.; Doolan, P.; Aherne, S. T.; Joyce, H.; O'Driscoll, L.; Gallagher, W. M.; Hennessy, B. T.; Moriarty, M.; Crown, J.; Kennedy, S.; Clynes, M., Correlating transcriptional networks to breast cancer survival: a large-scale coexpression analysis. *Carcinogenesis* **2013,** *34* (10), 2300-8.
34. Calcagno, A. M.; Salcido, C. D.; Gillet, J. P.; Wu, C. P.; Fostel, J. M.; Mumau, M. D.; Gottesman, M. M.; Varticovski, L.; Ambudkar, S. V., Prolonged drug selection of breast cancer cells and enrichment of cancer stem cell characteristics. *J Natl Cancer Inst* **2010,** *102* (21), 1637-52.
35. Duraiswamy, J.; Ibegbu, C. C.; Masopust, D.; Miller, J. D.; Araki, K.; Doho, G. H.; Tata, P.; Gupta, S.; Zilliox, M. J.; Nakaya, H. I.; Pulendran, B.; Haining, W. N.; Freeman, G. J.; Ahmed, R., Phenotype, function, and gene expression profiles of

programmed death-1(hi) CD8 T cells in healthy human adults. *J Immunol* **2011,** *186* (7), 4200-12.
36. Dalton, L. W.; Pinder, S. E.; Elston, C. E.; Ellis, I. O.; Page, D. L.; Dupont, W. D.; Blamey, R. W., Histologic grading of breast cancer: linkage of patient outcome with level of pathologist agreement. *Mod Pathol* **2000,** *13* (7), 730-5.
37. Liu, M. C.; Pitcher, B. N.; Mardis, E. R.; Davies, S. R.; Friedman, P. N.; Snider, J. E.; Vickery, T. L.; Reed, J. P.; DeSchryver, K.; Singh, B.; Gradishar, W. J.; Perez, E. A.; Martino, S.; Citron, M. L.; Norton, L.; Winer, E. P.; Hudis, C. A.; Carey, L. A.; Bernard, P. S.; Nielsen, T. O.; Perou, C. M.; Ellis, M. J.; Barry, W. T., PAM50 gene signatures and breast cancer prognosis with adjuvant anthracycline- and taxane-based chemotherapy: correlative analysis of C9741 (Alliance). *NPJ Breast Cancer* **2016,** *2*.
38. Chua, C. E.; Tang, B. L., The role of the small GTPase Rab31 in cancer. *J Cell Mol Med* **2015,** *19* (1), 1-10.
39. Chua, C. E.; Tang, B. L., Engagement of the small GTPase Rab31 protein and its effector, early endosome antigen 1, is important for trafficking of the ligand-bound epidermal growth factor receptor from the early to the late endosome. *J Biol Chem* **2014,** *289* (18), 12375-89.
40. Jin, K.; Kong, X.; Shah, T.; Penet, M. F.; Wildes, F.; Sgroi, D. C.; Ma, X. J.; Huang, Y.; Kallioniemi, A.; Landberg, G.; Bieche, I.; Wu, X.; Lobie, P. E.; Davidson, N. E.; Bhujwalla, Z. M.; Zhu, T.; Sukumar, S., The HOXB7 protein renders breast cancer cells resistant to tamoxifen through activation of the EGFR pathway. *Proc Natl Acad Sci U S A* **2012,** *109* (8), 2736-41.
41. Liang, Y. K.; Zeng; Xiao, Y. S.; Wu, Y.; Ouyang, Y. X.; Chen, M.; Li, Y. C.; Lin, H. Y.; Wei, X. L.; Zhang, Y. Q.; Kruyt, F. A.; Zhang, G. J., MCAM/CD146 promotes tamoxifen resistance in breast cancer cells through induction of epithelial-mesenchymal transition, decreased ERalpha expression and AKT activation. *Cancer Lett* **2017,** *386*, 65-76.
42. Shah, N.; Jin, K.; Cruz, L. A.; Park, S.; Sadik, H.; Cho, S.; Goswami, C. P.; Nakshatri, H.; Gupta, R.; Chang, H. Y.; Zhang, Z.; Cimino-Mathews, A.; Cope, L.; Umbricht, C.; Sukumar, S., HOXB13 mediates tamoxifen resistance and invasiveness in human breast cancer by suppressing ERalpha and inducing IL-6 expression. *Cancer Res* **2013,** *73* (17), 5449-58.
43. Kim, S.; Lee, J.; Oh, S. J.; Nam, S. J.; Lee, J. E., Differential effect of EGFR inhibitors on tamoxifen-resistant breast cancer cells. *Oncol Rep* **2015,** *34* (3), 1613-9.
44. Abba, M. C.; Hu, Y.; Sun, H.; Drake, J. A.; Gaddis, S.; Baggerly, K.; Sahin, A.; Aldaz, C. M., Gene expression signature of estrogen receptor alpha status in breast cancer. *BMC Genomics* **2005,** *6*, 37.
45. Jin, C.; Rajabi, H.; Pitroda, S.; Li, A.; Kharbanda, A.; Weichselbaum, R.; Kufe, D., Cooperative interaction between the MUC1-C oncoprotein and the Rab31 GTPase in estrogen receptor-positive breast cancer cells. *PLoS One* **2012,** *7* (7), e39432.

Figure 3.1 Targeted quantitative analysis of small GTPases in tamoxifen resistance.

(A) Schematic diagram showing the targeted quantitative proteomic analysis, relying on metabolic labeling with SILAC, SDS-PAGE fractionation, and scheduled LC-MRM analysis; **(B)** Heatmap showing the differential expression of small GTPases in MCF-7/WT and MCF-7/TamR cells. The log_2 ratios of relative levels of proteins in TamR over WT cells obtained from two forward and one reverse-SILAC labeling experiments (F1 and F2: forward experiments, R1: reverse experiment) are shown. As indicated by the scale bar, the red and blue bars designate the small GTPases that are up-and down-regulated, respectively, by at least 1.5-fold in the drug-resistant over parental MCF-7 cells; **(C)** Bar chart showing substantially up-regulated (>1.5-fold) and down-regulated (>1.5-fold) small GTPases quantified from three LC-MRM experiments.

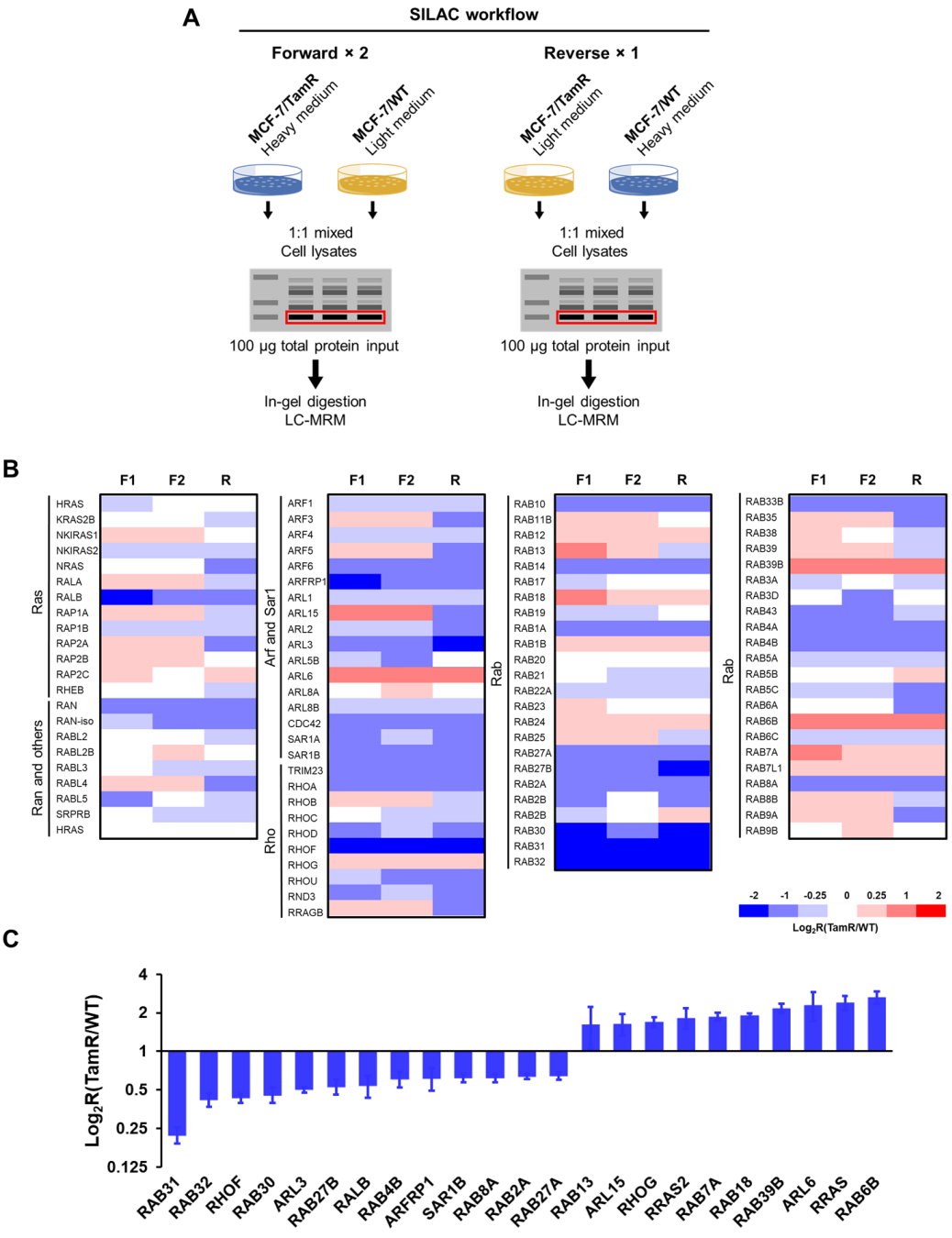

A
SILAC workflow
Forward × 2
Reverse × 1
MCF-7/TamR
Heavy medium
MCF-7/WT
Light medium
MCF-7/TamR
Light medium
MCF-7/WT
Heavy medium
1:1 mixed
Cell lysates
1:1 mixed
Cell lysates
100 µg total protein input
100 µg total protein input
In-gel digestion
LC-MRM
In-gel digestion
LC-MRM
B
F1
F2
R
Ras
HRAS
KRAS2B
NKIRAS1
NKIRAS2
NRAS
RALA
RALB
RAP1A
RAP1B
RAP2A
RAP2B
RAP2C
RHEB
Ran and others
RAN
RAN-iso
RABL2
RABL2B
RABL3
RABL4
RABL5
SRPRB
HRAS
Arf and Sar1
ARF1
ARF3
ARF4
ARF5
ARF6
ARFRP1
ARL1
ARL15
ARL2
ARL3
ARL5B
ARL6
ARL8A
ARL8B
CDC42
SAR1A
SAR1B
Rho
TRIM23
RHOA
RHOB
RHOC
RHOD
RHOF
RHOG
RHOU
RND3
RRAGB
Rab
RAB10
RAB11B
RAB12
RAB13
RAB14
RAB17
RAB18
RAB19
RAB1A
RAB1B
RAB20
RAB21
RAB22A
RAB23
RAB24
RAB25
RAB27A
RAB27B
RAB2A
RAB2B
RAB30
RAB31
RAB32
Rab
RAB33B
RAB35
RAB38
RAB39
RAB39B
RAB3A
RAB3D
RAB43
RAB4A
RAB4B
RAB5A
RAB5B
RAB5C
RAB6A
RAB6B
RAB6C
RAB7A
RAB7L1
RAB8A
RAB8B
RAB9A
RAB9B
-2
-1
-0.25
0
0.25
1
2
Log2R(TamR/WT)
C
Log2R(TamR/WT)
4
2
1
0.5
0.25
0.125
RAB31
RAB32
RHOF
RAB30
ARL3
RAB27B
RALB
RAB4B
ARFRP1
SAR1B
RAB8A
RAB2A
RAB27A
RAB13
ARL15
RHOG
RRAS2
RAB7A
RAB18
RAB39B
ARL6
RRAS
RAB6B

Figure 3.2 LC-MRM for the quantification of the relative levels of expressions of ARL3, RHOF, RAB30, and RRAS2 proteins in the paired MCF-7/WT and MCF-7/TamR cells.

(A) Representative extracted-ion chromatograms (XICs) for the quantification of the ARL3, RHOF, RAB30, and RRAS2 proteins in one forward- and one reverse-SILAC labeling experiments; **(B)** Quantification results of the LC-MRM analyses (n = 3). The *p* values were calculated by using a paired two-tailed Student's *t* test (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). Error bars represent standard deviations.

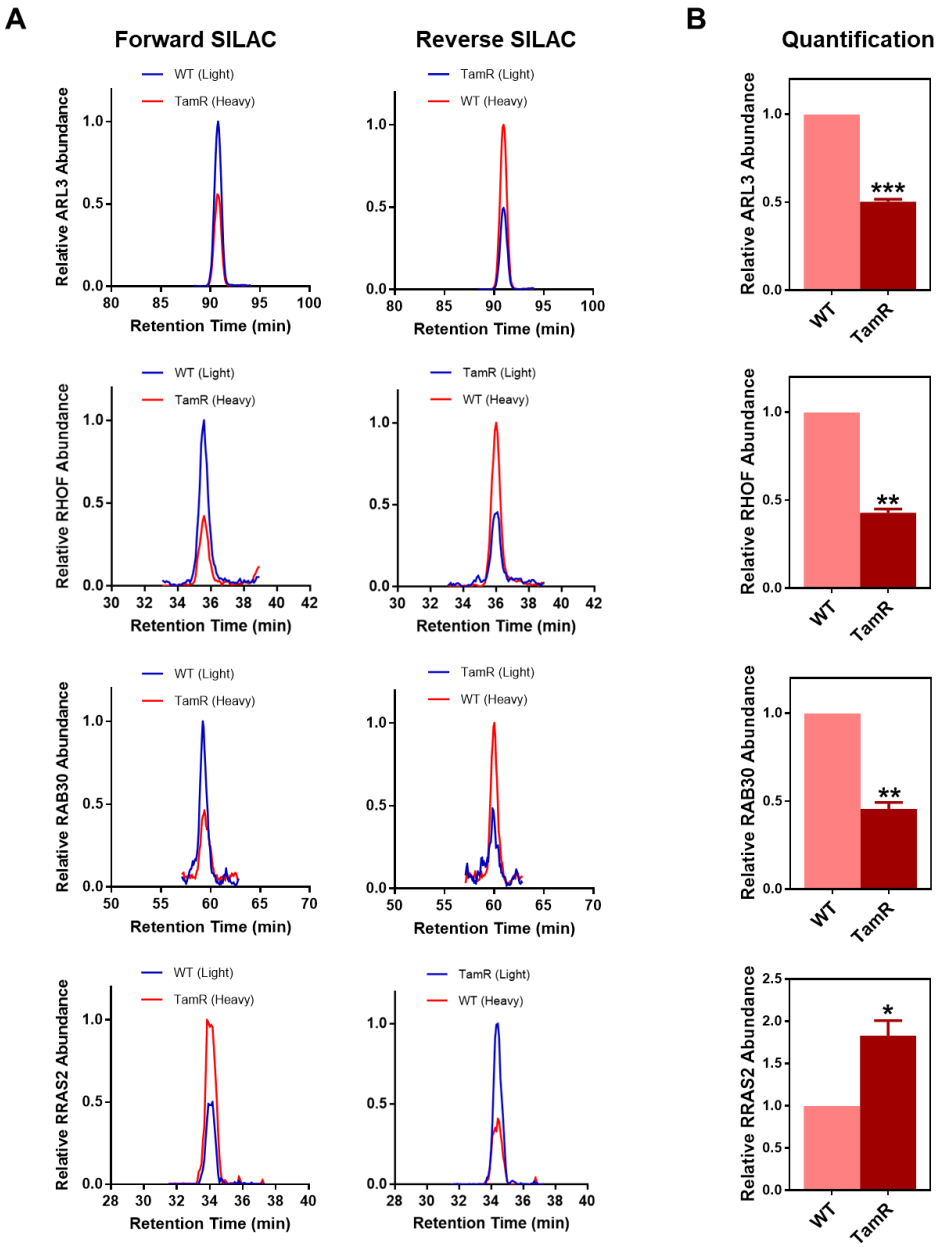


Figure 3.3 Performances of the scheduled LC-MRM method for targeted quantitative analysis of differential expression of small GTPases in tamoxifen resistance.

(A) Venn diagrams displaying the overlap between quantified small GTPases in the forward- and reverse-SILAC labeling experiments, as obtained from MRM analyses and DDA analyses, respectively, and the comparison about the performances of the two methods; **(B)** Correlation between the log_2-transformed SILAC ratios (Log_2R) obtained from one forward- and one reverse-SILAC labeling experiments with a relatively high correlation coefficient (R^2 = 0.9238); **(C)** Correlation between BSA standard iRT and measured RT values with a very high correlation coefficient (R^2 = 0.9996); **(D)** Measured RTs obtained in two forward labeling reactions (F1 vs. F2) with a very high correlation coefficient (R^2 = 0.9988); **(E)** Measured RTs obtained in one forward- and one reverse-SILAC labeling reactions (F1 vs. R) with a very high correlation coefficient (R^2 = 0.9987).

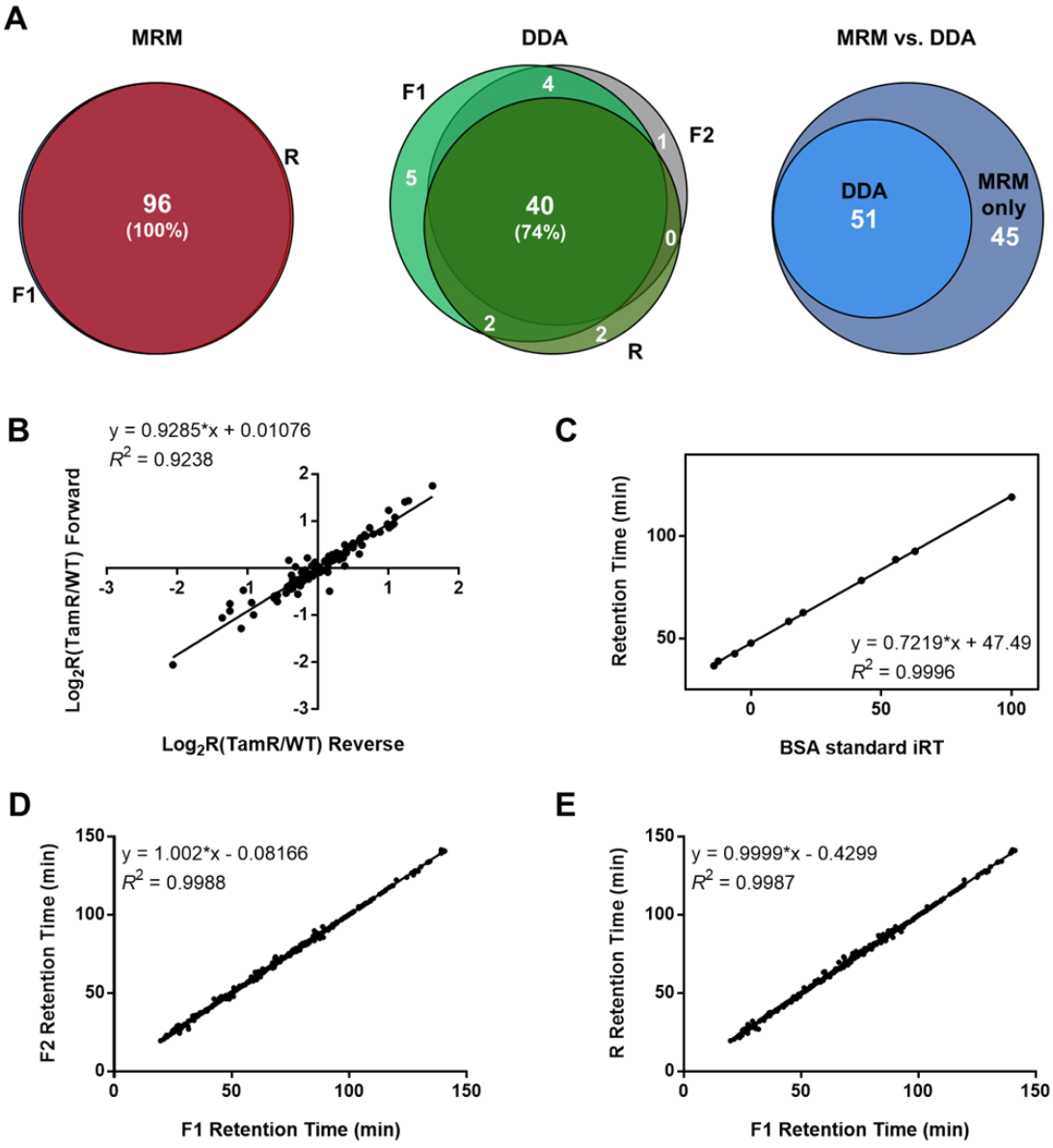

A
MRM
96
(100%)
R
F1
DDA
F1
4
5
40
(74%)
1
F2
0
2
2
R
MRM vs. DDA
DDA
51
MRM
only
45
B
y = 0.9285*x + 0.01076
R2 = 0.9238
Log2R(TamR/WT) Forward
Log2R(TamR/WT) Reverse
C
Retention Time (min)
y = 0.7219*x + 47.49
R2 = 0.9996
BSA standard iRT
D
y = 1.002*x - 0.08166
R2 = 0.9988
F2 Retention Time (mIn)
F1 Retention Time (min)
E
y = 0.9999*x - 0.4299
R2 = 0.9987
R Retention Time (min)
F1 Retention Time (min)

Figure 3.4 RAB31 is down-regulated in tamoxifen-resistant MCF-7 cells.

(A) Representative MRM traces for three transitions monitored for each of the three unique tryptic peptides derived from RAB31: FHSLAPMYYR (y_8, y_6, and y_5), QDSFYTLK (y_6, y_4, and y_3), and GSAAAVIVYDITK (y_7, y_6, and y_5); **(B)** Extracted-ion chromatograms (XICs) for the quantification of the three peptides in panel (A) in one forward- and one reverse-SILAC labeling experiments; **(C)** Validation of the differential expression of RAB31 in MCF-7/WT and MCF-7/TamR cells by Western-blot analysis; **(D)** Comparison of quantification results obtained from LC-MRM and Western-blot analyses (n = 3). The *p* values were calculated by using a paired two-tailed Student's *t* test (** $p < 0.01$). Error bars represent standard deviations.

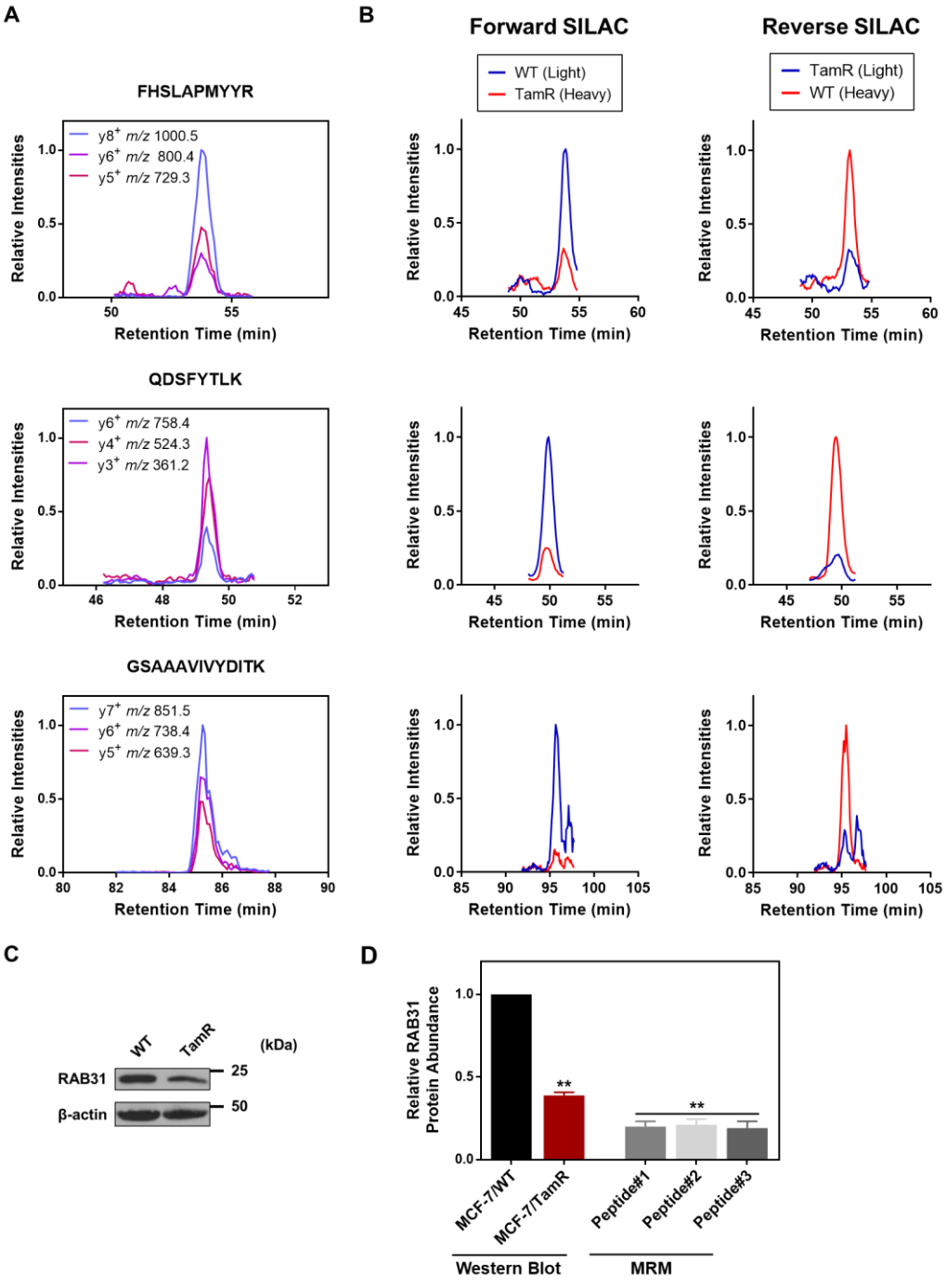

A
B
Forward SILAC
Reverse SILAC
FHSLAPMYYR
QDSFYTLK
GSAAAVIVYDITK
WT (Light)
TamR (Heavy)
TamR (Light)
WT (Heavy)
Relative Intensities
Retention Time (min)
C
WT
TamR
(kDa)
RAB31
β-actin
25
50
D
Relative RAB31 Protein Abundance
MCF-7/WT
MCF-7/TamR
Peptide#1
Peptide#2
Peptide#3
Western Blot
MRM

Figure 3.5 LC-MRM and Western-blot analyses for the quantification of the relative levels of expressions of RAB27A and RAB27B proteins in the paired MCF-7/WT and MCF-7/TamR cells.

(A) Representative XICs for the quantification of the tryptic peptide FLALGDSGVGK derived from RAB27A in one forward- and one reverse-SILAC labeling experiments; **(B)** Validation of the differential expression of RAB27A in MCF-7/WT and MCF-7/TamR cells by Western-blot analysis; **(C)** Comparison of quantification results obtained from LC-MRM and Western-blot analyses (n = 3) in panels (A) and (B); **(D)** Representative XICs for the quantification of the tryptic peptide LLALGDSGVGK derived from RAB27B in one forward- and one reverse-SILAC labeling experiments; **(E)** Validation of the differential expression of RAB27B in MCF-7/WT and MCF-7/TamR cells by Western-blot analysis; **(F)** Comparison of quantification results obtained from LC-MRM and Western-blot analyses (n = 3) in panels (D) and (E). The *p* values were calculated by using a paired two-tailed Student's *t* test (** $p < 0.01$). Error bars represent standard deviations.

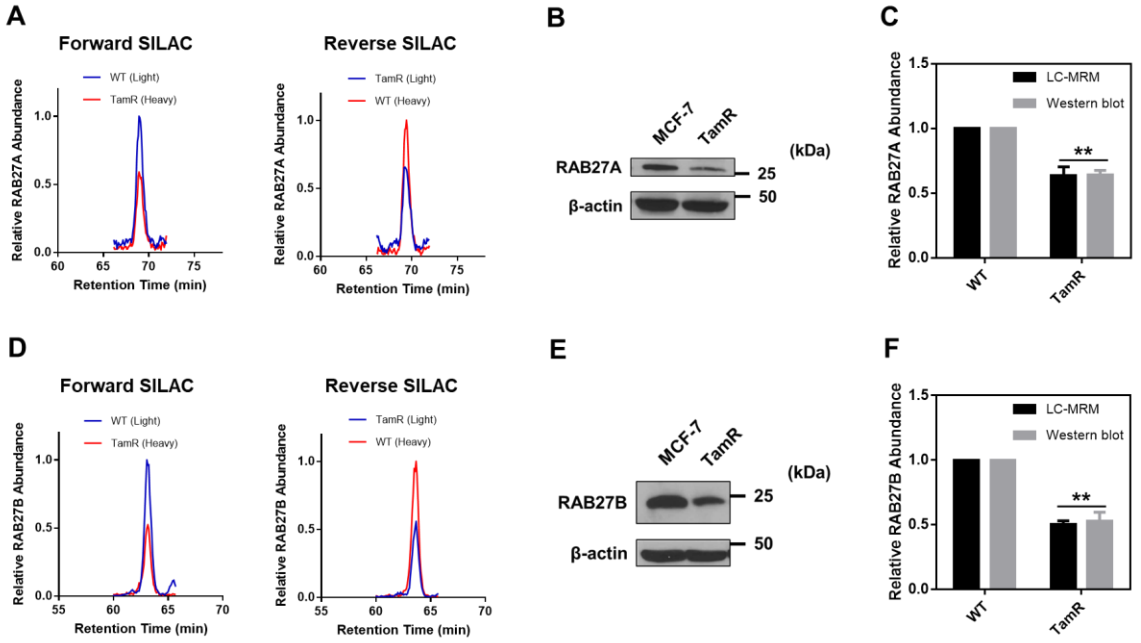


Figure 3.6 Representative Kaplan–Meier survival curves for the implications of *ARL3, RHOF, RAB30,* and *RRAS2* mRNA expressions in survival of breast cancer patients.

Kaplan–Meier survival analyses for relapse-free survival (RFS) in all breast cancer patients (left panel) and in ER-positive breast cancer patients receiving only treatment with tamoxifen but not chemotherapy (right panel). The patient population was stratified by median mRNA expression levels of the **(A)** *ARL3*, **(B)** *RAB30*, **(C)** *RHOF*, and **(D)** *RRAS2* genes, respectively. Analysis was performed using KM-plotter (kmplot.com/analysis). The *p* values were calculated by using a log-rank test. Number of patients, hazard ratios (HR) and 95% confidence intervals (95% CI) are indicated.

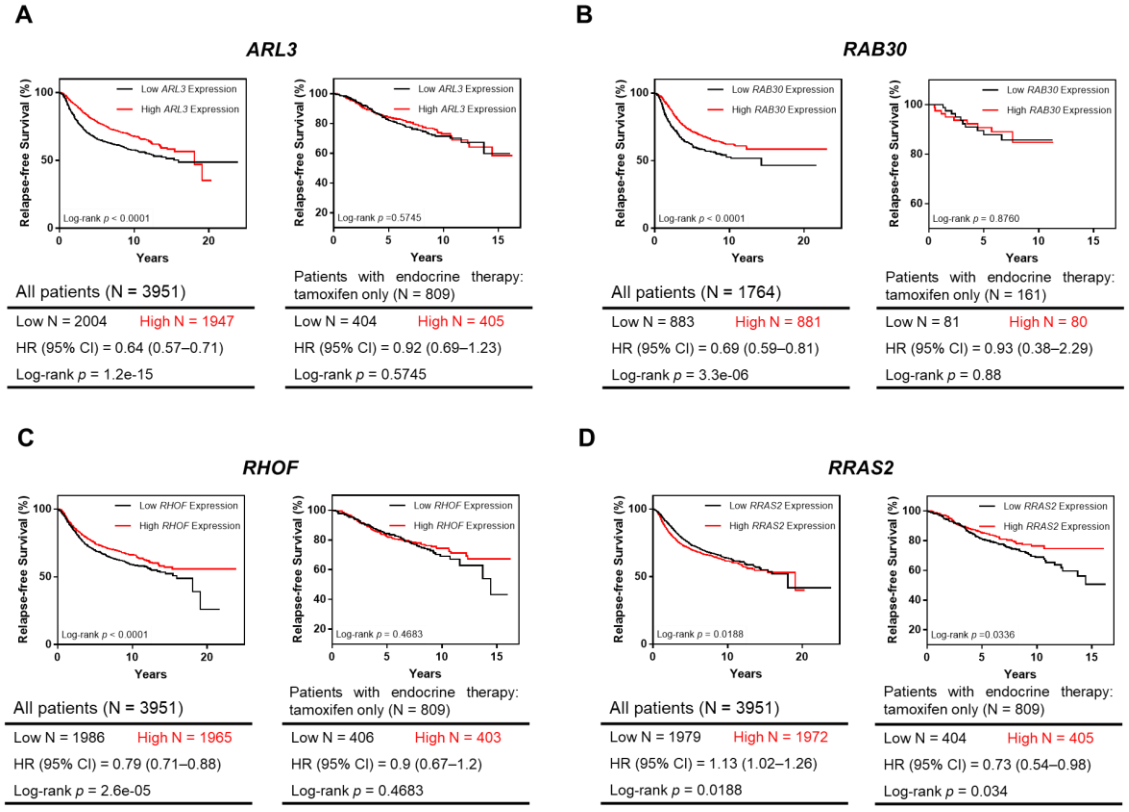


Figure 3.7 *RAB31* expression predicts breast cancer patient outcome.

Kaplan–Meier survival analyses of the implications of mRNA expression of the *RAB31* gene for relapse-free survival (RFS) in breast cancer patients: **(A)** all breast cancer patients (n = 3951); **(B)** ER-positive breast cancer patients receiving tamoxifen but without chemotherapy (n = 809); **(C)** ER-positive breast cancer patients (n = 2061); **(D)** ER-negative breast cancer patients (n = 801). The patient population was stratified by median *RAB31* mRNA expression levels. The *p* value was calculated by using a log-rank test. Number of patients, hazard ratios (HR) and 95% confidence intervals (95% CI) are indicated.

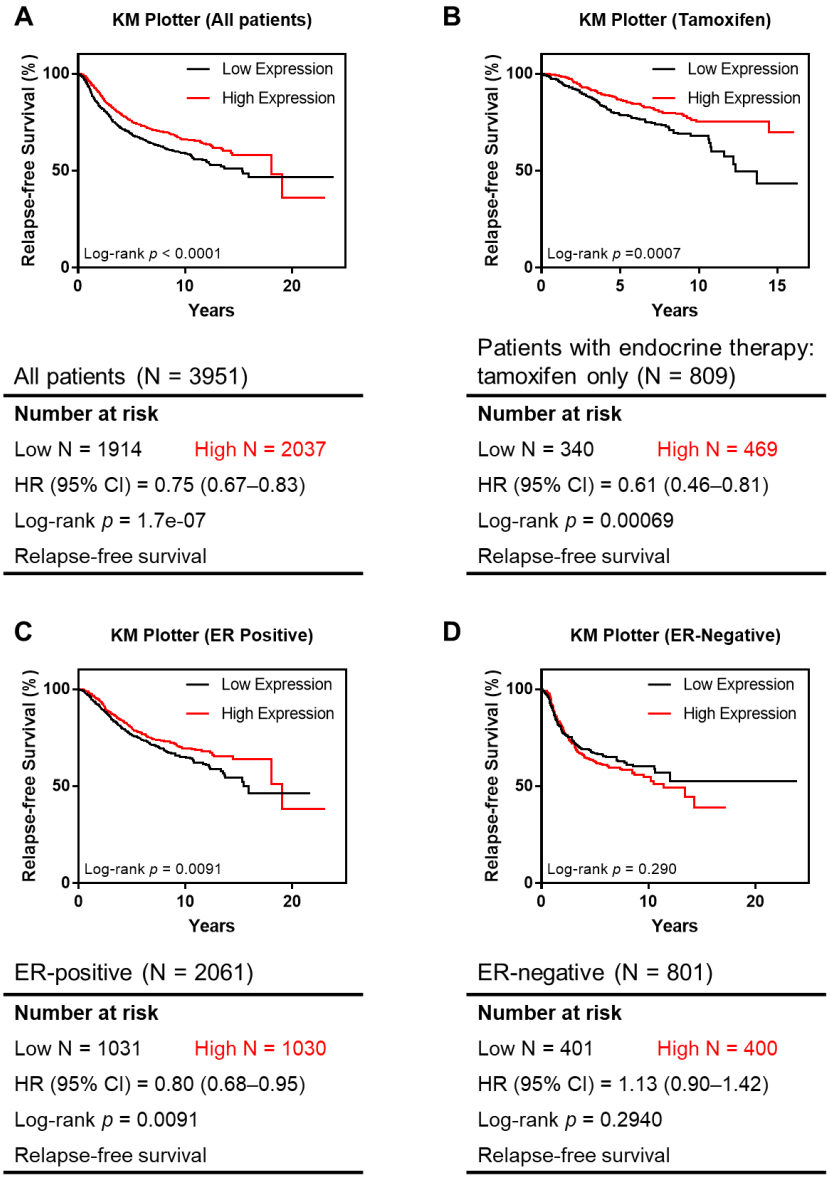


Figure 3.8 Bioinformatic analysis revealing RAB31 as a potential predictive marker for tamoxifen resistance.

Differential expression analysis for *RAB31* in: **(A)** docetaxel-, **(B)** tamoxifen- and **(C)** doxorubicin-resistant MCF-7 cell lines in three cohorts GSE6434, GSE26495 and GSE24460, respectively (Dox: docetaxel; Tam: tamoxifen; Adr: Adriamycin or doxorubicin). Kaplan–Meier survival analyses for: **(D)** relapse-free survival (RFS) in the GSE42568 data set; **(E)** overall survival (OS) in the GSE42568 data set; **(F)** disease-specific survival (DSS) in the GSE3494 data set; **(G)** differential *RAB31* expression in different grades of breast cancer progression in the GSE4922 data set. The patient population was stratified by median *RAB31* mRNA expression levels. The *p* values for the Kaplan–Meier curves were calculated by using the log-rank test. The *p* values for the box-whisker plots were calculated by using an unpaired two-tailed Student's *t* test (# $p > 0.05$, * $p < 0.05$, ** $p < 0.01$).

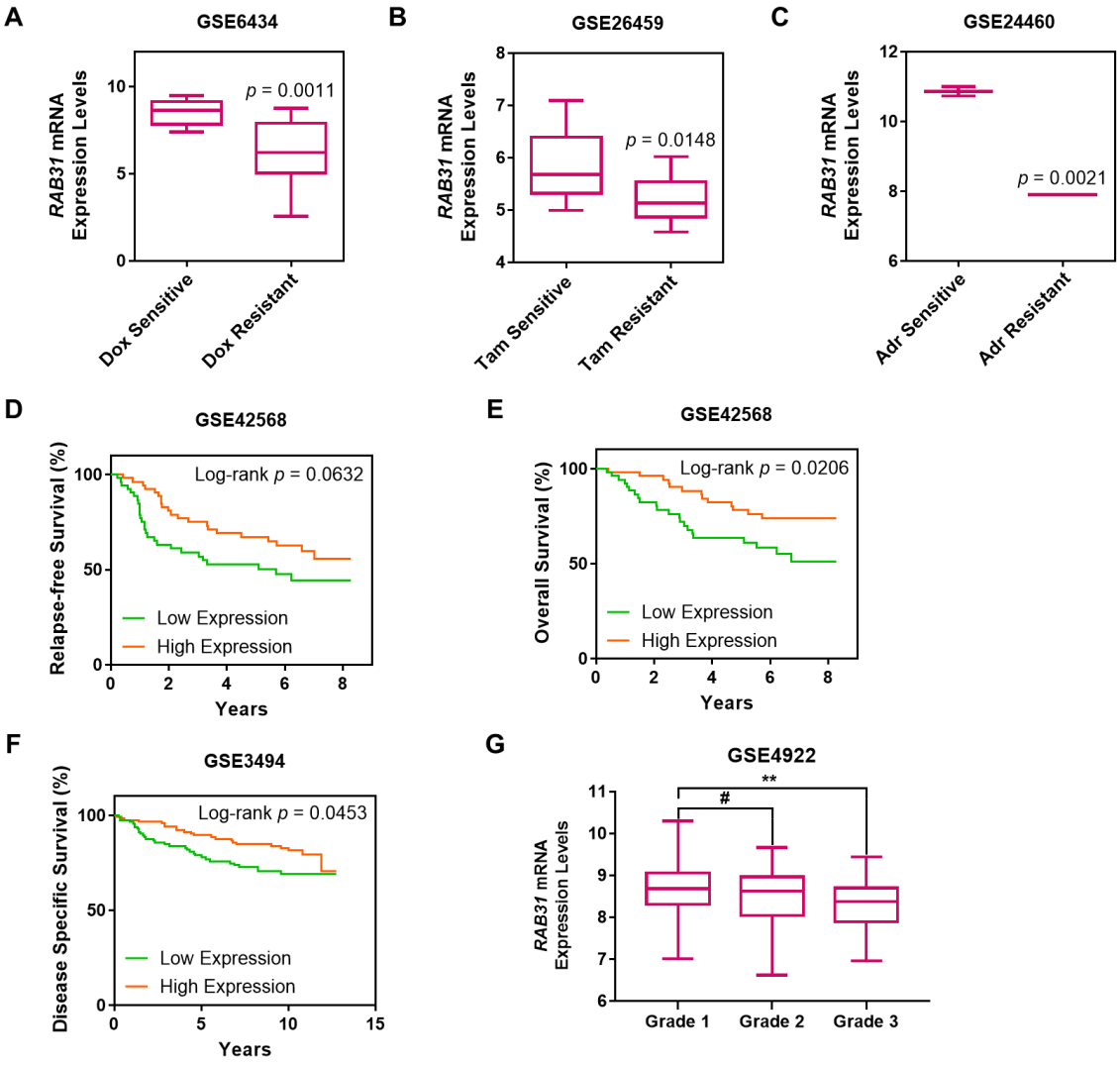

A
GSE6434
RAB31 mRNA Expression Levels
p = 0.0011
Dox Sensitive
Dox Resistant
B
GSE26459
RAB31 mRNA Expression Levels
p = 0.0148
Tam Sensitive
Tam Resistant
C
GSE24460
RAB31 mRNA Expression Levels
p = 0.0021
Adr Sensitive
Adr Resistant
D
GSE42568
Relapse-free Survival (%)
Log-rank p = 0.0632
Low Expression
High Expression
Years
E
GSE42568
Overall Survival (%)
Log-rank p = 0.0206
Low Expression
High Expression
Years
F
GSE3494
Disease Specific Survival (%)
Log-rank p = 0.0453
Low Expression
High Expression
Years
G
GSE4922
RAB31 mRNA Expression Levels
**
#
Grade 1
Grade 2
Grade 3

Figure 3.9 *RAB31* expression is correlated with ER status and breast cancer subtypes.

Box-whisker plot showing correlated expressions of *RAB31* and *ESR1*, the latter of which encodes estrogen receptor alpha (ERα), in the: **(A)** TCGA-BRCA cohort; **(B)** GSE4922 data set; **(C)** GSE23988 data set. Whisker shows 10th and 90th percentile; box boundaries show 25th and 75th percentile and the lines represent median values. **(D)** A scatter plot showing *RAB31* expression across the five molecular subtypes of breast cancer (n = 831) in the TCGA-BRCA cohort, and the lines represent median value. The *p* values for the box-whisker plots were calculated by using an unpaired two-tailed Student's *t* test.

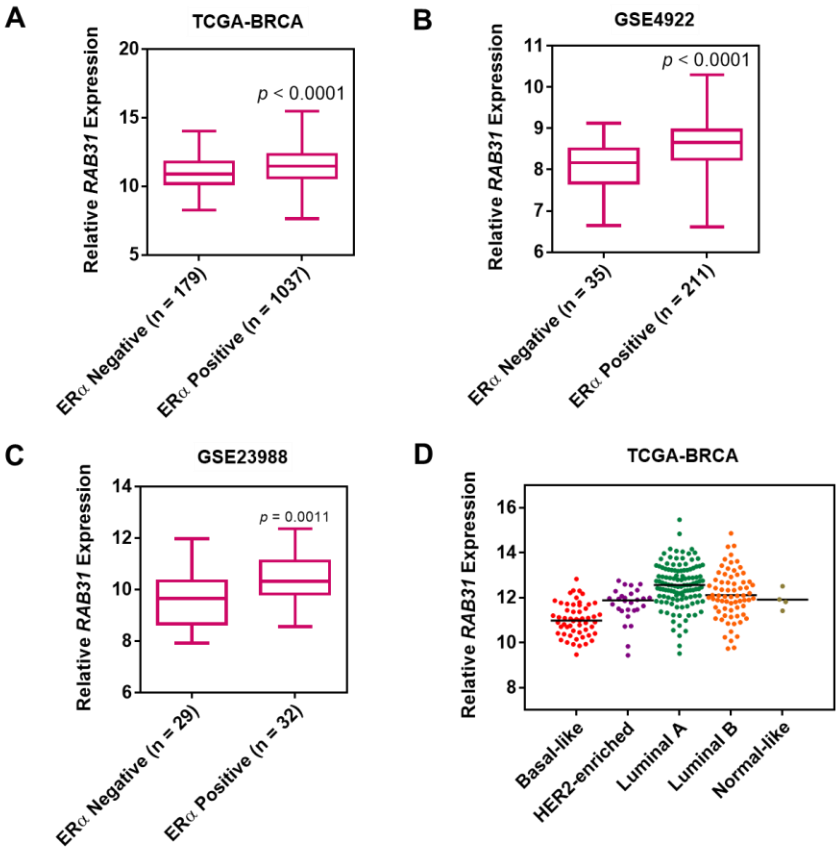


Figure 3.10 RAB31 knockdown confers increased tamoxifen resistance.

(A) Validation of stable knockdown of RAB31 in MCF-7 cells by Western-blot analysis; **(B)** Quantification results for Western-blot validation of stable shRAB31 knockdown MCF-7 cell lines; **(C)** Colony formation assay for stable shRAB31/shScramble MCF-7 cells; **(D)** Quantification results for colony formation assay; **(E)** Cell proliferation assay for stable shRAB31/shScramble MCF-7 cells upon treatment with different doses of 4-OHT. The *p* values were calculated by using a paired two-tailed Student's *t* test (** $p < 0.01$). Error bars represent standard deviations.

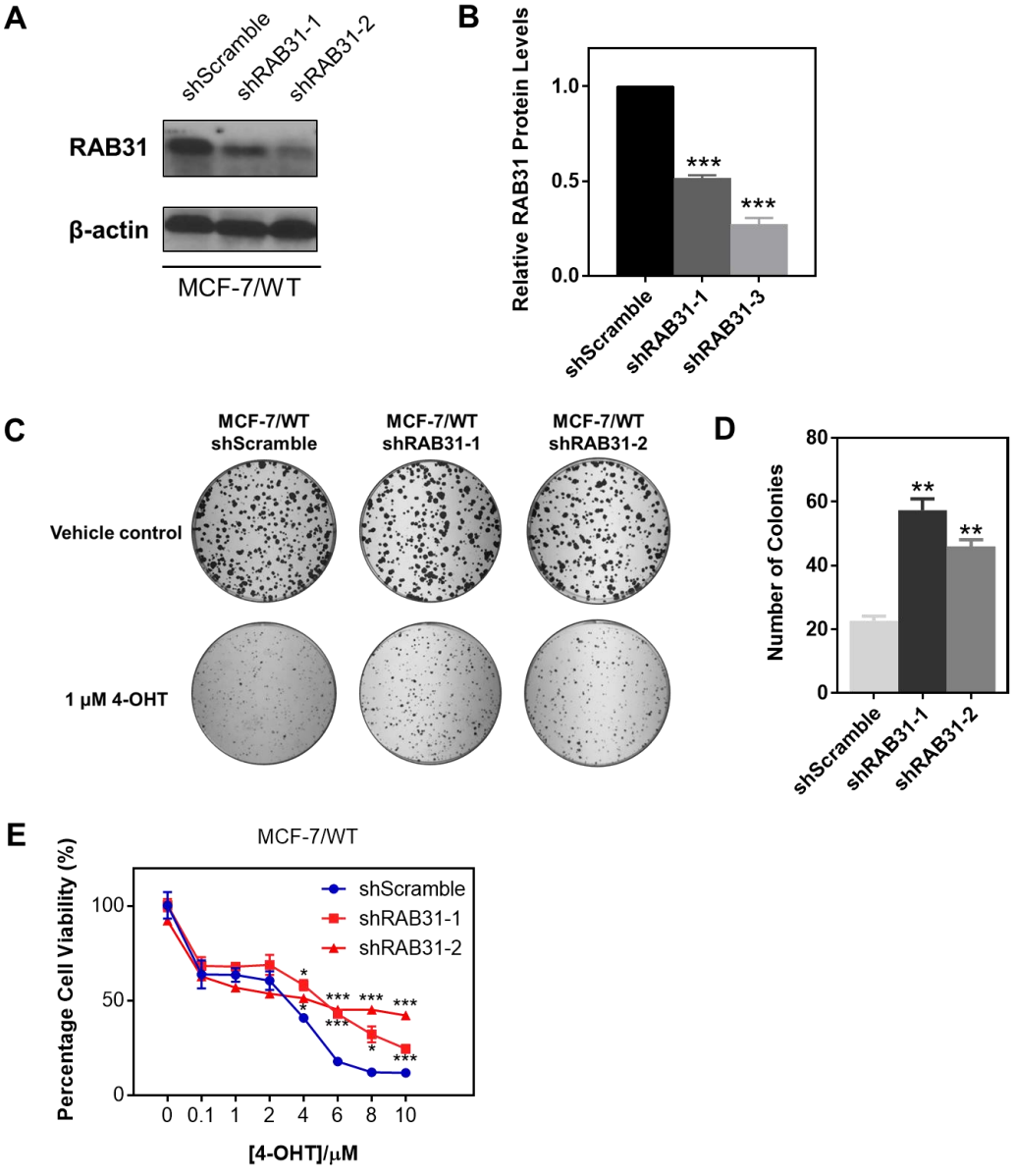


Figure 3.11 RAB31 knockdown renders elevated tamoxifen resistance and proliferation rates.

(A) Cell proliferation assay for stable shRAB31-2 and shScramble MCF-7 cells; **(B)** Validation of stable knockdown of RAB31 in T47D cell lines by Western-blot analysis; **(C)** Cell proliferation assay for stable shRAB31-2 and shScramble T47D cells; **(D)** Cell proliferation assay for stable shRAB31-2 and shScramble T47D cells under dose-dependent 4-OHT treatment; **(E)** Cell proliferation assay for stable shRAB31-2 and shScramble T47D cells under time-dependent 1 μM 4-OHT treatment. The data represent the means ± standard deviations of results from three parallel experiments. Some error bars appear to be smaller than the symbols.

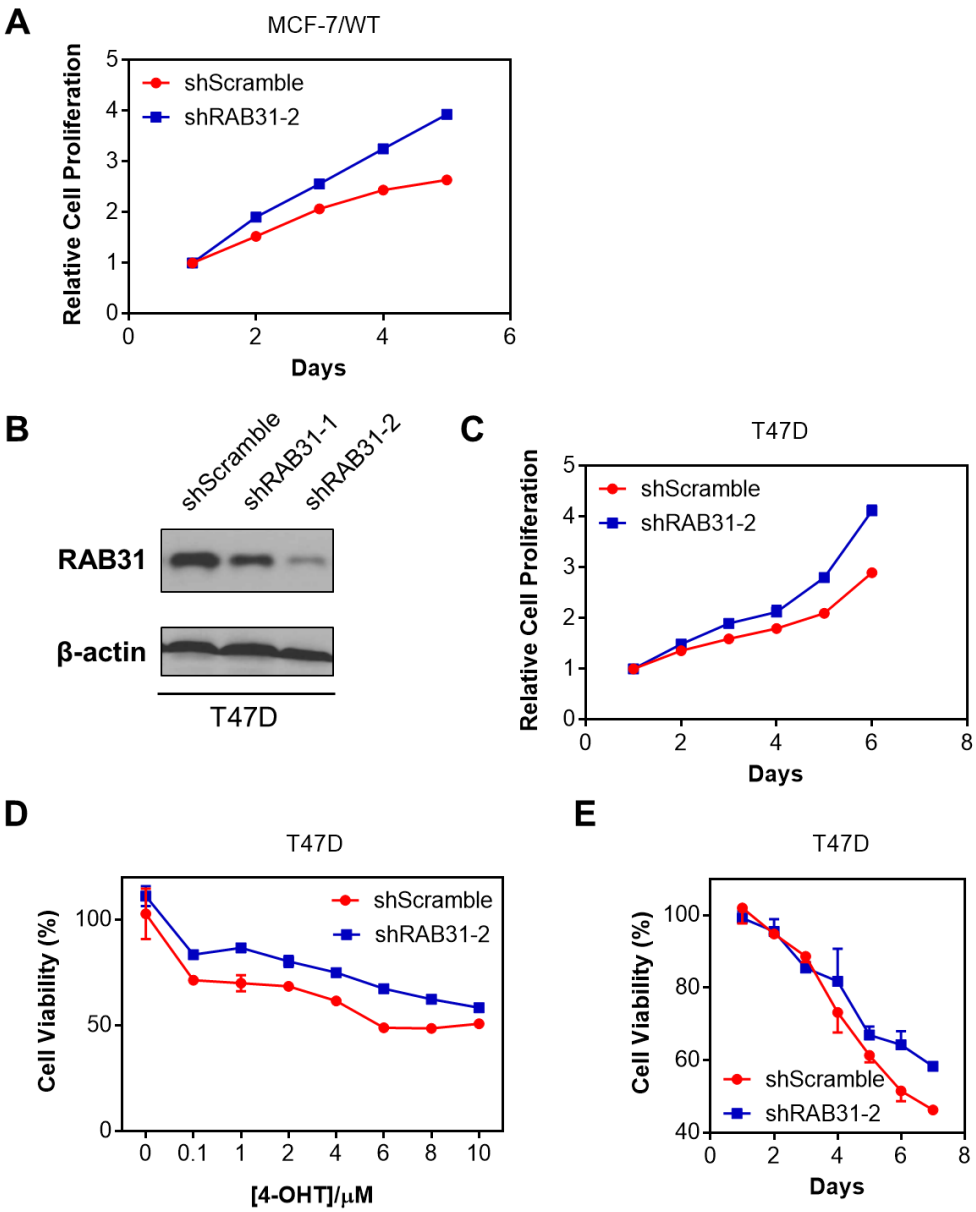
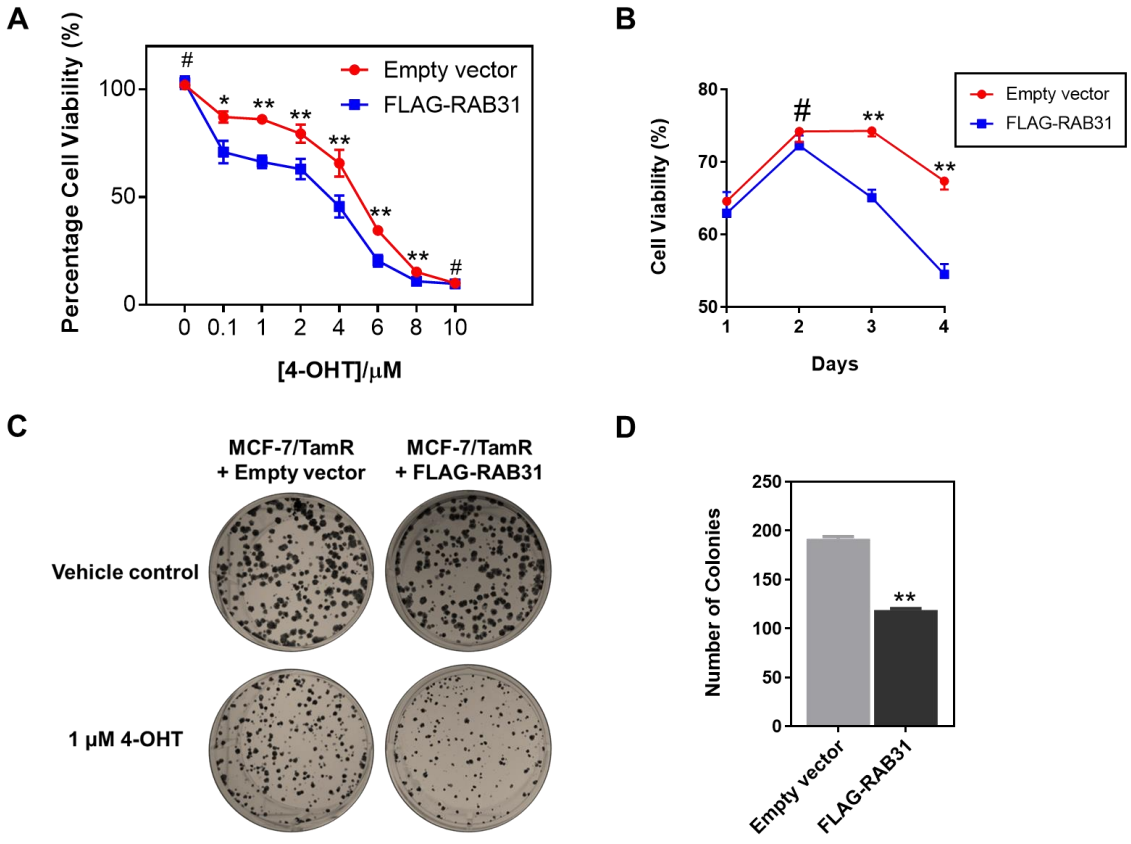

A
MCF-7/WT
Relative Cell Proliferation
shScramble
shRAB31-2
Days
B
shScramble
shRAB31-1
shRAB31-2
RAB31
β-actin
T47D
C
T47D
Relative Cell Proliferation
shScramble
shRAB31-2
Days
D
T47D
Cell Viability (%)
shScramble
shRAB31-2
[4-OHT]/μM
E
T47D
Cell Viability (%)
shScramble
shRAB31-2
Days

Figure 3.12 Ectopic expression of RAB31 led to elevated tamoxifen sensitivity.

(A) Cell proliferation assay for MCF-7/TamR cells after treatment with different doses of 4-OHT; **(B)** Cell proliferation assay for MCF-7/TamR cells after treatment with 4-OHT (1 μM) for different periods of time; **(C)** Colony formation assay for MCF-7/TamR cells expressing empty vector or FLAG-RAB31; **(D)** Quantification results for colony formation assay. The *p* values were calculated by using a paired two-tailed Student's *t* test (# $p > 0.05$, * $p < 0.05$, ** $p < 0.01$). Error bars represent standard errors of the means.



Chapter 4 Targeted Quantitative Proteomic Approach for Probing Altered Protein Expression of Small GTPases Associated with Colorectal Cancer

4.1 Introduction

Colorectal cancer (CRC) is a major cause of cancer-associated deaths worldwide, and it ranks third in terms of incidence but second in terms of mortality.[1-2] According to the stage definition by the American Joint Committee on Cancer (AJCC), five-year survival rates ranged from 93.2% for stage I to 8.1% for stage IV CRC patients.[3] Approximately 40–50% of all patients with CRC will ultimately develop into metastatic disease either at the time of diagnosis or develop distant relapses after therapy, of which the median overall survival is less than two years.[4] Therefore, a better understanding about the metastatic transformation of CRC cells could lead to more precise therapeutic strategies that improve patient outcomes.

The Ras superfamily of small (20–35 kDa) GTPases function as molecular switches that play important roles in carcinogenesis and tumor progression by regulating intracellular trafficking, cell signaling and malignant secretion.[5-6] Among them, *KRAS* is one of the most frequently mutated oncogenes that drives colorectal tumorigenesis, accounting for 40% of sporadic CRCs. In addition to mutations, aberrant regulation of small GTPase expression has been shown to have a role in CRC progression. For instance, overexpression of RAB3C promoted colorectal tumor progression *in vitro* and *vivo*.[7] Moreover, miR-27b and miR-204-5p suppressed proliferation and invasion of CRC cells by inhibiting RAB3D and RAB22A, respectively[8-9], and down-regulation of RAB25 was

found to be associated with reduced metastatic potential of CRC cells *in vivo*.[10] Therefore, we reason that a systematic investigation about small GTPases associated with metastatic transformation of CRC cells may lead to novel molecular targets for the improvement of therapeutic efficacy. In this study, we applied a recently established high-throughput, targeted proteomic approach to robustly quantify the differential expression of small GTPases in the SW480/SW620 matched primary/metastatic cells derived from the same patient.[11] We also examined, by using bioinformatic analyses, whether the aberrant expression of these small GTPases were dysregulated in clinical samples and showed prognostic values. Finally, we identified SAR1B as a potential metastasis suppressor and a prognostic biomarker that could modulate *in vitro* migration and invasion of CRC cells by regulating epithelial–mesenchymal transition (EMT).

4.2 Materials and Methods

4.2.1 Cell Culture

The CRC cell lines used in this study were SW480 (ATCC# CCL-228) and SW620 (ATCC# CCL-227), which were established from a primary human colon adenocarcinoma and a lymph node metastasis of the same patient, respectively. All cell lines were maintained in Dulbecco's Modified Eagle Medium (DMEM; Invitrogen-Gibco) supplemented with 10% fetal bovine serum (FBS; Invitrogen-Gibco) and penicillin/streptomycin (PS, 100 IU/mL) in a humidified atmosphere with 5% CO_2 at 37 °C, and the culture medium was changed in every 2 to 3 days as necessary.

For stable isotope labeling by amino acids in cell culture (SILAC) experiments, “heavy” and “light” SILAC DMEM media depleted of $_L$-lysine and $_L$-arginine (Thermo

Scientific™ Pierce) were freshly prepared by adding 0.146 g/L $^{13}C_6$ $^{15}N_2$ L-lysine (Lys-8) and 0.84 g/L $^{13}C_6$ L-arginine (Arg-6) (Cambridge Isotopes Inc.) or the corresponding un-labeled lysine (Lys-0) and arginine (Arg-0) to DMEM, which as supplemented with 10% dialyzed FBS (Corning) and PS (100 IU/mL). SW480 and SW620 cells were cultured in the heavy-DMEM medium for at least 10 days or six cell doublings to assure complete heavy-isotope incorporation.

4.2.2 Sample Preparation and LC-MRM Analysis

To assess the differential expression of small GTPases in SW480 and SW620 cells, we conducted SILAC-based quantitative proteomic experiments with forward and reverse labeling strategies. Briefly, we combined lysates of light-labeled SW480 cells and heavy-labeled SW620 cells at a 1:1 ratio in the forward labeling experiments. The reverse labeling experiment was conducted in the opposite way. The mixed cell lysates (100 μg in total) were loaded onto a 10% SDS-PAGE gel and separated by electrophoresis. The gel bands corresponding to the molecular weight range of 15–37 kDa were cut, reduced with 20 mM dithiothreitol, alkylated with 55 mM iodoacetamide, and digested in-gel with trypsin at an enzyme/protein ratio of 1:100. After tryptic digestion, the peptide mixtures were desalted and subjected to LC-MRM analyses on a TSQ Vantage triple-quadruple mass spectrometer, which was equipped with a nanoelectrospray ionization source and coupled with an EASY-nLC II HPLC system (Thermo Fisher Scientific, San Jose, CA). The samples were automatically loaded onto a 4-cm trapping column (150 μm i.d.) packed with ReproSil-Pur 120 C18-AQ resin (5 μm in particle size and 120 Å in pore size, Dr. Maisch GmbH HPLC) at 3 μL/min. The trapping column was connected to a 20-cm fused silica analytical column

(75 μm i.d.) packed with ReproSil-Pur 120 C18-AQ resin (3 μm in particle size and 120 Å in pore size, Dr. Maisch GmbH HPLC). A 157-min linear gradient of 2−35% acetonitrile in 0.1% formic acid was employed for peptide separation, and the flow rate was 230 nL/min. The spray voltage was set at 1.8 kV. Ions were isolated in both Q1 and Q3 using 0.7 fwhm resolution, where the cycle time was set as 5 s. The optimal collisional energy for each targeted peptide was calculated using a linear equation specific to the TSQ Vantage instrument and the precursor mass-to-charge ratio (*m/z*), according to the default setting in Skyline.[12]

To enable high-throughput quantitative analysis, we applied a previously developed scheduled LC-MRM method, where the mass spectrometer was programmed to acquire the MS/MS of the precursor ions for a limited number of peptides in each 6-min retention time (RT) window.[11] The MRM data for all targeted peptides were manually inspected to ensure correct peak picking. In this regard, the dot-plot or dot-product (dotp) value has to exceed 0.80.[13] In addition, the iRT value represents an intrinsic property (i.e., hydrophobicity) of a peptide; hence, a substantial deviation of measured RT from that projected from the linear plot of RT over iRT signals a false-positive detection.[14]

4.2.3 Immunoblotting

Total protein lysate was prepared from cell pellet using ice-cold CelLytic M cell lysis reagent (Sigma-Aldrich, MO) containing protease inhibitor cocktail (1:100). After cell lysis, the protein concentration was determined using Quick Start™ Bradford Protein Assay (Bio-Rad). Approximately 10–20 μg whole-cell protein lysates, mixed with 4× Laemmli SDS loading buffer, were loaded onto 10% polyacrylamide gels and, after

electrophoresis, the proteins were transferred onto nitrocellulose membranes. After blocking with 5% non-fat milk in PBS with 0.1% Tween-20 (PBST) at 25 °C for 1 h, the membranes were incubated with primary antibodies against human ARF4 (Proteintech; rabbit polyclonal, 1:5,000), RAB6A (38-TB, Santa Cruz; mouse polyclonal, 1:1,000), pan-RAB6 (3G3, Santa Cruz; mouse polyclonal, 1:1,000), RAB27A (Abcam; rabbit polyclonal, 1:5,000), RAB31 (4D12, Santa Cruz; rabbit polyclonal, 1:2,000), SAR1A (K-44, Santa Cruz; mouse polyclonal, 1:1,000), SAR1B (AT1C7, Santa Cruz; mouse polyclonal, 1:1,000), Slug (PA5-11922, Thermo Fisher; rabbit polyclonal, 1:1,000), or β-actin (Thermo Fisher; rabbit polyclonal, 1:10,000). After overnight incubation with primary antibodies at 4°C with 5% bovine serum albumin (BSA) in PBST, the membranes were then incubated with peroxidase-labeled donkey anti-rabbit secondary antibody (Thermo Fisher; 1:10,000) or mouse m-IgGκ BP-HRP (Santa Cruz; 1:10,000) for 1 h at 25 °C. Amersham ECL Prime Western Blot Detecting Reagent (GE Healthcare) was used for visualization of protein bands.

4.2.4 Data Sources for Bioinformatic Analyses

Patient RNA-Seq data were obtained from The Cancer Genome Atlas (TCGA) via cBioPortal (http://www.cbioportal.org/). The data of four different human expression microarrays with accession numbers of GSE14333, GSE17536, GSE21510, GSE39582, and the corresponding clinical information were downloaded from the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) (http://www.ncbi.nlm.nih.gov/geo/). The data were analyzed using R/Bioconductor (version 3.4.3).

4.2.5 Generation of Stable Knockdown Cell Lines and siRNA Transfection

The lentiviral vectors pLKO.1-shSAR1B and pLKO.1-scramble plasmids were constructed by inserting a short hairpin double-stranded oligonucleotide targeting human SAR1B or a scrambled sequence into the AgeI/EcoR1 sites of the shRNA vector pLKO.1 (Addgene #10878; Cambridge, MA). Expression of the shRNA was driven by the human U6 promoter. Recombinant lentiviruses were produced by co-transfection of HEK293T cells with the shRNA plasmids pLKO.1-shScramble or pLKO.1-shSAR1B, envelope plasmid pLTR-G (Addgene #17532) and packaging plasmid pCMV-dR8.2 dvpr (Addgene #8455). Lentivirus-containing supernatant was harvested and filtered through 0.45-μm pore size filters at 48 h post-transfection. Infection of SW480 cells with recombinant lentivirus was conducted in the presence of 5 μg/ml polybrene. After removal of virus, the cells were selected in 5 μg/mL puromycin-containing medium for 3 days to eliminate uninfected cells. After selection, the cells were maintained in a medium containing 2.5 μg/mL puromycin and were used for subsequent experiments. For siRNA transfection, siGENOME non-targeting (NT) siRNA control (D-001210-02-05) and an siRNA sequence (5′-UGAUGUUGUGGUCAAAGUGAU-3′) targeting RAB31 were purchased from Dharmacon (Lafayette, CO). siRNAs were transfected into SW480 cells using RNAiMAX (Invitrogen) following the manufacturer's protocol. Cells were either harvested or used for subsequent experiments at 72 h after transfection.

4.2.6 Cell Proliferation Assay

The cell proliferation was evaluated using a cell counting kit-8 (CCK-8; Dojindo Laboratories). Briefly, SW480 cells were seeded in a 96-well flat-bottomed microplate

(3000 cells/well) in complete growth medium (100 μL/well) for 24 h. At the desired time points, 10 μL of the CCK-8 dye was added to each well and, after incubation at 37°C for 4 h, the absorbance at 450 nm for the cells was recorded using the Synergy™ H1 Hybrid Multi-Mode Microplate Reader (BioTek Instruments).

4.2.7 Migration and Invasion Assays

Transwell chambers (Corning) were rehydrated at 37 °C and 1×10^5 cells were then added to the top chamber in serum-free medium and the bottom chamber was filled with medium containing 10% fetal bovine serum. The invasion assay was conducted under the same conditions except that the transwell membranes were pre-coated with Matrigel (Corning). The cells were cultured for 48 h at 37 °C in a 5% CO_2 humidified incubator. To quantify migrated or invaded cells, the cells from the top-side of the membrane were gently removed using a cotton-tipped swab and invading cells attached to bottom of the membrane were fixed with 70% ethanol and stained with 0.5% crystal violet. Cell numbers from 5 representative fields were counted for each insert.

4.2.8 Quantitative Reverse Transcription PCR (RT-qPCR)

Total RNA was extracted from cells using The E.Z.N.A.® Total RNA Kit I (Omega Bio-Tech), and cDNA was synthesized via oligo$(dT)_{18}$ primed reverse transcription by employing M-MLV reverse transcriptase (Promega). After a 60-min reaction at 42°C, the reverse transcriptase was deactivated by heating at 75°C for 5 min. RT-qPCR experiments was performed using iQ SYBR Green Supermix kit (Bio-Rad) on a Bio-Rad iCycler system (Bio-Rad), and the PCR conditions were as follows: 95°C for 3 min; 45 cycles at 95°C for 15 s, 55°C for 30 s, and 72°C for 45 s. The comparative cycle threshold (Ct) method (ΔΔCt)

was used for the relative quantification of gene expression.[15] Relative gene expression was normalized to that of the internal control (*GAPDH*).

4.3 Results

4.3.1 Targeted Quantitative Profiling of Differential Expression of Small GTPases during Metastatic Transformation of CRC Cells

The SW480/SW620 pair of isogenic, primary/metastatic CRC cell lines constitute a valuable *in vitro* cellular model for studying CRC metastasis. The SW620 cell line was isolated from the lymph node metastatic site of the same patient as its non-metastatic counterpart (i.e. SW480).[16] This pair of cell lines have been widely explored with various comparative shotgun and targeted proteomics approaches at the whole proteome and secretome levels to discover potential biomarkers and therapeutic targets for CRC.[17-22]

To systematically investigate the differential protein expression of small GTPases in the paired SW480/SW620 cells, we employed a previously established targeted proteomic workflow, which involves SILAC, SDS-PAGE fractionation, and scheduled multiple-reaction monitoring (MRM) analysis.[11] To obtain reliable quantification results, we carried out SILAC experiments in triplicate, with two sets of forward labeling (Figure 4.1) and one set of reverse labeling. In this vein, the throughput for the method is high, where the entire library of small GTPase tryptic peptides could be monitored in two LC-MRM runs with retention time scheduling by using the iRT algorithm.[14] We also examined the run-to-run reproducibility of the MRM-based quantification. As illustrated in Figure 4.1, the $\log_2$-transformed SILAC ratios for all the quantified small GTPases obtained from forward and reverse SILAC labeling experiments exhibited an excellent linear fit ($R^2 = 0.928$). MRM

analyses facilitated reproducible quantification of 83 small GTPases in each of the three SILAC labeling experiments (Figure 4.2). By contrast, analysis of the same samples by LC-MS/MS in the data-dependent acquisition (DDA) mode only led to the identification of 37 and 35 small GTPases in the two forward SILAC samples (F1 and F2, respectively), and 31 small GTPases in the reverse SILAC sample (R) (Figure 4.1). Moreover, MRM exhibited better sensitivity than the shotgun proteomic approach, as reflected by the substantially increased coverage of the small GTPase proteome in two LC-MRM runs without further sample pre-fractionation (Figure 4.1).

We also analyzed the previously published shotgun or targeted proteomic data acquired from the SW480/SW620 paired cell lines. Figure 4.3 showed the Venn diagrams representing small GTPases quantified in this study as compared to those reported by three independent proteomic studies.[18, 20-21] The results clearly revealed that SDS-PAGE fractionation followed by scheduled MRM analysis allowed for a more in-depth coverage of the small GTPase proteome. It is of note that the method reported by Cai *et al*.[21] , which utilizes the acyl-phosphate GTP affinity probe, enabled proteome-wide enrichment of GTP-binding proteins, not solely restricted to small GTPases.

The above results demonstrated that the scheduled LC-MRM approach facilitates highly sensitive and reproducible quantitative profiling of small GTPases in CRC cell lines with an elevated throughput and depth of coverage.

4.3.2 Validation of Differential Protein Expression by Western-blot Analyses

Among all the quantified small GTPases, 18 and 7 proteins were significantly down- and up-regulated (with >1.5-fold change), respectively, in the metastatic SW620 cells

relative to primary SW480 cells (Figure 4.1). In this vein, we employed 1.5-fold change as a cutoff based on the mean relative standard deviation (RSD = 6.7%) for all the quantified small GTPases. Notably, several Rho small GTPases exhibited significantly altered expression, including RHOB, RHOF, and RHOG. Figure 4.4 displayed the selected-ion chromatograms (SICs) for these small GTPases. Down-regulation of RHOB and up-regulation of RHOG in SW620 relative to SW480 cells were previously reported in two independent studies.[21, 23] Importantly, RHOB is a well-recognized tumor suppressor for CRC and clear-cell renal cell carcinoma (ccRCC).[23-25] Since the protein abundance of RHOB was down-regulated by ~2-fold in SW620 over SW480 cells, our results validated a known suppressor for CRC metastasis.

Among the small GTPases with altered protein expression, RAB6B displayed a pronouncedly decreased expression (by ~10-fold) in SW620 relative to SW480 cells (Figure 4.1). Owing to the high degree of sequence homology among the three RAB6 isoforms (RAB6A/B/C), common antibody-based approach including Western blot may be susceptible to cross-reactivity and hence could be challenging in distinguishing different isoforms. By selecting proteotypic peptides which have distinct sequences for different protein isoforms, our MRM-based approach successfully revealed the substantial down-regulation of RAB6B, but not RAB6A/C, in SW620 relative to SW480 cells (Figure 4.4). Meanwhile, Western-blot analyses validated the MRM data for RAB6A and pan-RAB6 proteins (RAB6A/B/C) but failed to distinguish RAB6A/C from RAB6B (Figure 4.4). In addition, we verified the robustness and accuracy of the MRM-based quantitation for other GTPases. As shown in Figure 4.5 and Figure 4.6, we confirmed the up-regulation of

SAR1A and RAB27A, and down-regulation of SAR1B, RAB31 and ARF4 in the metastatic SW620 cells relative to the primary SW480 cells. Taken together, the MRM method constitutes a highly accurate quantitative tool with excellent throughput and specificity.

4.3.3 Potential Roles of SAR1A and SAR1B in CRC Progression

We sought to examine, among the differentially expressed small GTPase proteins, which can potentially promote or suppress CRC progression. Therefore, we assessed the gene expression data from patients in The Cancer Genome Atlas Colon Adenocarcinoma (TCGA-COAD) data set, and we found that there were significantly lower mRNA expressions of *ARF4*, *RAB6B*, *RHOF*, and *SAR1B* genes in CRC tissues than in normal tissues ($n = 50$; $p < 0.05$, paired Student's *t*-test) (Figure 4.7).

We also interrogated other public CRC data sets accessible from the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) database, including GSE14333 (n = 290, Ludwig Institute for Cancer Research), GSE17538 (n = 232, Vanderbilt University), GSE21510 (n = 274, Juntendo University), and GSE39582 (n = 585, Ligue Nationale contre le Cancer).[26-29] We again observed significantly lower mRNA expressions of *ARF4*, *RAB6B*, *RHOF*, and *SAR1B* in CRC tissues than in normal tissues (Figure 4.8). However, higher mRNA expression of *ARF4*, *RAB6B* and *RHOF* was significantly associated with poorer CRC patient survival, which is contradictory to their plausible roles in suppressing metastasis (Figure 4.8). The discrepancy between the proteomic data and the bioinformatic data may arise from high degree of patient-to-patient variability. Meanwhile, *SAR1A* expression is consistently up-regulated in human CRC

tissues compared to the adjacent normal tissues in both TCGA-COAD and GSE21510 cohorts (Figure 4.9). Kaplan–Meier survival analyses showed that CRC patients with higher *SAR1A* expression presented a higher risk of death [hazard ratio (HR) = 2.353, 95% confidence interval (CI) = 1.349 to 4.106, p = 0.0026], and those with higher *SAR1B* expression displayed a more favorable outcome (HR = 0.5915, 95% CI = 0.3723 to 0.9398, p = 0.0262) (Figures 4.7 and 4.8). These observations are in keeping with the quantitative proteomic data showing the down-regulated SAR1A and up-regulated SAR1B in SW620 over SW480 cells. Furthermore, lower *SAR1B* expression is significantly correlated with more advanced pathological stages of CRC (Figure 4.8). Collectively, these findings substantiated SAR1B as a potential suppressor and a prognostic biomarker for CRC progression, and SAR1A and SAR1B may assume distinct roles in disease development.

4.3.4 SAR1B Knockdown Led to Elevated in vitro Migration and Invasion of SW480 Cells by Modulating Epithelial–Mesenchymal Transition (EMT)

To explore the potential roles of SAR1B in CRC metastasis, we next sought to examine how the migratory and invasive abilities of CRC cells are modulated by genetic depletion of SAR1B. We found that shRNA-mediated knockdown of SAR1B in SW480 cells elicited marked elevations in both the migratory and invasive abilities, as manifested by results from both transwell migration/invasion assay and wound-healing assay (Figure 4.10). These observations were in accordance with the acquisition of a mesenchymal phenotype.

We next performed real-time quantitative PCR (RT-qPCR) experiments to assess the expression of EMT markers in SW480 cells upon shRNA-mediated stable knockdown of SAR1B. The results indeed demonstrated that knockdown of SAR1B led to a significant

reduction in the levels of epithelial marker E-cadherin (*CDH1*) and increased mRNA expression of the mesenchymal marker Vimentin (*VIM*), indicating a shift to a more mesenchymal phenotype (Figure 4.10). We, however, did not observe any significant changes in mRNA expression for other mesenchymal markers including *SNAI1/2* and *ZEB1/2*. In addition, the mRNA level of *CDH2* gene was not detectable in SW480 or SW620 cells. Likewise, loss of SAR1B did not alter the proliferation rate of SW480 cells (Figure 4.10). Together, SAR1B may contribute to suppression of CRC motility and metastasis partly through modulating EMT, but not by regulating proliferation.

4.3.5 RAB31 Promotes Proliferation, Migration and Invasion of SW480 Cells in vitro

RAB31 was previously shown to play a role in breast cancer proliferation and acquired tamoxifen resistance.[30-31] Our targeted proteomic data revealed a more than 4-fold down-regulation of RAB31 in the metastatic SW620 cells, which is in keeping with the quantitation results from Western-blot analysis (Figure 4.11). However, there were no significant differences observed for the *RAB31* expression between CRC and normal tissues in the TCGA-COAD data set (Figure 4.11). Moreover, Kaplan–Meier survival analyses revealed that higher *RAB31* mRNA levels are correlated with worse disease-free survival (DFS), which argues against the hypothesis that RAB31 may serve as a suppressor for CRC metastasis (Figure 4.11).

We further investigated whether knock-down of RAB31 affects the behavior of SW480 cells by assessing cell proliferation and transwell migration/invasion. The siRNA-mediated knockdown of RAB31 in SW480 cells was confirmed by Western-blot analysis (Figure 4.12). We observed significantly diminished cell proliferation, *in vitro* migration and

invasion of SW480 cells after genetic depletion of *RAB31* expression, which is accompanied by diminished level of mesenchymal marker Slug (Figure 4.12). Although RAB31 is substantially down-regulated in SW620 cells, it appeared to be a metastasis driver instead of a suppressor based on the cellular experiments. Therefore, these results underscored that some of the dysregulated small GTPases may be simply accompanied with CRC metastasis and they may not modulate metastatic transformation of CRC. This observation emphasizes the importance of validating the findings made from proteomic experiments with bioinformatic analysis and cell-based assay.

4.4 Conclusions

In this study, we employed our recently developed MRM-based targeted proteomic method for high-throughput, reproducible and in-depth quantification of small GTPases in the matched primary/metastatic SW480/SW620 CRC cells. Among the 83 quantified small GTPases, we found that SAR1A and SAR1B proteins were significantly elevated and reduced, respectively, in SW620 over SW480 cells. Furthermore, low level of *SAR1B* expression was significantly associated with poor survival in CRC patients and with elevated disease stages, suggesting its potential role in suppressing CRC progression. Our results also indicated that loss of SAR1B conferred elevated *in vitro* migratory and invasive abilities of SW480 cells, accompanied by reduced E-cadherin and elevated Vimentin expression. Metastasis consists of multiple steps of cellular transformation, during which the acquisition of enhanced cell motility, migratory capability and invasiveness is essential to the early stage of metastatic dissemination.[32] Hence, we reason that diminished SAR1B

promotes motility and *in vitro* metastasis of SW480 cells partly through modulating EMT maker gene expression.

The two SAR1 proteins, SAR1A and SAR1B, are ubiquitously expressed in many types of cancer. Aberrant expression or mutations in the *SAR1A and SAR1B* genes were shown to play an important role in regulating cholesterol biosynthesis.[33-34] SAR1B promotes secretion of both apoB48- and apoB100-lipoproteins, whereas overexpression of SAR1A displays the opposite effect, preferentially blocking the secretion of the lipid-laden particles.[34] Furthermore, reduction in total serum cholesterol could be a signal of occult CRC.[35] We therefore speculate that the crucial role of SAR1B in lipid biosynthesis and secretion may contribute to CRC progression. Interestingly, a previous study showed that the expressions of SAR1B and SAR1A were significantly decreased and increased, respectively, in the intestinal biopsies of patients with Anderson's disease as compared to healthy individuals.[36] Hence, they may display functionally redundant, yet distinct roles in regulating vesicular trafficking and cargo transport. Although we demonstrated that SAR1A is up-regulated in the metastatic SW620 cells and displayed prognostic values in large patient cohorts, the detailed molecular mechanisms underlying the role of SAR1A in CRC progression warrant further investigation.

In addition to SAR1 proteins, ARF4 was found to be down-regulated in the metastatic SW620 cells relative to the primary SW480 cells, as shown by both MRM and Western-blot analyses. In a large-scale secretome study reported by Barderas *et al.*[37] using paired CRC cell lines KM12C/KM12SM, ARF4 was found to be down-regulated in the secretome of the highly metastatic KM12SM cells relative to that of the primary KM12C cells. ARF4

induces breast cancer cell migration and functions as an anti-apoptotic protein in human glioblastoma cells.[38-39] Therefore, ARF4 may participate in the secretory pathways to suppress the malignant transformation of CRC cells.

In conclusion, we applied an MRM-based targeted proteomic method for high-throughput quantitative measurement of small GTPases associated with CRC metastasis by employing paired SW480/SW620 cell lines. Among the 25 small GTPases that exhibited differential expression with at least a 1.5-fold change in metastatic SW620 relative to primary SW480 cells, we found that SAR1B was significantly down-regulated during metastatic transformation and the *SAR1B* mRNA expression was significantly correlated with disease stages and prognosis in CRC patient cohorts. We also reported that SAR1B depletion in SW480 cells promoted cell motility and transwell invasion, partly through modulation of EMT. Collectively, these results suggested that SAR1B may serve as a prognostic biomarker and a potential suppressor for CRC metastasis.

1. Siegel, R. L.; Miller, K. D.; Fedewa, S. A.; Ahnen, D. J.; Meester, R. G. S.; Barzi, A.; Jemal, A., Colorectal cancer statistics, 2017. *Ca-Cancer J Clin* **2017,** *67*, 177-193.
2. Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R. L.; Torre, L. A.; Jemal, A., Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* **2018,** *68*, 394-424.
3. O'Connell, J. B.; Maggard, M. A.; Ko, C. Y., Colon cancer survival rates with the new American Joint Committee on cancer sixth edition staging. *J Natl Cancer I* **2004,** *96*, 1420-1425.
4. Cartwright, T. H., Treatment decisions after diagnosis of metastatic colorectal cancer. *Clin Colorectal Cancer* **2012,** *11*, 155-66.
5. Colicelli, J., Human RAS superfamily proteins and related GTPases. *Sci STKE* **2004,** *2004*, RE13.
6. Porther, N.; Barbieri, M. A., The role of endocytic Rab GTPases in regulation of growth factor signaling and the migration and invasion of tumor cells. *Small GTPases* **2015,** *6*, 135-44.
7. Chang, Y. C.; Su, C. Y.; Chen, M. H.; Chen, W. S.; Chen, C. L.; Hsiao, M., Secretory RAB GTPase 3C modulates IL6-STAT3 pathway to promote colon cancer metastasis and is associated with poor prognosis. *Mol Cancer* **2017,** *16*, 135.
8. Luo, Y.; Yu, S. Y.; Chen, J. J.; Qin, J.; Qiu, Y. E.; Zhong, M.; Chen, M., MiR-27b directly targets Rab3D to inhibit the malignant phenotype in colorectal cancer. *Oncotarget* **2018,** *9*, 3830-3841.
9. Yin, Y.; Zhang, B.; Wang, W.; Fei, B.; Quan, C.; Zhang, J.; Song, M.; Bian, Z.; Wang, Q.; Ni, S.; Hu, Y.; Mao, Y.; Zhou, L.; Wang, Y.; Yu, J.; Du, X.; Hua, D.; Huang, Z., miR-204-5p inhibits proliferation and invasion and enhances chemotherapeutic sensitivity of colorectal cancer cells by downregulating RAB22A. *Clin Cancer Res* **2014,** *20*, 6187-99.
10. Goldenring, J. R.; Nam, K. T., Rab25 as a tumour suppressor in colon carcinogenesis. *Br J Cancer* **2011,** *104*, 33-36.
11. Huang, M.; Qi, T. Y. F.; Li, L.; Zhang, G.; Wang, Y. S., A targeted quantitative proteomic approach assesses the reprogramming of small GTPases during melanoma metastasis. *Cancer Res* **2018,** *78*, 5431-5445.
12. MacLean, B.; Tomazela, D. M.; Shulman, N.; Chambers, M.; Finney, G. L.; Frewen, B.; Kern, R.; Tabb, D. L.; Liebler, D. C.; MacCoss, M. J., Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* **2010,** *26*, 966-8.
13. Kawahara, R.; Bollinger, J. G.; Rivera, C.; Ribeiro, A. C.; Brandao, T. B.; Paes Leme, A. F.; MacCoss, M. J., A targeted proteomic strategy for the measurement of oral cancer candidate biomarkers in human saliva. *Proteomics* **2016,** *16*, 159-73.
14. Escher, C.; Reiter, L.; MacLean, B.; Ossola, R.; Herzog, F.; Chilton, J.; MacCoss, M. J.; Rinner, O., Using iRT, a normalized retention time for more targeted measurement of peptides. *Proteomics* **2012,** *12*, 1111-21.
15. Livak, K. J.; Schmittgen, T. D., Analysis of relative gene expression data using real-time quantitative PCR and the 2(T)(-Delta Delta C) method. *Methods* **2001,** *25*, 402-408.

16. Leibovitz, A.; Stinson, J. C.; McCombs, W. B., 3rd; McCoy, C. E.; Mazur, K. C.; Mabry, N. D., Classification of human colorectal adenocarcinoma cell lines. *Cancer Res* **1976,** *36*, 4562-9.
17. Xue, H.; Lu, B.; Zhang, J.; Wu, M.; Huang, Q.; Wu, Q.; Sheng, H.; Wu, D.; Hu, J.; Lai, M., Identification of serum biomarkers for colorectal cancer metastasis using a differential secretome approach. *J Proteome Res* **2010,** *9*, 545-55.
18. Ghosh, D.; Yu, H.; Tan, X. F.; Lim, T. K.; Zubaidah, R. M.; Tan, H. T.; Chung, M. C.; Lin, Q., Identification of key players for colorectal cancer metastasis by iTRAQ quantitative proteomics profiling of isogenic SW480 and SW620 cell lines. *J Proteome Res* **2011,** *10*, 4373-87.
19. Lei, Y. L.; Huang, K.; Gao, C.; Lau, Q. C.; Pan, H.; Xie, K.; Li, J. Y.; Liu, R.; Zhang, T.; Xie, N.; Nai, H. S.; Wu, H.; Dong, Q.; Zhao, X.; Nice, E. C.; Huang, C. H.; Wei, Y. Q., Proteomics identification of ITGB3 as a key regulator in reactive oxygen species-induced migration and invasion of colorectal cancer cells. *Mol Cell Proteomics* **2011,** *10*, M110.005397.
20. Lee, J. G.; McKinney, K. Q.; Pavlopoulos, A. J.; Park, J. H.; Hwang, S., Identification of anti-metastatic drug and natural compound targets in isogenic colorectal cancer cells. *J Proteomics* **2015,** *113*, 326-36.
21. Cai, R.; Huang, M.; Wang, Y. S., Targeted quantitative profiling of GTP-binding proteins in cancer cells using isotope-coded GTP probes. *Anal Chem* **2018,** *90*, 14339-14346.
22. Torres, S.; Garcia-Palmero, I.; Marin-Vicente, C.; Bartolome, R. A.; Calvino, E.; Fernandez-Acenero, M. J.; Casal, J. I., Proteomic characterization of transcription and splicing factors associated with a metastatic phenotype in colorectal cancer. *J Proteome Res* **2018,** *17*, 252-264.
23. Arsic, N.; Ho-Pun-Cheung, A.; Evelyne, C.; Assenat, E.; Jarlier, M.; Anguille, C.; Colard, M.; Pezet, M.; Roux, P.; Gadea, G., The p53 isoform delta133p53 beta regulates cancer cell apoptosis in a RhoB-dependent manner. *PLoS One* **2017,** *12*, e0172125.
24. Liu, M.; Tang, Q.; Qiu, M.; Lang, N.; Li, M.; Zheng, Y.; Bi, F., miR-21 targets the tumor suppressor RhoB and regulates proliferation, invasion and apoptosis in colorectal cancer cells. *FEBS Lett* **2011,** *585*, 2998-3005.
25. Chen, W. H.; Niu, S. X.; Ma, X.; Zhang, P.; Gao, Y.; Fan, Y.; Pang, H. G.; Gong, H. J.; Shen, D. L.; Gu, L. Y.; Zhang, Y.; Zhang, X., RhoB acts as a tumor suppressor that inhibits malignancy of clear cell renal cell carcinoma. *PLoS One* **2016,** *11*, e0157599.
26. Jorissen, R. N.; Gibbs, P.; Christie, M.; Prakash, S.; Lipton, L.; Desai, J.; Kerr, D.; Aaltonen, L. A.; Arango, D.; Kruhoffer, M.; Orntoft, T. F.; Andersen, C. L.; Gruidl, M.; Kamath, V. P.; Eschrich, S.; Yeatman, T. J.; Sieber, O. M., Metastasis-associated gene expression changes predict poor outcomes in patients with Dukes stage B and C colorectal cancer. *Clin Cancer Res* **2009,** *15*, 7642-7651.
27. Freeman, T. J.; Smith, J. J.; Chen, X.; Washington, M. K.; Roland, J. T.; Means, A. L.; Eschrich, S. A.; Yeatman, T. J.; Deane, N. G.; Beauchamp, R. D., Smad4-mediated signaling inhibits intestinal neoplasia by inhibiting expression of beta-catenin. *Gastroenterology* **2012,** *142*, 562-571 e2.

28. Tsukamoto, S.; Ishikawa, T.; Iida, S.; Ishiguro, M.; Mogushi, K.; Mizushima, H.; Uetake, H.; Tanaka, H.; Sugihara, K., Clinical significance of osteoprotegerin expression in human colorectal cancer. *Clin Cancer Res* **2011,** *17*, 2444-50.
29. Marisa, L.; de Reynies, A.; Duval, A.; Selves, J.; Gaub, M. P.; Vescovo, L.; Etienne-Grimaldi, M. C.; Schiappa, R.; Guenot, D.; Ayadi, M.; Kirzin, S.; Chazal, M.; Flejou, J. F.; Benchimol, D.; Berger, A.; Lagarde, A.; Pencreach, E.; Piard, F.; Elias, D.; Parc, Y.; Olschwang, S.; Milano, G.; Laurent-Puig, P.; Boige, V., Gene expression classification of colon cancer into molecular subtypes: characterization, validation, and prognostic value. *PLoS Med* **2013,** *10*, e1001453.
30. Grismayer, B.; Solch, S.; Seubert, B.; Kirchner, T.; Schafer, S.; Baretton, G.; Schmitt, M.; Luther, T.; Kruger, A.; Kotzsch, M.; Magdolen, V., Rab31 expression levels modulate tumor-relevant characteristics of breast cancer cells. *Mol Cancer* **2012,** *11*, 62.
31. Huang, M.; Wang, Y., Roles of small GTPases in acquired tamoxifen resistance in MCF-7 cells revealed by targeted, quantitative proteomic analysis. *Anal Chem* **2018,** *90*, 14551-14560.
32. Justus, C. R.; Leffler, N.; Ruiz-Echevarria, M.; Yang, L. V., In vitro cell migration and invasion assays. *J Vis Exp* **2014**, 51046.
33. Jones, B.; Jones, E. L.; Bonney, S. A.; Patel, H. N.; Mensenkamp, A. R.; Eichenbaum-Voline, S.; Rudling, M.; Myrdal, U.; Annesi, G.; Naik, S.; Meadows, N.; Quattrone, A.; Islam, S. A.; Naoumova, R. P.; Angelin, B.; Infante, R.; Levy, E.; Roy, C. C.; Freemont, P. S.; Scott, J.; Shoulders, C. C., Mutations in a Sar1 GTPase of COPII vesicles are associated with lipid absorption disorders. *Nat Genet* **2003,** *34*, 29-31.
34. Fryer, L. G. D.; Jones, B.; Duncan, E. J.; Hutchison, C. E.; Ozkan, T.; Williams, P. A.; Alder, O.; Nieuwdorp, M.; Townley, A. K.; Mensenkamp, A. R.; Stephens, D. J.; Dallinga-Thie, G. M.; Shoulders, C. C., The endoplasmic reticulum coat protein II transport machinery coordinates cellular lipid secretion and cholesterol biosynthesis. *J Biol Chem* **2014,** *289*, 4244-4261.
35. Mamtani, R.; Lewis, J. D.; Scott, F. I.; Ahmad, T.; Goldberg, D. S.; Datta, J.; Yang, Y. X.; Boursi, B., Disentangling the association between statins, cholesterol, and colorectal cancer: a nested case-control study. *PLoS Med* **2016,** *13*, e1002007.
36. Georges, A.; Bonneau, J.; Bonnefont-Rousselot, D.; Champigneulle, J.; Rabes, J. P.; Abifadel, M.; Aparicio, T.; Guenedet, J. C.; Bruckert, E.; Boileau, C.; Morali, A.; Varret, M.; Aggerbeck, L. P.; Samson-Bouma, M. E., Molecular analysis and intestinal expression of SAR1 genes and proteins in Anderson's disease (Chylomicron retention disease). *Orphanet J Rare Dis* **2011,** *6*, 1.
37. Barderas, R.; Mendes, M.; Torres, S.; Bartolome, R. A.; Lopez-Lucendo, M.; Villar-Vazquez, R.; Pelaez-Garcia, A.; Fuente, E.; Bonilla, F.; Casal, J. I., In-depth characterization of the secretome of colorectal cancer metastatic cells identifies key proteins in cell adhesion, migration, and invasion. *Mol Cell Proteomics* **2013,** *12*, 1602-20.
38. Jang, S. Y.; Jang, S. W.; Ko, J., Regulation of ADP-ribosylation factor 4 expression by small leucine zipper protein and involvement in breast cancer cell migration. *Cancer Lett* **2012,** *314*, 185-97.
39. Wu, Q.; Ren, X.; Zhang, Y.; Fu, X.; Li, Y.; Peng, Y.; Xiao, Q.; Li, T.; Ouyang, C.; Hu, Y.; Zhang, Y.; Zhou, W.; Yan, W.; Guo, K.; Li, W.; Hu, Y.; Yang, X.; Shu, G.; Xue,

H.; Wei, Z.; Luo, Y.; Yin, G., MiR-221-3p targets ARF4 and inhibits the proliferation and migration of epithelial ovarian cancer cells. *Biochem Biophys Res Commun* **2017,** *497*, 1162-1170.

Figure 4.1 MRM-based targeted quantitative profiling of small GTPases associated with CRC metastasis.

(A) Schematic diagram depicting the targeted quantitative proteomic workflow, relying on forward SILAC labeling, in-gel fractionation, and scheduled LC-MRM analysis; **(B)** Correlation between the log_2-transformed SILAC ratios (log_2R) obtained from one forward- and one reverse-SILAC labeling experiments with a relatively high correlation coefficient ($R^2 = 0.9280$); **(C)** Venn diagrams displaying the overlap between quantified small GTPases in the triplicate SILAC experiments obtained from MRM and DDA analyses, respectively, and the comparison between the performances of the two methods; **(D)** Bar charts showing significantly up- and down-regulated (>1.5-fold) small GTPases quantified from three LC-MRM experiments.

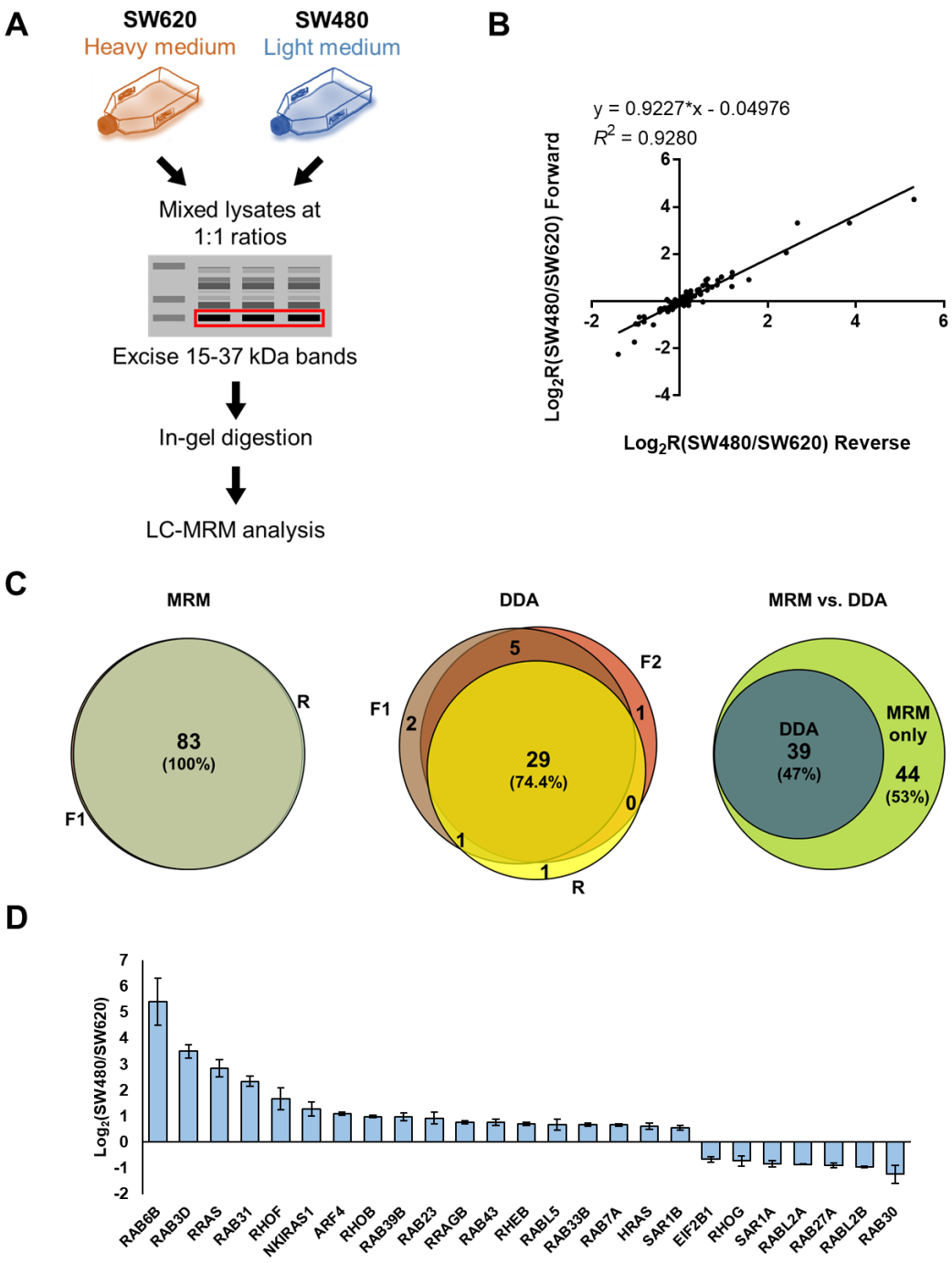

A
SW620
Heavy medium
SW480
Light medium
Mixed lysates at
1:1 ratios
Excise 15-37 kDa bands
In-gel digestion
LC-MRM analysis
B
y = 0.9227*x - 0.04976
R2 = 0.9280
Log2R(SW480/SW620) Forward
Log2R(SW480/SW620) Reverse
C
MRM
DDA
MRM vs. DDA
83
(100%)
F1
R
5
F2
F1
2
1
29
(74.4%)
0
1
1
R
DDA
39
(47%)
MRM
only
44
(53%)
D
Log2(SW480/SW620)
RAB6B
RAB3D
RRAS
RAB31
RHOF
NKIRAS1
ARF4
RHOB
RAB39B
RAB23
RRAGB
RAB43
RHEB
RABL5
RAB33B
RAB7A
HRAS
SAR1B
EIF2B1
RHOG
SAR1A
RABL2A
RAB27A
RABL2B
RAB30

Figure 4.2 Scheduled MRM analysis of differential expression of small GTPases in paired SW480/SW620 cells.

Heatmap showing the differential expression of small GTPases in paired SW480/SW620 cells. Shown are the Log_2R(SW620/SW480) values obtained from two forward and one reverse SILAC labeling experiments (F1 and F2: forward experiments, R1: reverse experiment). As indicated by the scale bar, the red and blue bars designate those small GTPases that are up-and down-regulated, respectively, by at least 1.5-fold in the SW620 over SW480 cells.

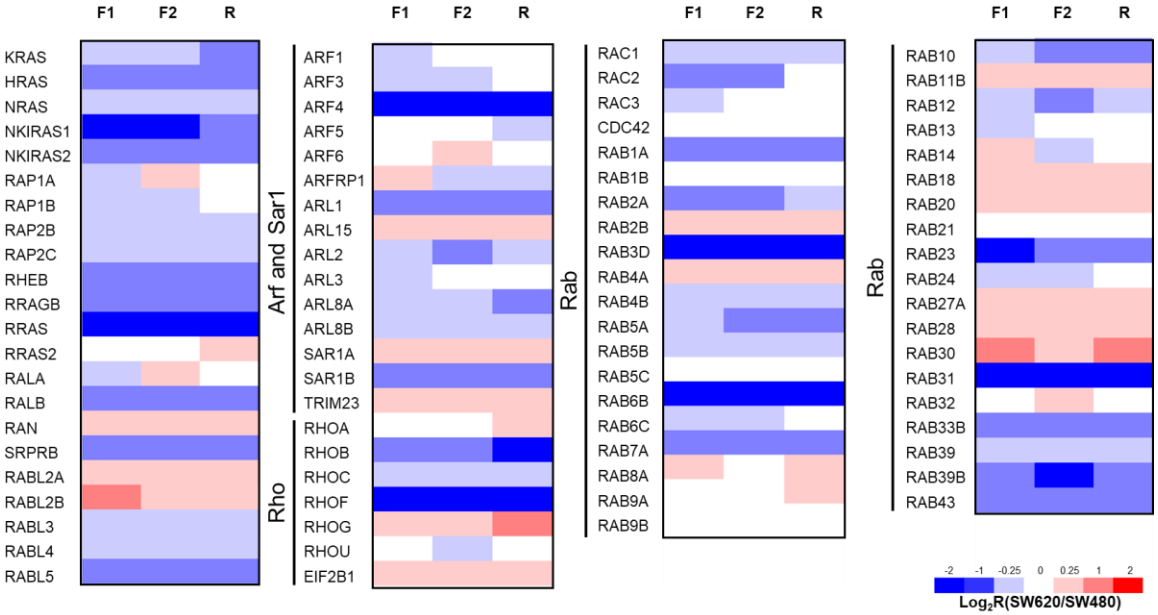


Figure 4.3 In-depth coverage of the small GTPase proteome facilitated by the Ge-LC-MRM-based quantification.

Venn diagrams showing the overlapped small GTPases quantified in this study and three independent proteomic studies.

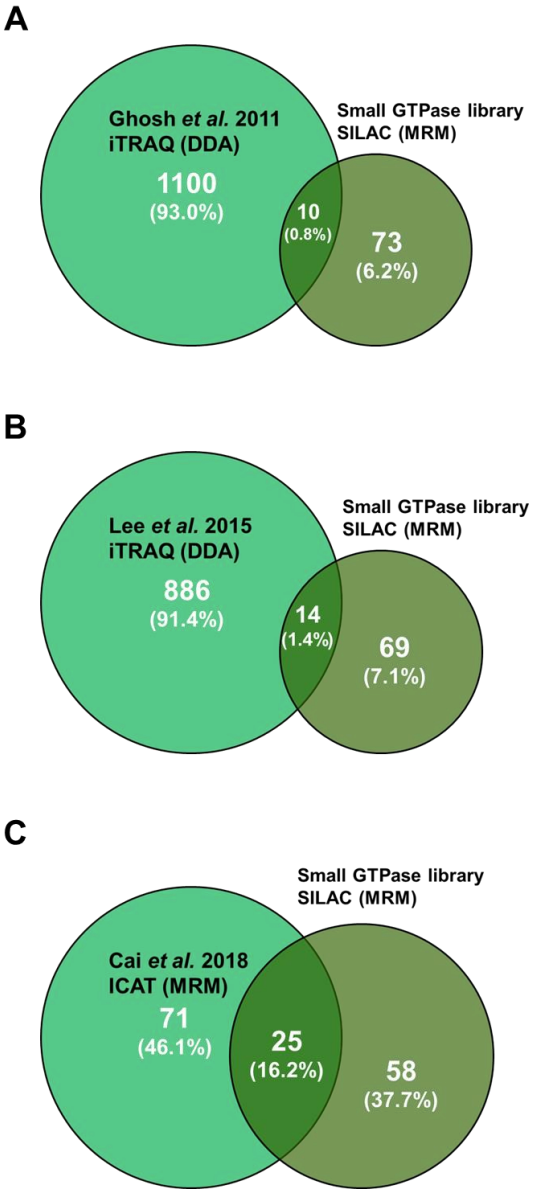


Figure 4.4 LC-MRM quantification of RHOB, RHOF, RHOG, and RAB6 in paired SW480/SW620 cells.

(A) Selected-ion chromatograms (SICs) for the quantification of tryptic peptides IQAYDYLEC̲SAK (RHOB), AALYLEC̲SAK (RHOF) and YLEC̲SALQQDGVK (RHOG) in one forward and one reverse SILAC labeling experiments; **(B)** Western-blot validation of the protein abundance of pan-RAB6 (RABA/B/C) and RAB6A, and the SICs for the quantification of tryptic peptides ELNVMFIETSAK (RAB6A), QITIEEGEQR (RAB6B) and TDLADKR (RABA/C).

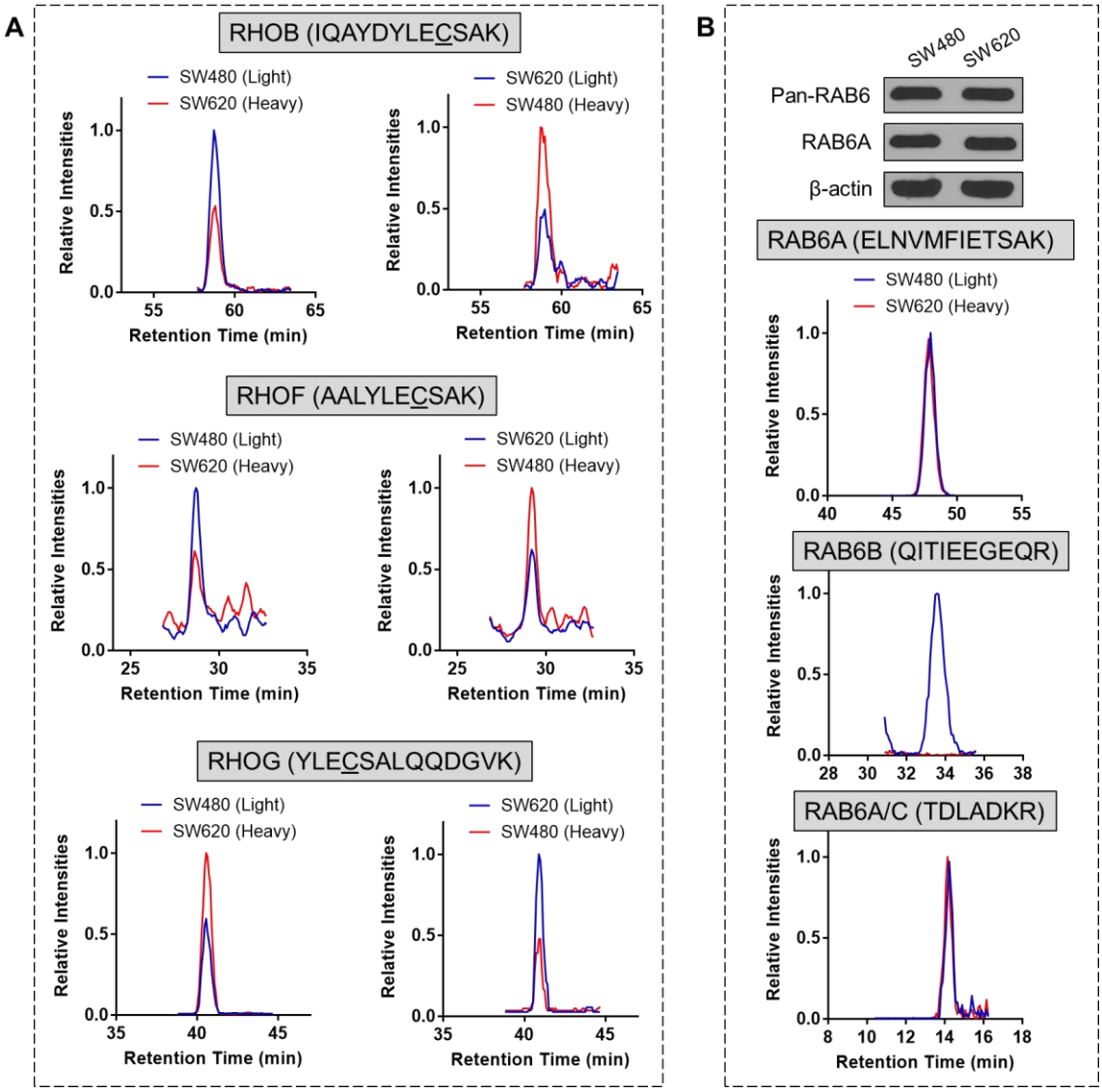


Figure 4.5 Validation of differential expression of SAR1A and SAR1B in SW480/SW620 cells.

Selected-ion chromatograms (SICs) for the quantification of tryptic peptides: **(A)** TDAISEEK from SAR1A and **(B)** EMFGLYGQTTGK from SAR1B, in one forward- and one reverse-SILAC labeling experiments; Western-blot validation of the differentially expressed SAR1A **(A)** and SAR1B **(B)** proteins in SW480/SW620 cells, and quantitative comparison of protein ratios obtained from LC-MRM and Western-blot analyses (n = 3). The *p* values were calculated by using a paired two-tailed Student's *t* test (**p* < 0.05, ***p* < 0.01). The data represent the mean and standard deviation of results obtained from three parallel experiments.

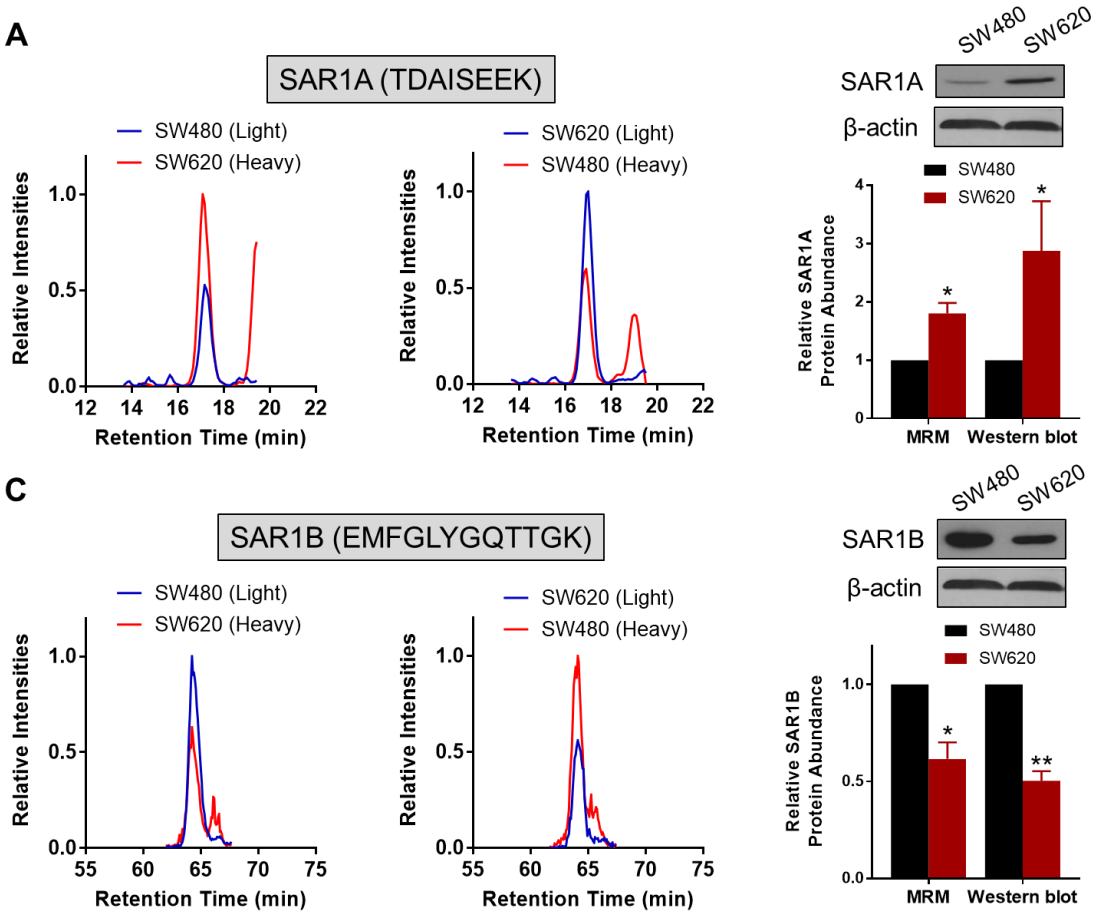


Figure 4.6 Validation of differential expression of RAB27A and ARF4 in SW480/SW620 cells.

SICs for the quantification of tryptic peptides **(A)** SWIPEGVVR from RAB27A and **(B)** LGLQSLR from ARF4, in one forward- and one reverse-SILAC labeling experiments; Western-blot validation of the differentially expressed **(A)** RAB27A and **(B)** ARF4 in SW480/SW620 cells, and quantitative comparison of protein ratios obtained from LC-MRM and Western-blot analyses (n = 3). The *p* values were calculated by using a paired two-tailed Student's *t* test (*$0.01 \leq p < 0.05$, **$0.001 \leq p < 0.01$, ***$p < 0.001$). The data represent the mean and standard deviation of results obtained from three parallel experiments.

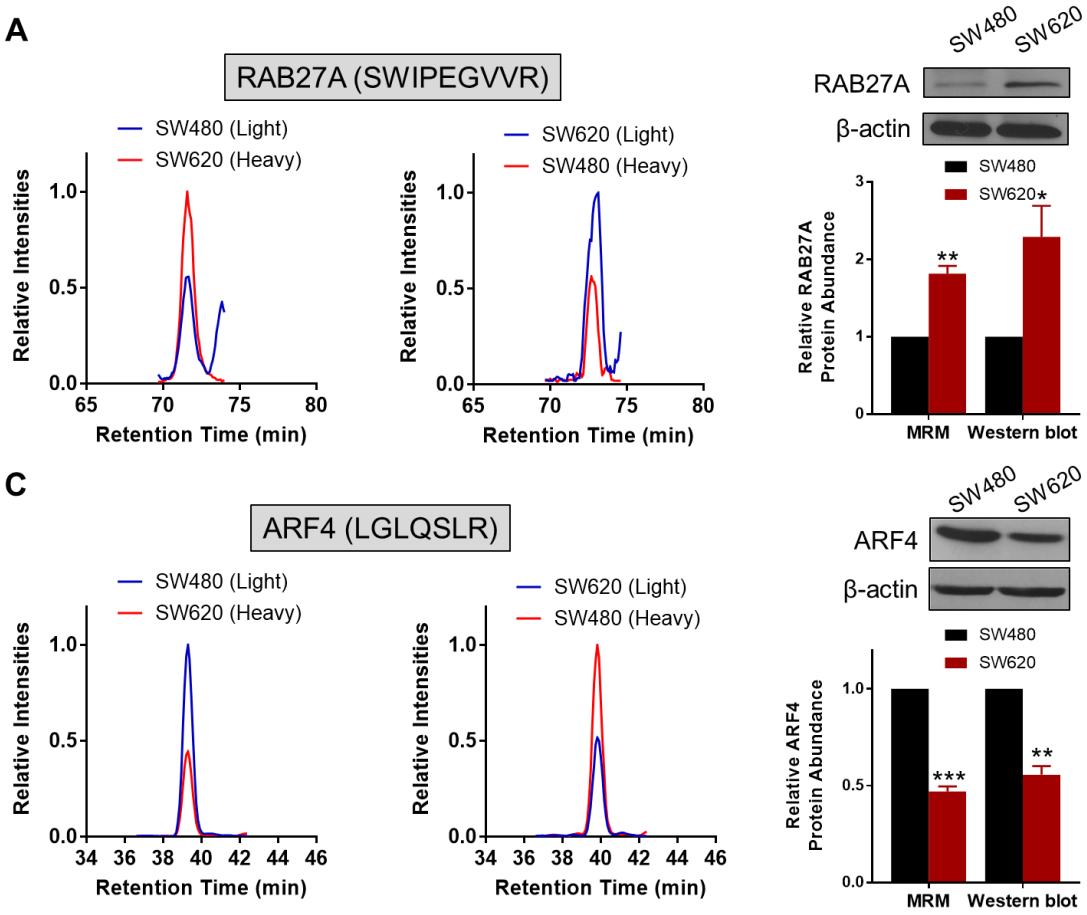


Figure 4.7 Down-regulated SAR1B in CRC confers better patient prognosis and associates with lower disease stages.

Comparison of *SAR1B* expression levels in paired CRC tissues (CRC) with adjacent non-tumor tissues (normal) in **(A)** TCGA-COAD (n = 50) and **(B)** GSE21510 (n = 44) cohorts; The *p* values were calculated by using a paired two-tailed Student's *t* test (****, $p < 0.0001$). Kaplan–Meier survival analysis of CRC patients stratified by the median *SAR1B* mRNA expression in **(C)** GSE17538 (n = 232) and **(D)** GSE39582 (n = 585) cohorts; The log-rank (Mantel–Cox) test was used to calculate the *p* values. **(E)** Correlation of *SAR1B* mRNA expression with different CRC stages in the GSE39582 (n = 585) cohort. The *p* values were calculated by using an unpaired two-tailed Student's *t* test (*$p < 0.05$).

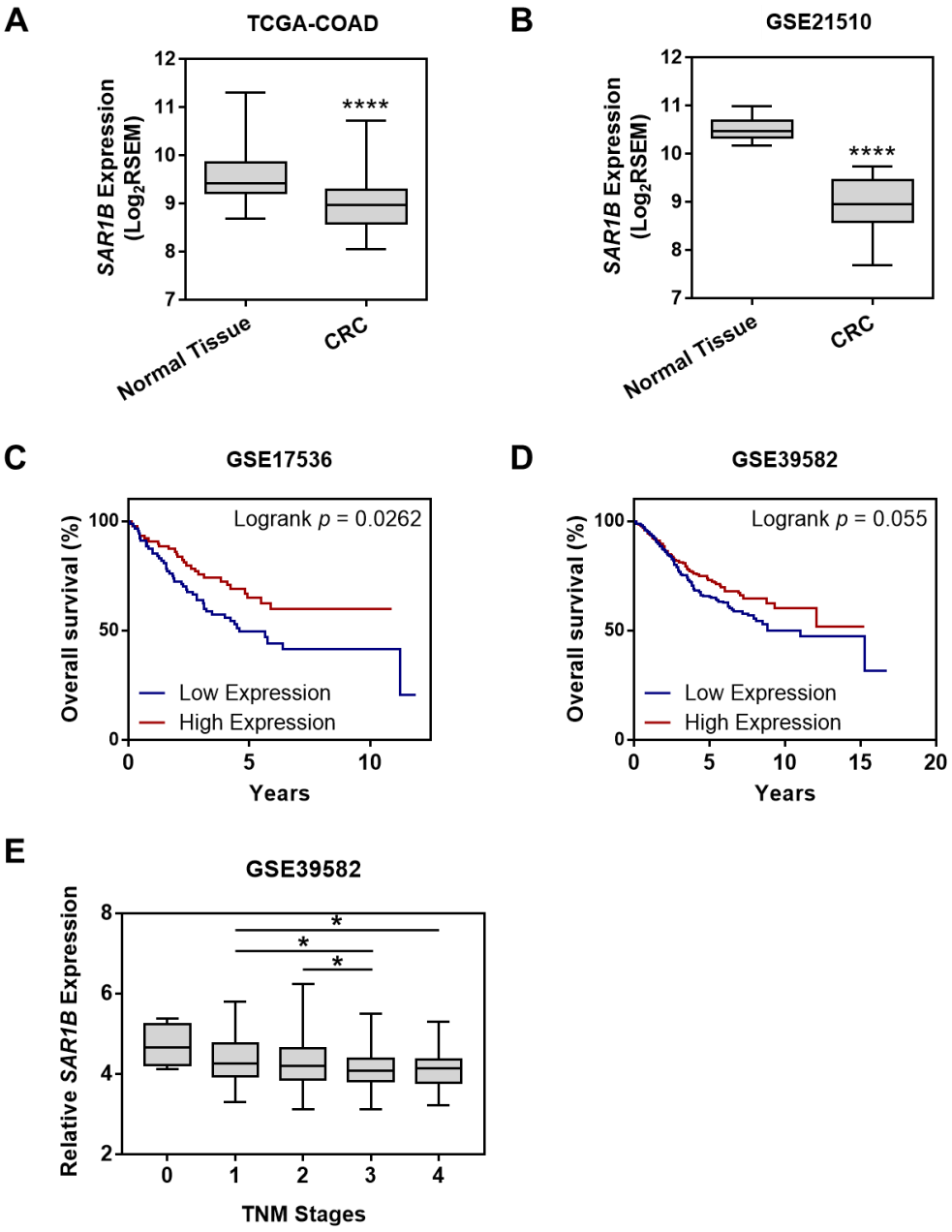

A
TCGA-COAD
SAR1B Expression (Log2RSEM)
12
11
10
9
8
7
****
Normal Tissue
CRC
B
GSE21510
SAR1B Expression (Log2RSEM)
12
11
10
9
8
7
****
Normal Tissue
CRC
C
GSE17536
Logrank p = 0.0262
Overall survival (%)
100
50
0
Low Expression
High Expression
0
5
10
Years
D
GSE39582
Logrank p = 0.055
Overall survival (%)
100
50
0
Low Expression
High Expression
0
5
10
15
20
Years
E
GSE39582
Relative SAR1B Expression
8
6
4
2
*
*
*
0
1
2
3
4
TNM Stages

Figure 4.8 Prognostic values of ARF4, RHOF and RAB6B in CRC patient cohorts.

Differential mRNA expression of *ARF4*, *RHOF and RAB6B* in the **(A, D, G)** GSE21510 (n = 44) **(B, E, H)** TCGA-COAD (n = 50) cohorts. The *p* values were calculated by using an unpaired two-tailed Student's *t* test (ns, $p > 0.05$; *, $0.01 \leq p < 0.05$; **, $0.001 \leq p < 0.01$; ****, $p < 0.0001$). Kaplan–Meier survival analysis of CRC patients stratified by the median mRNA expression levels of **(C)** *ARF4* and **(F)** *RHOF* in the GSE14333 cohort (n = 290) and **(I)** *RAB6B* in the GSE17536 cohort (n = 232). The log-rank (Mantel–Cox) test was performed to calculate the *p* values.

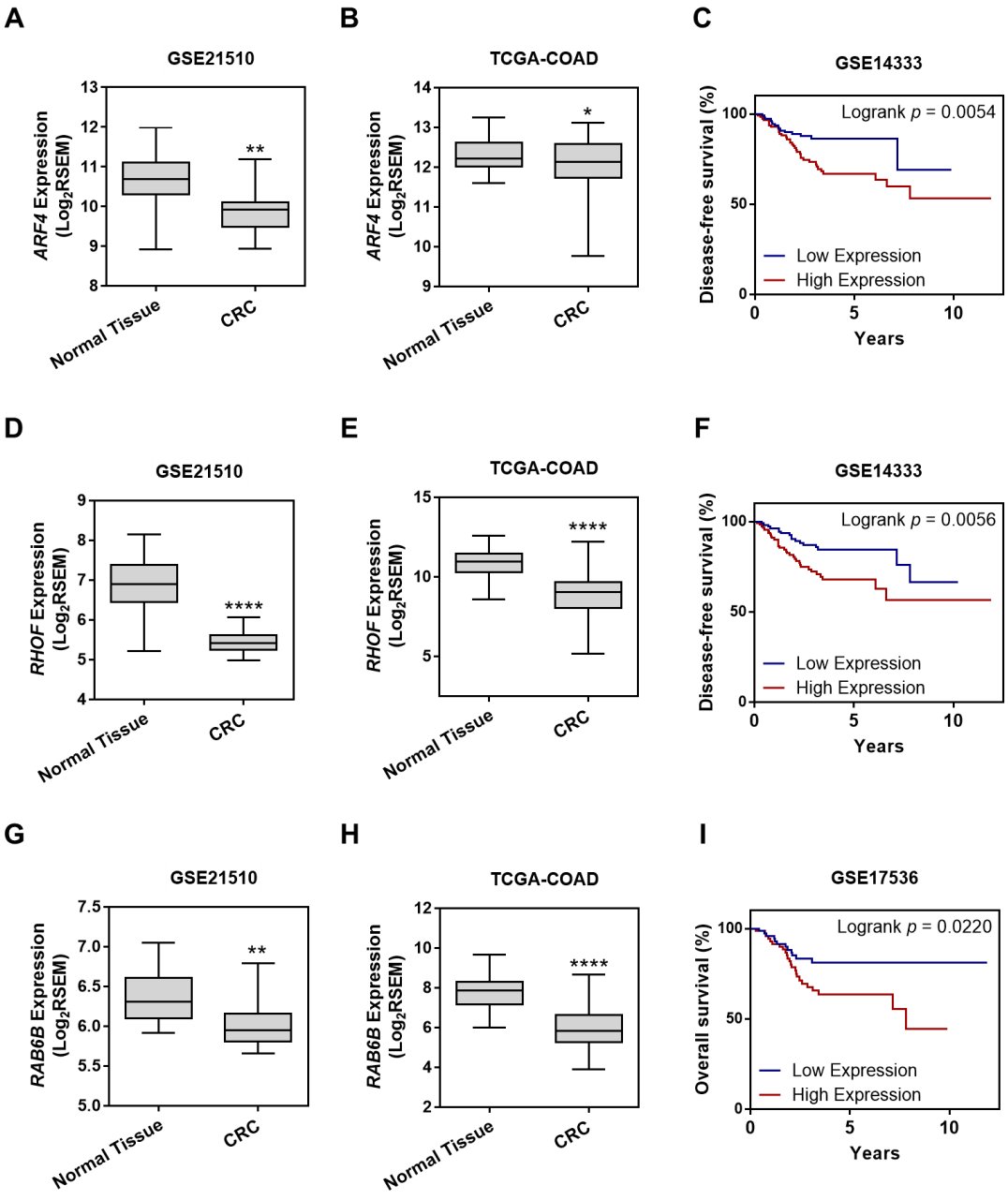

A
GSE21510
ARF4 Expression (Log2RSEM)
13
12
11
10
9
8
**
Normal Tissue
CRC
B
TCGA-COAD
ARF4 Expression (Log2RSEM)
14
13
12
11
10
9
*
Normal Tissue
CRC
C
GSE14333
Logrank p = 0.0054
Disease-free survival (%)
100
50
0
Low Expression
High Expression
0
5
10
Years
D
GSE21510
RHOF Expression (Log2RSEM)
9
8
7
6
5
4
****
Normal Tissue
CRC
E
TCGA-COAD
RHOF Expression (Log2RSEM)
15
10
5
****
Normal Tissue
CRC
F
GSE14333
Logrank p = 0.0056
Disease-free survival (%)
100
50
0
Low Expression
High Expression
0
5
10
Years
G
GSE21510
RAB6B Expression (Log2RSEM)
7.5
7.0
6.5
6.0
5.5
5.0
**
Normal Tissue
CRC
H
TCGA-COAD
RAB6B Expression (Log2RSEM)
12
10
8
6
4
2
****
Normal Tissue
CRC
I
GSE17536
Logrank p = 0.0220
Overall survival (%)
100
50
0
Low Expression
High Expression
0
5
10
Years

Figure 4.9 Higher expression of *SAR1A* in CRC confers unfavorable patient prognosis.

Comparison of *SAR1A* expression levels in paired CRC tissues (CRC) with adjacent non-tumor tissues (Normal Tissue) in the **(A)** TCGA-COAD (n = 50) and **(B)** GSE21510 (n = 44) cohorts; The *p* values were calculated by using a paired two-tailed Student's *t* test (ns, $p > 0.05$, ** $0.001 \leq p < 0.01$). Kaplan–Meier survival analysis of CRC patients stratified by the median *SAR1A* mRNA expression in **(C)** GSE14333 (n = 290) and **(D)** GSE17538 (n = 232) cohorts. The log-rank (Mantel–Cox) test was performed to calculate the *p* values.

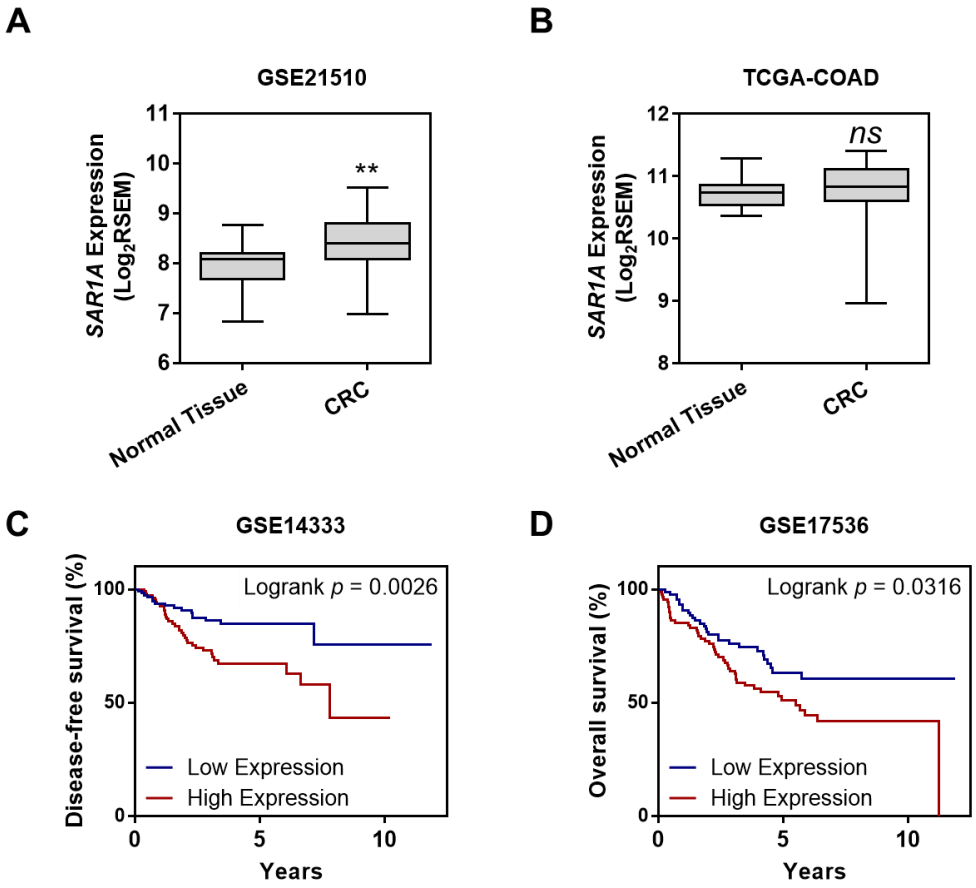


Figure 4.10 Depletion of SAR1B modulates the migratory and invasive capacities of SW480 cells.

(A) Representative images and **(B)** quantification results for migration/invasion assay showing the effects of stable SAR1B knockdown (shSAR1B) on the *in vitro* migratory and invasive abilities of SW480 cells compared to the control (shScramble). **(C)** Representative images and **(D)** quantification results for wound healing assay showing the effects of stable SAR1B knockdown (shSAR1B) on the wound healing abilities of SW480 cells compared to the control (shScramble). **(E)** RT-qPCR expression analyses of EMT marker genes in SW480 cells with stable shScramble or shSAR1B knockdown. Gene expression was normalized to the relative expression of *GAPDH*. "n.d.", not detectable. **(F)** Cell proliferation of shScramble/shSAR1B SW480 cells. The *p* values were calculated by using a paired two-tailed Student's *t* test (*, $0.01 \leq p < 0.05$, **, $0.001 \leq p < 0.01$). The data represent the means and standard deviations of results obtained from three parallel experiments.

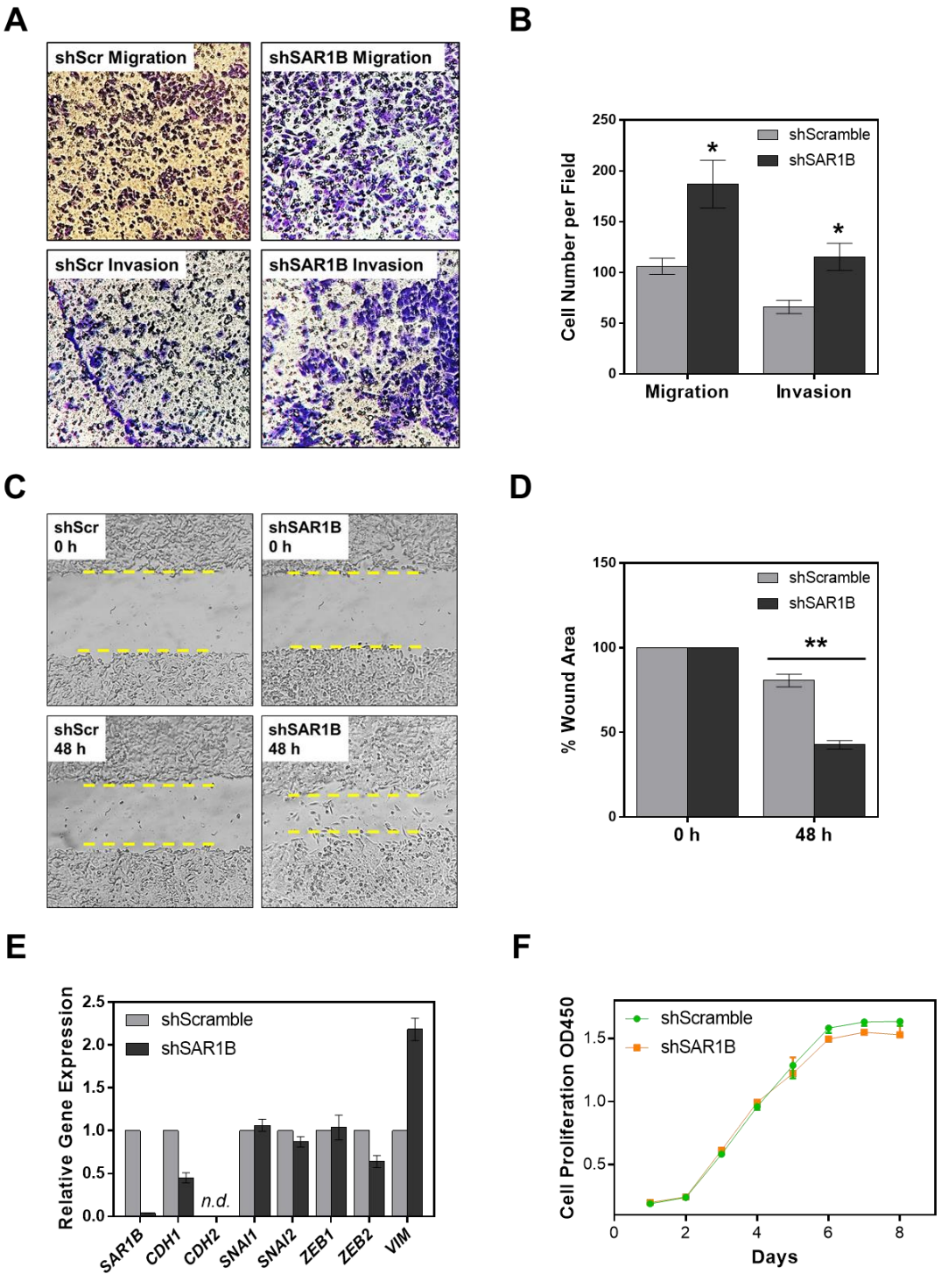

A
shScr Migration
shSAR1B Migration
shScr Invasion
shSAR1B Invasion
B
Cell Number per Field
250
200
150
100
50
0
shScramble
shSAR1B
*
*
Migration
Invasion
C
shScr
0 h
shSAR1B
0 h
shScr
48 h
shSAR1B
48 h
D
% Wound Area
150
100
50
0
shScramble
shSAR1B
**
0 h
48 h
E
Relative Gene Expression
2.5
2.0
1.5
1.0
0.5
0.0
shScramble
shSAR1B
n.d.
SAR1B
CDH1
CDH2
SNAI1
SNAI2
ZEB1
ZEB2
VIM
F
Cell Proliferation OD450
1.5
1.0
0.5
shScramble
shSAR1B
0
2
4
6
8
Days

Figure 4.11 Higher expression of *RAB31* is associated with worse CRC patient outcome and higher disease stages.

(A) SICs for the quantification of a tryptic peptide GSAAAVIVYDITK from RAB31; **(B)** Western blot validation of the differentially expressed RAB31 in SW480/SW620 cells, and quantitative comparison of protein ratios obtained from LC-MRM and Western blot analyses (n = 3). The *p* values were calculated by using a paired two-tailed Student's *t* test (*, $0.01 \leq p < 0.05$; **, $0.001 \leq p < 0.01$). The data represent the mean and standard deviation of results obtained from three parallel experiments. **(C)** Comparison of *RAB31* expression levels in paired CRC tissues (CRC) with adjacent non-tumor tissues (Normal Tissue) in the TCGA-COAD cohort (n = 50); **(D)** Correlation of *RAB31* mRNA expression with different CRC stages in the GSE39582 cohort (n = 585). The *p* values were calculated by using an unpaired two-tailed Student's *t* test (ns, $p > 0.05$; *, $0.01 \leq p < 0.05$). Kaplan–Meier survival analysis of CRC patients stratified by the median *RAB31* mRNA expression in **(E)** GSE14333 (n = 290) and **(F)** GSE17536 (n = 232) cohorts. The log-rank (Mantel–Cox) test was performed to calculate the *p* values.

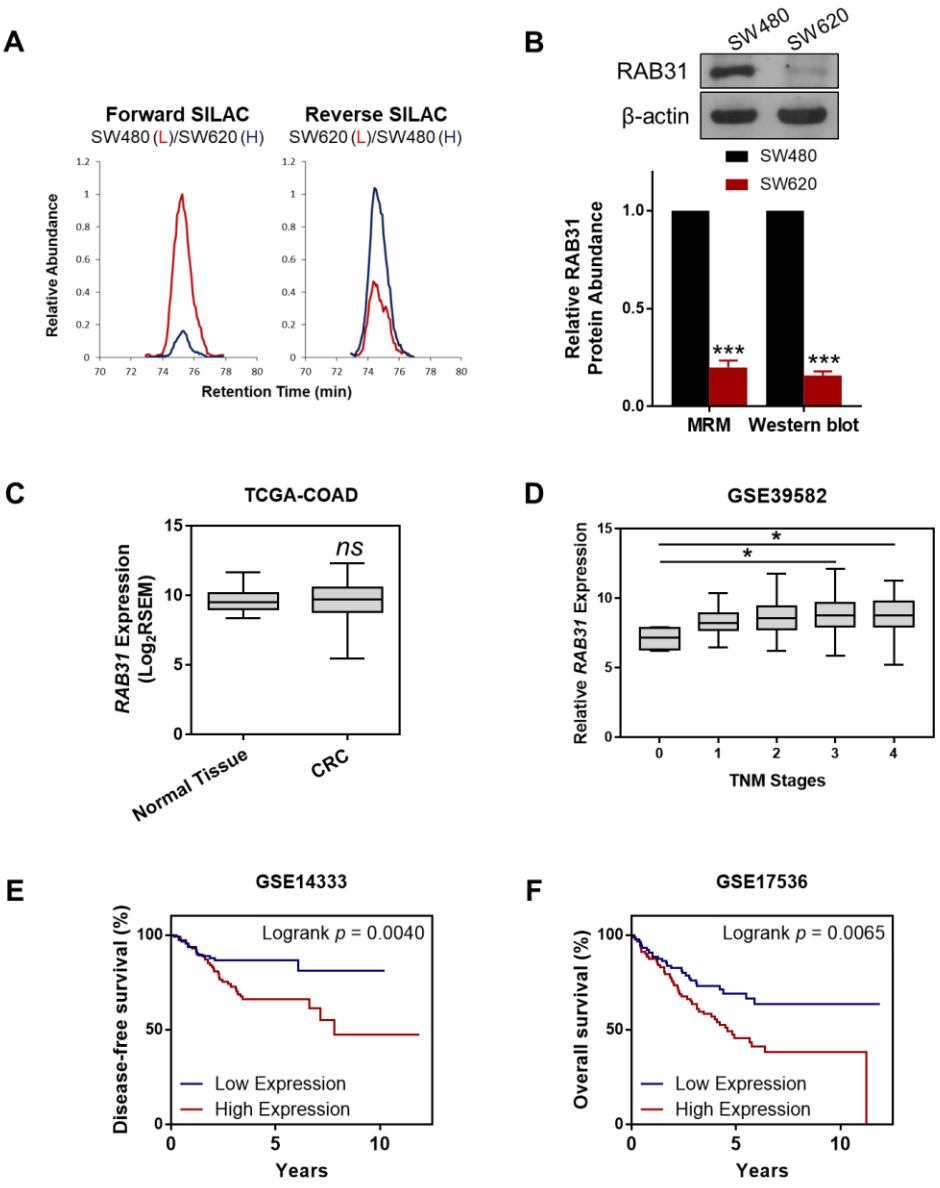
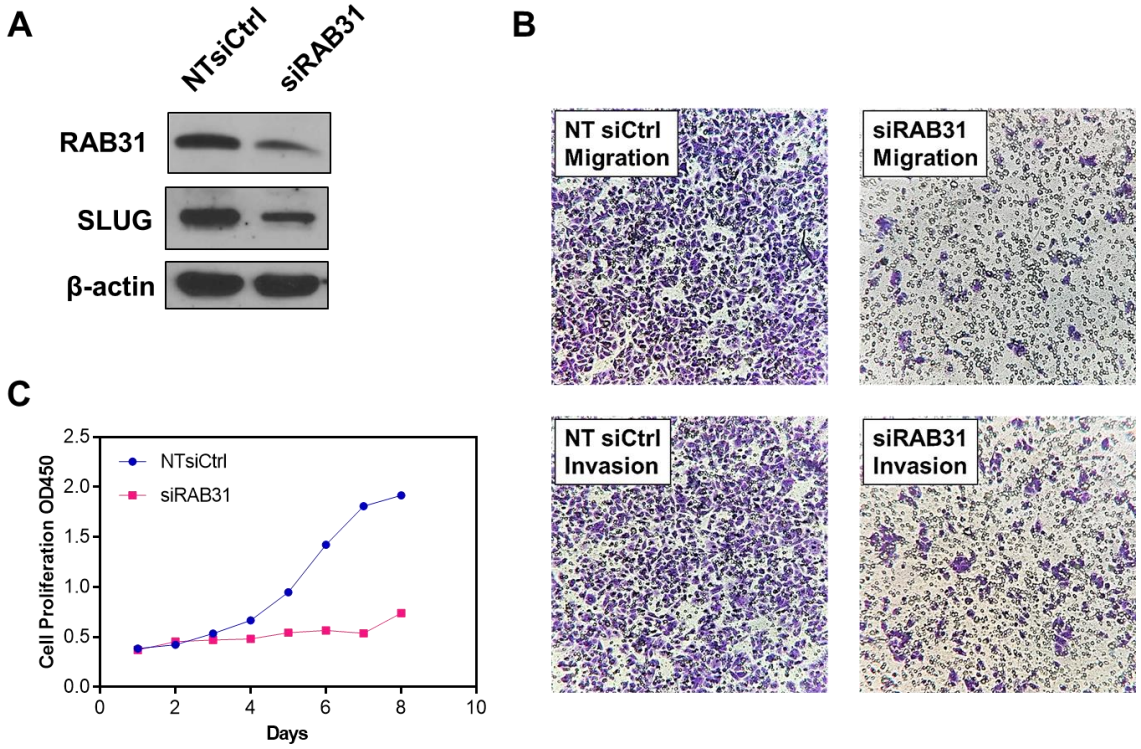

A
Forward SILAC
SW480 (L)/SW620 (H)
Reverse SILAC
SW620 (L)/SW480 (H)
Relative Abundance
Retention Time (min)
B
SW480
SW620
RAB31
β-actin
SW480
SW620
Relative RAB31 Protein Abundance
***
***
MRM
Western blot
C
TCGA-COAD
RAB31 Expression (Log2RSEM)
ns
Normal Tissue
CRC
D
GSE39582
Relative RAB31 Expression
*
*
TNM Stages
E
GSE14333
Logrank p = 0.0040
Disease-free survival (%)
Low Expression
High Expression
Years
F
GSE17536
Logrank p = 0.0065
Overall survival (%)
Low Expression
High Expression
Years

Figure 4.12 RAB31 knockdown modulates the migratory and invasive capacities and cell proliferation of SW480 cells.

(A) Western blot results showing the siRNA-mediated knock-down of RAB31 in SW480 cells and the resulting altered level of Slug protein; **(B)** Representative images depicting the migratory and invasive abilities of SW480 cells upon siRNA-mediated depletion of RAB31; **(C)** Cell proliferation of SW480 cells upon siRNA-mediated knockdown of RAB31.



Chapter 5 A Targeted Quantitative Proteomic Approach for High-throughput Quantitative Profiling of Small GTPases in Brain Tissues of Alzheimer's Disease Patients

5.1 Introduction

Alzheimer's disease (AD) is among the most common neurodegenerative diseases worldwide and is a leading cause for dementia in the elderly.[1] In the United States, AD affects approximately 10% people with ages over 65 and remains the 5th leading cause of death in this population.[2] The hallmarks of AD pathology in the brains include synapse loss, the accumulation of extracellular beta-amyloid (Aβ) plaques and twisted strands of the hyper-phosphorylated tau protein (neurofibrillary tangles) inside neurons, which promote neuronal death and ultimately dementia.[3] As region-specific neurodegenerative disorders, AD can lead to neuronal loss that is predominantly found in the cerebral cortex and hippocampus, where the cerebral cortex can be further divided into different lobes, the frontal, parietal, temporal, and occipital.[4]

The advent of quantitative proteomics in recent years has allowed for high-throughput interrogation of proteomic alterations accompanied by the early development and progression of AD in various sample types or post-mortem brain regions including temporal neocortex,[5] cerebral cortex,[6] cerebrospinal fluids,[7-8] hippocampus,[9-10] frontal cortex[11-13] and anterior cingulate gyrus.[14] Very recently, McKetney *et al.*[15] reported a proteomic resource comprising nine anatomically distinct sections from post-mortem AD patient brain tissues. Among the various proteomic techniques, targeted proteomics was widely applied in clinics for validation of disease biomarker for AD.[8, 16]

Small GTPases of the Ras superfamily are essential regulators of intracellular trafficking and signal transduction, and therefore, they represent a potential therapeutic strategy in disease treatment.[17] A growing body of literature has emphasized the importance of Rab small GTPases as crucial signaling modules in the brain.[18-20] Rab7A and Rab35 regulate secretion and endolysosomal degradation of the tau protein, respectively.[21-22] Moreover, Rab6A is among the 16 hub genes associated with metastable AD subproteome,[23] and Rab6 showed elevated expression levels in the temporal cortex, but not hippocampus of AD brains compared to non-demented controls.[24] In addition, the mRNA expressions of Rab4, Rab5 and Rab7 were up-regulated in the hippocampus from individuals with AD.[25] Augmented mRNA expressions of Rab4, Rab5, Rab7, and Rab27 were also observed in the basal forebrain of AD patients.[26] In addition to Rab GTPases, Arf and Rho GTPases also play important roles in regulation of membrane dynamics and trafficking in neuronal cells.[27] Given the importance of synaptic trafficking in neurodegenerative diseases, we reason that a comprehensive investigation of the association of small GTPases with AD could further our understanding of disease etiology and explore their potential as possible molecular targets for therapeutic intervention. Here we utilized a novel multiplexed and high-throughput targeted quantitative proteomic approach that involves multiple-reaction monitoring (MRM) and the use of stable isotope-labeled standard (SIS) peptides. We further applied this method to assess the altered expression of small GTPases in the frontal cortex region of post-mortem patient brain tissues with various AD stages.

5.2 Materials and Methods

5.2.1 MRM Library Construction

We recently established a Skyline library incorporating more than 130 proteins from the small GTPase of Ras superfamily based on in-house shotgun proteomic data.[28] Built upon the old library and an online targeted proteomic experiment design tool Picky (https://picky.mdc-berlin.de),[29] the current MRM library incorporates 148 proteins that represent 138 unique gene IDs along with a total of 1670 MRM transitions. Cysteine carbamidomethylation (C +57.0215 Da) was set as a fixed modification, and heavy lysine (K +8.0 Da) and heavy arginine (R +10.0 Da) modifications were incorporated as heavy-isotope labels.

5.2.2 Brain Tissue Homogenization and Protein Extraction

Snap-frozen human brain tissues from medial frontal cortex were obtained from 15 individuals, and the AD cases were neuropathologically diagnosed as Braak stages 3−6. Among them, 5 individuals with high AD pathology (Braak stages 5−6, "high AD" group), 5 individuals with medium AD pathology (Braak stages 3−4, "intermediate AD"), and 5 age-matched individuals with no or low AD pathology (Braak stage 1 or none, "no AD"). Each tissue piece (approximately 100 mg in wet weight) was homogenized in a 300 μL of RIPA lysis buffer (20 mM Tris-HCl, 150 mM NaCl, 1 mM Na_2EDTA, 1% NP-40, pH 7.5), supplemented with protease inhibitor cocktail (1:100, v/v) using a Bullet Blender (Next Advance) and 100 mg of 0.5 mm zirconium oxide beads (Next Advance). After centrifugation at 15,000 RPM for 30 min at 4 °C, the supernatant was transferred to another

vial. Total protein concentration was determined using the Quick Start Bradford Protein Assay (Bio-Rad).

5.2.3 MS Sample Preparation

Approximately 10 μg of tissue lysate was separated on a 10% SDS-PAGE gel, followed by excision of gel bands in the molecular weight range of 15−37 kDa. The excised gel bands were cut into pieces (1-mm^3 volume), reduced with dithiothreitol, alkylated with iodoacetamide, and digested in-gel with trypsin (enzyme/substrate 1:50) at 37 °C overnight.

The peptide samples were desalted, reconstituted in 40 μL of buffer A (0.08% formic acid) and 36 μL of sample was spiked with 4 μL of crude SIL peptide stock solution (20 fmol/μL) prior to scheduled LC-MRM analysis, to yiled a final spike-in concentration of 2 fmol/μL. Approximately 4 μL (20 ng equivalent) of sample was loaded onto a 4-cm long in-house prepared trapping column packed with 5 μm Reprosil-Pur C18-AQ resin (Dr Maisch) and separated on a 25-cm long 75 μm ID fused silica columns packed in-house with 3 μm Reprosil-Pur C18-AQ resin (Dr Maisch). The LC-MS platforms consisted of a Dionex UltiMate 3000 RSLCnano UPLC system coupled to a TSQ Altis triple-stage quadrupole mass spectrometer with a Flex nano-electrospray ion source (Thermo Fisher). Sample elution was performed over a gradient of 12−33% Buffer B (0.08% formic acid in 80% acetonitrile) over 70 min, and from 33−100% B over 5 min at a flow rate of 300 nL/min. The column was reconditioned with 100% B for 3 min and equilibrated with 1% B for 10 min at a flow rate of 300 nL/min. The mass spectrometer was operated with an ion spray voltage of 2200 V, a capillary offset voltage of 35 V, a skimmer offset voltage

of -5 V, and a capillary inlet temperature of 325 °C. Both Q1 and Q3 were set at a resolution of 0.7 fwhm, and Q2 gas pressure used for peptide fragmentation was set at 1.5 mTorr. Collision energies specific to peptide precursors were calculated in Skyline software (version 4.2.0). A modified iRT calculator was employed for retention time prediction and generation of the MRM method in Skyline, as described recently.[28]

5.2.4 Data Analysis

The raw data were directly imported into Skyline for visualization of chromatograms of target peptides to manually determine the detectability of target peptides. For each peptide, the transitions without matrix interference were used for precise quantification. Three criteria were used to determine the peak detection and integration: (1) scheduled retention time, (2) relative distribution of each transition (dotp value), and (3) co-elution of endogenous (light) peptides and heavy IS peptides. All the data were manually inspected to ensure that the intensity distribution of selected transitions match with theoretical distributions in the spectral library, with dotp being greater than 0.8.[30] MRM peak area normalization was applied with the use of crude SIL peptides as IS or SS.

5.2.5 Heavy Isotope-labeled Synthetic Peptides

Crude synthetic peptides labeled with $^{13}C/^{15}N$ on C-terminal lysine and arginine (>70% in peptide purity and >99% in isotopic purity) were synthesized and purified by New England Peptides (Cambridge, MA). The lyophilized form of peptides was reconstituted with 15% acetonitrile in 0.1% formic acid, aliquoted and stored at -80 °C until use to avoid freeze/thaw cycles. The crude heavy peptide standards of small GTPases were added to each sample at a final concentration of 0.5, 2, and 5 fmol/μL.

5.2.6 Immunoblotting

After tissue homogenization, the protein concentration in the total brain tissue lysates was determined using Quick Start Bradford Protein Assay (Bio-Rad). Approximately 10 μg of lysates were mixed with 4× Laemmli SDS loading buffer, boiled at 95 °C for 5 min, and loaded onto 10% polyacrylamide gels. After electrophoresis, the proteins were transferred onto nitrocellulose membranes. After blocking with 5% non-fat milk in PBS with 0.1% Tween-20 (PBST) at 25 °C for 1 h, the membranes were incubated with primary antibodies against human Rab27B (Proteintech; rabbit polyclonal, 1:2,000), RAB31 (4D12, Santa Cruz; rabbit polyclonal, 1:2,000), or β-actin (Thermo Fisher; rabbit polyclonal, 1:10,000). After overnight incubation with primary antibodies at 4°C with 5% bovine serum albumin (BSA) in PBST, the membranes were then incubated with peroxidase-labeled donkey anti-rabbit secondary antibody (Thermo Fisher; 1:10,000) or mouse m-IgGκ BP-HRP (Santa Cruz; 1:10,000) for 1 h at 25 °C. Amersham ECL Prime Western Blot Detecting Reagent (GE Healthcare) was used for visualization of protein bands.

5.3 Results

5.3.1 Targeted Assay Development

In the past decade, LC-MS/MS in the multiple-reaction monitoring (MRM) mode has allowed for sensitive, reproducible and reliable protein quantitation in different biological matrices through multiplexed and targeted quantification of peptides.[31] In MRM-based quantification, specific transitions of precursor ions selected in Q1 to fragment ions selected in Q3 are monitored using a triple-quadrupole instrument, generating signals for

quantification. Therefore, analytical robustness of the method is highly dependent upon the selection of optimal proteotypic peptides and transitions that represent the proteins of interest.[32] Recently, we established a Skyline MRM spectral library in combination with the metabolic labeling by stable isotope labeling of amino acids in cell culture (SILAC) for high-throughput quantitative profiling of small GTPases in human cancer cell lines.[28] This library was constructed primarily using shotgun proteomic data generated in-house, covering 134 proteins representing 113 unique gene IDs. With retention time scheduling, all targeted transitions for small GTPases can be monitored in two LC-MRM runs with the use of a 6-min retention time window.

To further expand the MRM library, we referred to an online targeted proteomic experiment design tool Picky (https://picky.mdc-berlin.de) to map comprehensively the candidate peptides for small GTPases of human or mouse origins.[29] The peptides were chosen based on their intensities as well as sequence uniqueness. With the use of this tool, we were able to expand the library to 148 proteins representing 138 unique gene IDs, covering 90% of the human small GTPase proteome. To the best of our knowledge, this is the most comprehensive MRM library for small GTPases thus far. Because a relatively large number of transitions (~1680) were monitored in a single LC run, we employed retention time scheduling using iRT.[33] All targeted transitions for small GTPases can be monitored in a single LC-MRM run in scheduled MRM analysis, with a retention time window of 4 min to yield the maximum concurrent transitions of ~50 (Figure 5.1).

5.3.2 Analytical Performance of Crude SIL Peptides

The absolute quantification (AQUA), which involves the spiking of samples with high-purity (>98%) stable isotope-labeled peptides as internal standards (SIS), enables highly reliable and specific targeted quantification of protein or post-translational modifications.[34] Peptides labelled with a stable isotope (^{13}C and ^{15}N) are chemically identical to their native counterparts and thus have identical chromatographic behavior. The AQUA method usually requires the generation of calibration curves by stable isotope dilution using high-purity SIS peptides and is therefore more expensive and labor-intensive.

Other studies used MRM signals for a single or selected group of peptides derived from housekeeping proteins (e.g. GAPDH or actin) to normalize MRM signals from the targeted peptides.[35-36] Additionally, the labeled reference peptide (LRP) approach emerged as a cost-effective normalization strategy for a larger number of target proteins, in which a single isotope-labeled peptide standard is used as the reference for all target peptides in an analysis.[35] Compared to label-free quantification, the LRP method can compensate for technical variations during sample analysis and thus yield more reliable normalization.[37]

Crude synthetic peptides, on the other hand, display quite limited uses in targeted quantitation due to their large variation in purities. Owing to simplified purification, a larger set of candidate peptides can be incorporated into MRM assays. Despite preliminary purification, lower-purity crude peptides may still contain a wide range of impurities such as residual salts, deblocking and scavenger reagents, and truncated and partially deblocked peptides.[38] To this end, we synthesized 131 peptides carrying a uniformly [^{13}C,^{15}N]-labeled arginine or lysine at the C termini. In order to eliminate possible quantification bias

introduced by the impurities present in the crude standards, we first evaluated, by employing MRM analysis, the analytical performance of the small GTPase crude peptides synthesized by NEP in a 96-well plate format. We confirmed the absence of isotope interference for all peptides (Figure 5.2). In addition, we detected 114 out of 131 crude SIL peptides in a single LC-MRM run (Figure 5.3), which exhibited a broad dynamic range spanning 3 orders of magnitude (Figure 5.3). Failure in detection of the other peptides may be attributed to ion suppression resulting from matrix effects or low abundance of the peptides in the crude peptide pool.

5.3.3 Evaluation of Relative Quantitation by Crude SIL Peptides

Next, we sought to investigate the use of crude peptides in performing relative quantitation. Despite their lower purities, crude SIL peptides can be added at an equivalent concentration to all samples thus can be used for correction of sample losses during transfer and normalization across runs due to LC injection variations. Figure 5.1 displays the linear correlations obtained by the spike-in concentration and the peak area ratios (light/heavy), where most peptides exhibit high linear regression coefficients ($R^2 > 0.995$). The results again supported the feasibility of spiking crude SIL peptides in samples for relative quantitation. We further optimized the concentration of crude peptides spiked in the samples as internal standards (IS). Among the three spike-in concentrations (0.5, 2, 5 fmol/μL), we found that using 2 fmol/μL could result in more properly distributed light/heavy ratios and dynamic range (Figures 5.1−5.3). Collectively, we determined 2 fmol/μL as the spike-in concentration used for the subsequent experiments.

For other library peptides without their heavy counterparts as IS, we adapted slightly from the LRP method and adopted the concept of RT-defined surrogate standards (SS). Unlike IS peptides, which share high chemical and chromatographic similarities with target peptides, heavy SS peptides were chosen based upon their high intensity, lower variability and similar chromatographic behaviors. Therefore, we selected twelve SS peptides that elute at varied retention time windows across the entire gradient (Figure 5.4). As depicted in Figure 5.5, we found that the SS peptides eluting at similar retention times exhibited more similar distribution of MRM signal intensity or peak area, compared to those eluting at different retention times. Relative to existing normalization methods used in MRM-based quantitation, such as label-free and LRP method using a single peptide derived from house-keeping proteins, we reasoned that the use of crude SIL peptides offers a cost-effective and reliable alternative in MRM-based quantification.

5.3.4 Targeted Proteomic Analysis of Small GTPases in AD Brain Tissue Samples

To explore the potential roles of small GTPases in AD progression, we applied the established MRM assay to assess the differential expressions of small GTPases in the frontal cortex region of post-mortem patient brain tissues with differed disease stages (Figure 5.6). In total, we were able to quantify reproducibly more than 80 small GTPases among the 15 brain tissue samples (Figure 5.7). As the availability of such clinical material for analysis is often limited, we also demonstrated that a relatively small amount of protein input (~10 μg/sample) is sufficient to achieve a good coverage of target proteins by the established LC-MRM method. Figure 5.7 shows a heatmap for a more detailed comparison of the mean quantification results obtained from each small GTPase among the two disease

groups (intermediate AD, and high AD) normalized to the mean values of the normal group (no AD).

5.3.5 Altered Expression of Small GTPases Involved with Synaptic Functions

Synaptic vesical trafficking that involves continuous cycles of exocytosis and endocytosis is required for the maintenance of proper synaptic functions.[39-40] By performing network analysis, Kokotos *et al*.[41] revealed that Rab small GTPases constitute a key functional hub within the activity-dependent bulk endocytosis (ADBE) proteome in cerebellar granule neurons. Table S2 shows the network analysis of the synaptic Rab small GTPases reported in literature.

A loss of synaptic contacts in both the neocortex and hippocampus represents one of the major neuropathological hallmarks associated with AD.[42] A couple of Rab GTPases were previously determined by Pavlos *et al*.[43] as the exocytotic (Rab3A, Rab3B, Rab3C, and Rab27B) and endocytic (Rab4B, Rab5A/B, Rab10, Rab11B, and Rab14) Rab machinery of synaptic vesicles (SVs). Interestingly, the LC-MRM data revealed aberrant expressions of several synaptic small GTPases, such as Rab3A/C, Rab4A/B and Rab27B in disease progression of AD. Among the several highly homologous Rab3 isoforms (Rab3A, Rab3B, Rab3C and Rab3D), Rab3A is the most abundant in the brain, where it is localized on SVs and participates in Ca^{2+}-triggered neurotransmitter release.[44] In general, these Rab3 isoforms usually play largely overlapping yet redundant secretory functions in neurons and are crucial to synaptic integrity.[45] Rab3B was not quantified by the LC-MRM method, which is presumably due to its low protein abundance. Based on the LC-MRM results, we observed increased protein expression of Rab3A, Rab3C and Rab3D in later

stages of AD (Figure 5.7), which indicate important, yet previously unrecognized functions of small GTPase-regulated synaptic trafficking and signaling in promoting AD pathology.

Furthermore, two synaptic small GTPases, Rab4A and Rab4B, were significantly up-regulated in the "high AD" group of patients compared to the "intermediate AD" group (Figure 5.8). However, there were no significant changes when the comparison was imposed on "no AD" group and "high AD" group. It is also noteworthy that Rab4 was found to be up-regulated at mRNA levels in the basal forebrain and the hippocampus regions of AD brains in two independent studies.[25-26] In contrast to the microarray-based studies, our MRM-based approach indicated that both isoforms of Rab4 (Rab4A and Rab4B) may be involved with synaptic functions that could potentially implicate with AD progression. For other synaptic GTPases such Rab10, there were no obvious alteration in expression among the three patient groups (Figure 5.8). Collectively, we showed that the MRM quantitation results obtained by two normalization methods (IS-based and SS-based) were highly similar, and that dysregulated synaptic trafficking could play potential roles in disease progression of AD.

5.3.6 Validation of Proteomic Data by Western Blot Analysis

In addition to Rab3, Rab27B also play a distinct yet overlapping role in SV trafficking and involves with Ca^{2+}-dependent exocytosis.[43] Quite interestingly, we also observed substantial, yet statistically insignificant, increases of Rab27B in higher stages of AD (Figure 5.9). The endogenous/heavy peptide LLALGDSGVGK showed a dotp value between 0.93 and 0.97, and the endogenous peptide FITTVGIDFR displayed a dotp value between 0.89 and 0.93 (Figure 5.10). Furthermore, the quantification results among the 15

brain samples normalized by both IS and SS peptides showed a reasonably high linear regression (R^2 = 0.961) (Figure 5.10). Again, we demonstrated that the quantitation performances between these two normalization methods are adequately reliable.

To validate the proteomic results, we also performed western blot analysis and found that there was a moderate correlation between the MRM-based and the immunoblot-based analyses, which displayed a modest linear regression coefficient of R^2 = 0.5889 (Figure 5.9). These results indicate the differences in dynamic range of the quantification results obtained from the two methods, which led to a higher degree of variation with increased protein abundance.

5.4 Discussion

Previous systems biology analyses and proteomic studies revealed substantial reprogramming of a plethora of cellular processes involved with AD pathology and disease progression, such as energy metabolism, glycolysis, oxidative stress, apoptosis, signal transduction, and synaptic functioning.[5] Potential roles of small GTPases participating in synaptic trafficking and modulation of neurodegeneration have been increasingly discussed recently.[20] In this study, we found a substantial up-regulation of Rab27B protein level in higher disease stage of AD, which is in accordance with the dysregulated Rab27B in AD reported in previous proteomic and transcriptomic studies. Furthermore, several other synaptic GTPases, such as Rab3A/C and Rab4A/B, were found to be up-regulated in brain tissues with higher AD levels. It is also expected that the proteomic profiles within different brain regions are likely to change qualitatively and/or quantitatively during aging and/or in different disease states. Thus, identification of proteins unique to each brain

region, those associated with neurodegenerative mechanisms could yield opportunities to overcome major obstacles in the development of new protective and restorative therapies for prominent neurodegenerative diseases.

5.5 Conclusions

In this study, we developed a targeted quantitative MRM-based proteomic assay for quantification with >1,600 transitions representing 550 peptide precursors in a 45-min LC gradient. With the use of crude SIL peptides, we further applied the established MRM assay to assess quantitatively the differential expression of small GTPases in the frontal cortex regions of post-mortem brain tissue samples acquired from patients with different AD stages. Furthermore, we describe the proof-of-principle for targeted measurement of small GTPases in human brain tissues by performing relative quantitation employing crude SIL peptides as both internal standards (IS) and surrogate standards (SS). Our results led to the discovery of the altered expressions of several synaptic GTPase proteins, including Rab3A/C, Rab4A/B and Rab27B in higher AD stages, suggesting a potential role of synaptic trafficking in AD progression.

1. Blennow, K.; de Leon, M. J.; Zetterberg, H., Alzheimer's disease. *Lancet* **2006,** *368*, 387-403.
2. Assoc, A. s., 2018 Alzheimer's disease facts and figures. *Alzheimers Dement* **2018,** *14*, 367-425.
3. Spires-Jones, T. L.; Hyman, B. T., The intersection of amyloid beta and tau at synapses in Alzheimer's disease. *Neuron* **2014,** *82*, 756-71.
4. Zhang, J.; Goodlett, D. R.; Montine, T. J., Proteomic biomarker discovery in cerebrospinal fluid for neurodegenerative diseases. *J Alzheimers Dis* **2005,** *8*, 377-86.
5. Musunuri, S.; Wetterhall, M.; Ingelsson, M.; Lannfelt, L.; Artemenko, K.; Bergquist, J.; Kultima, K.; Shevchenko, G., Quantification of the brain proteome in Alzheimer's disease using multiplexed mass spectrometry. *J Proteome Res* **2014,** *13*, 2056-68.
6. Seyfried, N. T.; Dammer, E. B.; Swarup, V.; Nandakumar, D.; Duong, D. M.; Yin, L.; Deng, Q.; Nguyen, T.; Hales, C. M.; Wingo, T.; Glass, J.; Gearing, M.; Thambisetty, M.; Troncoso, J. C.; Geschwind, D. H.; Lah, J. J.; Levey, A. I., A Multi-network Approach Identifies Protein-Specific Co-expression in Asymptomatic and Symptomatic Alzheimer's Disease. *Cell Syst* **2017,** *4*, 60-72 e4.
7. Sathe, G.; Na, C. H.; Renuse, S.; Madugundu, A. K.; Albert, M.; Moghekar, A.; Pandey, A., Quantitative Proteomic Profiling of Cerebrospinal Fluid to Identify Candidate Biomarkers for Alzheimer's Disease. *Proteomics Clin Appl* **2018**, e1800105.
8. Lleo, A.; Nunez-Llaves, R.; Alcolea, D.; Chiva, C.; Balateu-Panos, D.; Colom-Cadena, M.; Gomez-Giro, G.; Munoz, L.; Querol-Vilaseca, M.; Pegueroles, J.; Rami, L.; Llado, A.; Molinuevo, J. L.; Tainta, M.; Clarimon, J.; Spires-Jones, T.; Blesa, R.; Fortea, J.; Martinez-Lage, P.; Sanchez-Valle, R.; Sabido, E.; Bayes, A.; Belbin, O., Changes in Synaptic Proteins Precede Neurodegeneration Markers in Preclinical Alzheimer's Disease Cerebrospinal Fluid. *Mol Cell Proteomics* **2019,** *18*, 546-560.
9. Begcevic, I.; Kosanam, H.; Martinez-Morillo, E.; Dimitromanolakis, A.; Diamandis, P.; Kuzmanov, U.; Hazrati, L. N.; Diamandis, E. P., Semiquantitative proteomic analysis of human hippocampal tissues from Alzheimer's disease and age-matched control brains. *Clin Proteomics* **2013,** *10*, 5.
10. Hondius, D. C.; van Nierop, P.; Li, K. W.; Hoozemans, J. J.; van der Schors, R. C.; van Haastert, E. S.; van der Vies, S. M.; Rozemuller, A. J.; Smit, A. B., Profiling the human hippocampal proteome at all pathologic stages of Alzheimer's disease. *Alzheimers Dement* **2016,** *12*, 654-68.
11. Andreev, V. P.; Petyuk, V. A.; Brewer, H. M.; Karpievitch, Y. V.; Xie, F.; Clarke, J.; Camp, D.; Smith, R. D.; Lieberman, A. P.; Albin, R. L.; Nawaz, Z.; El Hokayem, J.; Myers, A. J., Label-free quantitative LC-MS proteomics of Alzheimer's disease and normally aged human brains. *J Proteome Res* **2012,** *11*, 3053-67.
12. Zhang, Q.; Ma, C.; Gearing, M.; Wang, P. G.; Chin, L. S.; Li, L., Integrated proteomics and network analysis identifies protein hubs and network alterations in Alzheimer's disease. *Acta Neuropathol Com* **2018,** *6*.
13. Hales, C. M.; Dammer, E. B.; Deng, Q.; Duong, D. M.; Gearing, M.; Troncoso, J. C.; Thambisetty, M.; Lah, J. J.; Shulman, J. M.; Levey, A. I.; Seyfried, N. T., Changes in the detergent-insoluble brain proteome linked to amyloid and tau in Alzheimer's Disease progression. *Proteomics* **2016,** *16*, 3042-3053.

14. Ping, L.; Duong, D. M.; Yin, L.; Gearing, M.; Lah, J. J.; Levey, A. I.; Seyfried, N. T., Global quantitative analysis of the human brain proteome in Alzheimer's and Parkinson's Disease. *Sci Data* **2018,** *5*, 180036.
15. McKetney, J.; Runde, R. M.; Hebert, A. S.; Salamat, S.; Roy, S.; Coon, J. J., Proteomic Atlas of the Human Brain in Alzheimer's Disease. *J Proteome Res* **2019,** *18*, 1380-1391.
16. Shi, T. J.; Song, E. W.; Nie, S.; Rodland, K. D.; Liu, T.; Qian, W. J.; Smith, R. D., Advances in targeted proteomics and applications to biomedical research. *Proteomics* **2016,** *16*, 2160-2182.
17. Prieto-Dominguez, N.; Parnell, C.; Teng, Y., Drugging the Small GTPase Pathways in Cancer Treatment: Promises and Challenges. *Cells* **2019,** *8*.
18. Govek, E. E.; Hatten, M. E.; Van Aelst, L., The role of Rho GTPase proteins in CNS neuronal migration. *Dev Neurobiol* **2011,** *71*, 528-53.
19. Li, G., Rab GTPases, membrane trafficking and diseases. *Curr Drug Targets* **2011,** *12*, 1188-93.
20. Kiral, F. R.; Kohrs, F. E.; Jin, E. J.; Hiesinger, P. R., Rab GTPases and Membrane Trafficking in Neurodegeneration. *Curr Biol* **2018,** *28*, R471-R486.
21. Rodriguez, L.; Mohamed, N. V.; Desjardins, A.; Lippe, R.; Fon, E. A.; Leclerc, N., Rab7A regulates tau secretion. *J Neurochem* **2017,** *141*, 592-605.
22. Vaz-Silva, J.; Gomes, P.; Jin, Q.; Zhu, M.; Zhuravleva, V.; Quintremil, S.; Meira, T.; Silva, J.; Dioli, C.; Soares-Cunha, C.; Daskalakis, N. P.; Sousa, N.; Sotiropoulos, I.; Waites, C. L., Endolysosomal degradation of Tau and its role in glucocorticoid-driven hippocampal malfunction. *EMBO J* **2018,** *37*.
23. Kundra, R.; Ciryam, P.; Morimoto, R. I.; Dobson, C. M.; Vendruscolo, M., Protein homeostasis of a metastable subproteome associated with Alzheimer's disease. *Proc Natl Acad Sci U S A* **2017,** *114*, E5703-E5711.
24. Scheper, W.; Hoozemans, J. J.; Hoogenraad, C. C.; Rozemuller, A. J.; Eikelenboom, P.; Baas, F., Rab6 is increased in Alzheimer's disease brain and correlates with endoplasmic reticulum stress. *Neuropathol Appl Neurobiol* **2007,** *33*, 523-32.
25. Ginsberg, S. D.; Alldred, M. J.; Counts, S. E.; Cataldo, A. M.; Neve, R. L.; Jiang, Y.; Wuu, J.; Chao, M. V.; Mufson, E. J.; Nixon, R. A.; Che, S., Microarray analysis of hippocampal CA1 neurons implicates early endosomal dysfunction during Alzheimer's disease progression. *Biol Psychiatry* **2010,** *68*, 885-93.
26. Ginsberg, S. D.; Mufson, E. J.; Alldred, M. J.; Counts, S. E.; Wuu, J.; Nixon, R. A.; Che, S., Upregulation of select rab GTPases in cholinergic basal forebrain neurons in mild cognitive impairment and Alzheimer's disease. *J Chem Neuroanat* **2011,** *42*, 102-10.
27. Klassen, M. P.; Wu, Y. E.; Maeder, C. I.; Nakae, I.; Cueva, J. G.; Lehrman, E. K.; Tada, M.; Gengyo-Ando, K.; Wang, G. J.; Goodman, M.; Mitani, S.; Kontani, K.; Katada, T.; Shen, K., An Arf-like small G protein, ARL-8, promotes the axonal transport of presynaptic cargoes by suppressing vesicle aggregation. *Neuron* **2010,** *66*, 710-23.
28. 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* **2018,** *78*, 5431-5445.
29. Zauber, H.; Kirchner, M.; Selbach, M., Picky: a simple online PRM and SRM method designer for targeted proteomics. *Nat Methods* **2018,** *15*, 156-157.

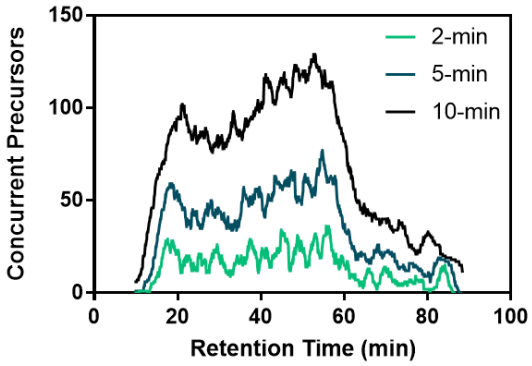
30. Sherwood, C. A.; Eastham, A.; Lee, L. W.; Risler, J.; Vitek, O.; Martin, D. B., Correlation between y-type ions observed in ion trap and triple quadrupole mass spectrometers. *J Proteome Res* **2009,** *8*, 4243-51.
31. Picotti, P.; Aebersold, R., Selected reaction monitoring-based proteomics: workflows, potential, pitfalls and future directions. *Nat Methods* **2012,** *9*, 555-66.
32. Kawahara, R.; Bollinger, J. G.; Rivera, C.; Ribeiro, A. C.; Brandao, T. B.; Paes Leme, A. F.; MacCoss, M. J., A targeted proteomic strategy for the measurement of oral cancer candidate biomarkers in human saliva. *Proteomics* **2016,** *16*, 159-73.
33. Escher, C.; Reiter, L.; MacLean, B.; Ossola, R.; Herzog, F.; Chilton, J.; MacCoss, M. J.; Rinner, O., Using iRT, a normalized retention time for more targeted measurement of peptides. *Proteomics* **2012,** *12*, 1111-21.
34. 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* **2003,** *100*, 6940-5.
35. Zhang, H.; Liu, Q.; Zimmerman, L. J.; Ham, A. J.; Slebos, R. J.; Rahman, J.; Kikuchi, T.; Massion, P. P.; Carbone, D. P.; Billheimer, D.; Liebler, D. C., Methods for peptide and protein quantitation by liquid chromatography-multiple reaction monitoring mass spectrometry. *Mol Cell Proteomics* **2011,** *10*, M110 006593.
36. Whiteaker, J. R.; Lin, C.; Kennedy, J.; Hou, L.; Trute, M.; Sokal, I.; Yan, P.; Schoenherr, R. M.; Zhao, L.; Voytovich, U. J.; Kelly-Spratt, K. S.; Krasnoselsky, A.; Gafken, P. R.; Hogan, J. M.; Jones, L. A.; Wang, P.; Amon, L.; Chodosh, L. A.; Nelson, P. S.; McIntosh, M. W.; Kemp, C. J.; Paulovich, A. G., A targeted proteomics-based pipeline for verification of biomarkers in plasma. *Nat Biotechnol* **2011,** *29*, 625-34.
37. Liebler, D. C.; Zimmerman, L. J., Targeted quantitation of proteins by mass spectrometry. *Biochemistry* **2013,** *52*, 3797-806.
38. Hoofnagle, A. N.; Whiteaker, J. R.; Carr, S. A.; Kuhn, E.; Liu, T.; Massoni, S. A.; Thomas, S. N.; Townsend, R. R.; Zimmerman, L. J.; Boja, E.; Chen, J.; Crimmins, D. L.; Davies, S. R.; Gao, Y.; Hiltke, T. R.; Ketchum, K. A.; Kinsinger, C. R.; Mesri, M.; Meyer, M. R.; Qian, W. J.; Schoenherr, R. M.; Scott, M. G.; Shi, T.; Whiteley, G. R.; Wrobel, J. A.; Wu, C.; Ackermann, B. L.; Aebersold, R.; Barnidge, D. R.; Bunk, D. M.; Clarke, N.; Fishman, J. B.; Grant, R. P.; Kusebauch, U.; Kushnir, M. M.; Lowenthal, M. S.; Moritz, R. L.; Neubert, H.; Patterson, S. D.; Rockwood, A. L.; Rogers, J.; Singh, R. J.; Van Eyk, J. E.; Wong, S. H.; Zhang, S.; Chan, D. W.; Chen, X.; Ellis, M. J.; Liebler, D. C.; Rodland, K. D.; Rodriguez, H.; Smith, R. D.; Zhang, Z.; Zhang, H.; Paulovich, A. G., Recommendations for the Generation, Quantification, Storage, and Handling of Peptides Used for Mass Spectrometry-Based Assays. *Clin Chem* **2016,** *62*, 48-69.
39. Sudhof, T. C., The synaptic vesicle cycle. *Annu Rev Neurosci* **2004,** *27*, 509-47.
40. Rizzoli, S. O., Synaptic vesicle recycling: steps and principles. *EMBO J* **2014,** *33*, 788-822.
41. Kokotos, A. C.; Peltier, J.; Davenport, E. C.; Trost, M.; Cousin, M. A., Activity-dependent bulk endocytosis proteome reveals a key presynaptic role for the monomeric GTPase Rab11. *Proc Natl Acad Sci U S A* **2018,** *115*, E10177-E10186.

42. Scheff, S. W.; Price, D. A.; Schmitt, F. A.; Mufson, E. J., Hippocampal synaptic loss in early Alzheimer's disease and mild cognitive impairment. *Neurobiol Aging* **2006,** *27*, 1372-84.
43. Pavlos, N. J.; Gronborg, M.; Riedel, D.; Chua, J. J.; Boyken, J.; Kloepper, T. H.; Urlaub, H.; Rizzoli, S. O.; Jahn, R., Quantitative analysis of synaptic vesicle Rabs uncovers distinct yet overlapping roles for Rab3a and Rab27b in Ca2+-triggered exocytosis. *J Neurosci* **2010,** *30*, 13441-53.
44. Takai, Y.; Sasaki, T.; Matozaki, T., Small GTP-binding proteins. *Physiol Rev* **2001,** *81*, 153-208.
45. Jahn, R., Principles of exocytosis and membrane fusion. *Ann N Y Acad Sci* **2004,** *1014*, 170-8.

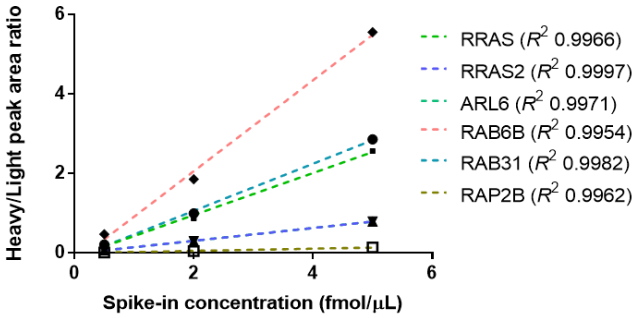
Figure 5.1 Sensitivity, multiplexing capability, reproducibility, and accuracy of the MRM-based quantification at the peptide level.

(A) Number of concurrent transitions scheduled in each cycle with a 2-min, 5-min and 10-min retention time window, respectively; **(B)** Linear regression of the peak area ratios (heavy/light) for 6 representative peptides obtained from three LC-MRM experiments with different spike-in concentrations. Crude SIL standards were spiked at 0.5, 2 and 5 fmol/μL, respectively. **(C)** Scatter plots depicting distribution of the peak area ratios (heavy/light) for all quantified peptides obtained from three LC-MRM experiments with different spike-in concentrations.

A


Concurrent Precursors
150
100
50
0
0
20
40
60
80
100
Retention Time (min)
2-min
5-min
10-min

B


Heavy/Light peak area ratio
6
4
2
0
0
2
4
6
Spike-in concentration (fmol/μL)
RRAS (R^2 0.9966)
RRAS2 (R^2 0.9997)
ARL6 (R^2 0.9971)
RAB6B (R^2 0.9954)
RAB31 (R^2 0.9982)
RAP2B (R^2 0.9962)

C

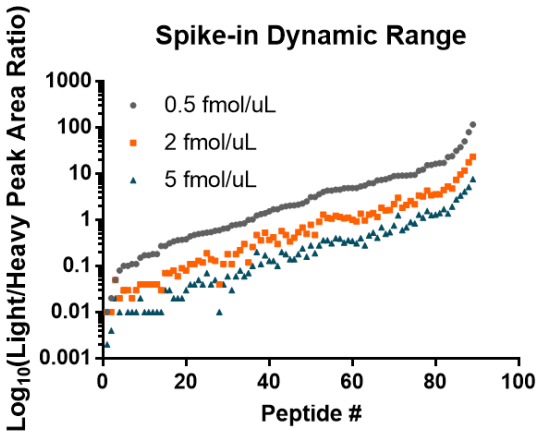

Spike-in Dynamic Range
Log10(Light/Heavy Peak Area Ratio)
1000
100
10
1
0.1
0.01
0.001
0
20
40
60
80
100
Peptide #
0.5 fmol/uL
2 fmol/uL
5 fmol/uL

Figure 5.2 Confirmation of the qualities of the crude stable isotope-labeled (SIL) peptides by MRM analyses.

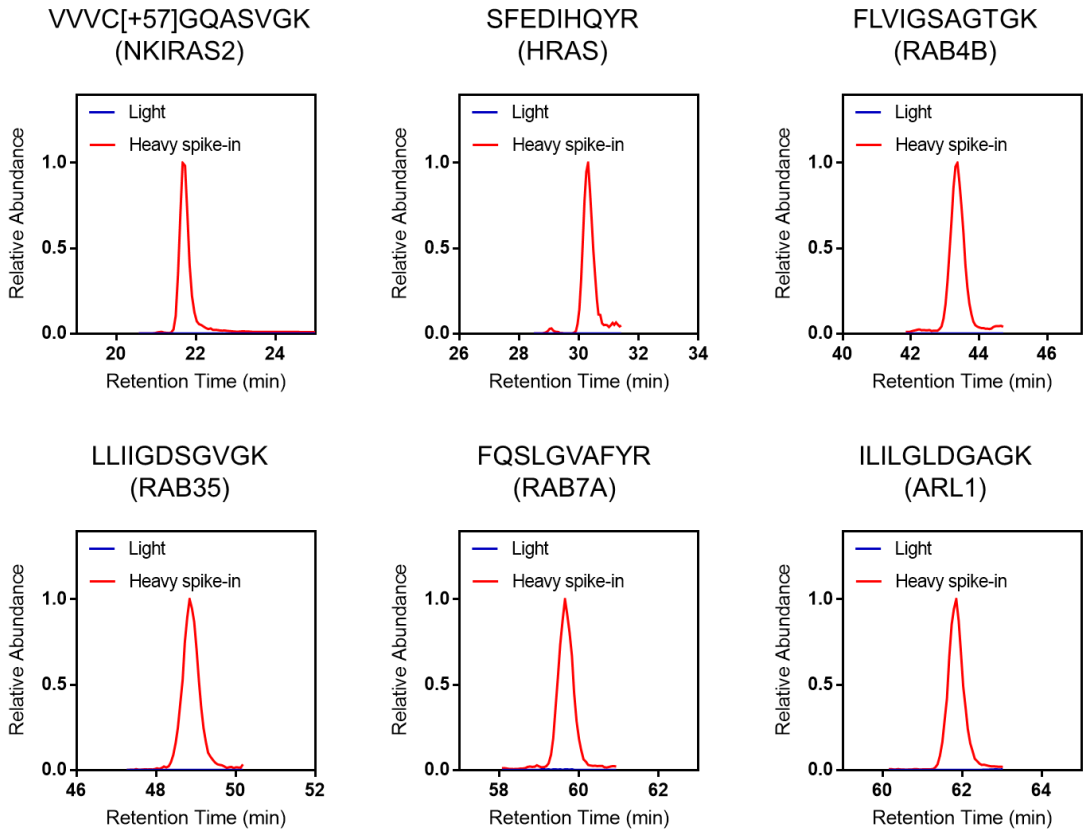
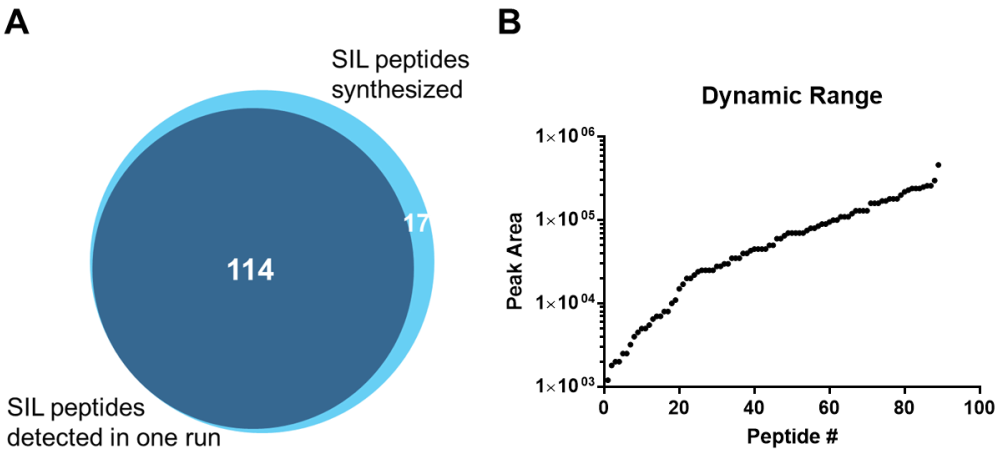


Figure 5.3 Coverage, dynamic range and spike-in concentration optimization for the crude SIL peptides in the MRM analyses.

(A) A Venn diagram showing the coverage of the SIL peptides detected in a single LC-MRM run; **(B)** A scatter plot depicting the dynamic range of the peak areas resulting from LC-MRM analysis of 2 fmol of crude SIL peptides; **(C)** A table describing the peak area ratios (light/heavy) obtained from three LC-MRM experiments.



C

	Light/Heavy Ratio			
Conc.	**Lowest**	**Median**	**Highest**	**Average**
0.5 fmol/μL	0.01	2.04	115	7.41
2.0 fmol/μL	0.005	0.34	23.4	1.71
5.0 fmol/μL	0.002	0.14	7.6	0.58

Figure 5.4 Selection for the twelve surrogate standard (SS) peptides across the designated retention time windows.

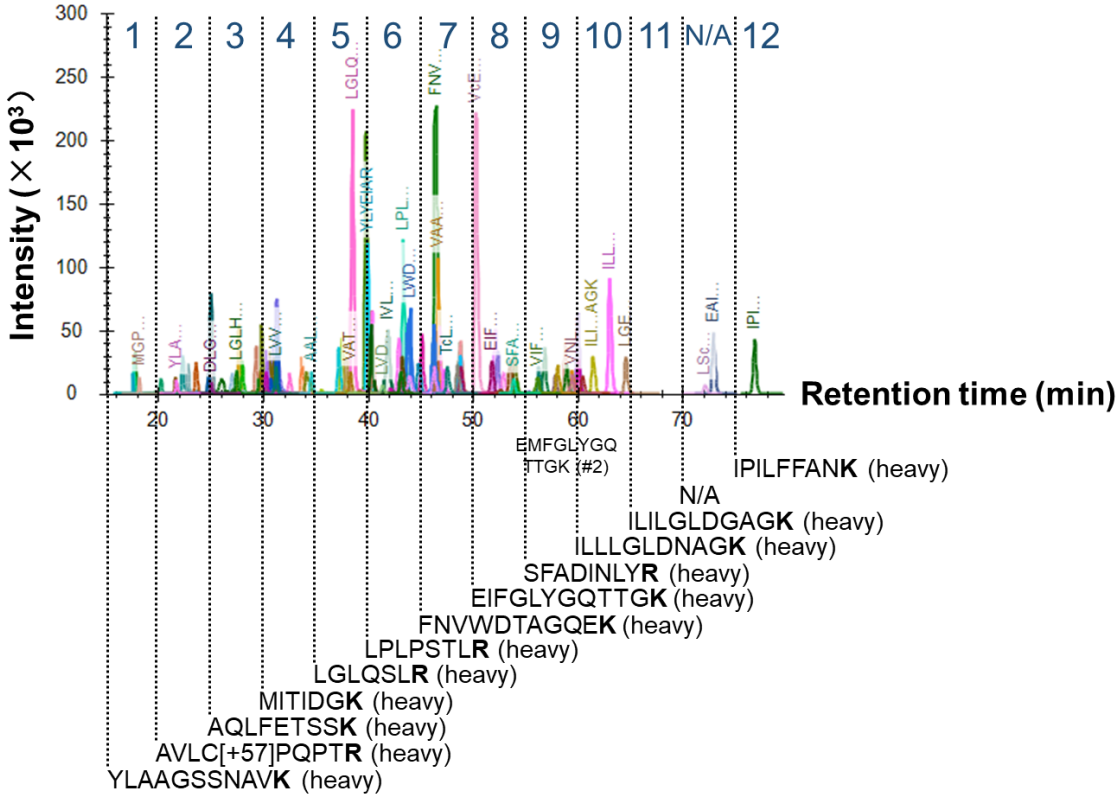


Figure 5.5 Comparison of the peak areas extracted for the selected SS peptides in the MRM analyses across similar or different retention time windows.

(A) A connected scatter plot illustrating the peak area variations across three SS peptides (#1, #2, #3) eluted at similar retention time windows; **(B)** A connected scatter plot illustrating the peak area variations across three SS peptides (#1, #7, #12) eluted at different retention time windows.

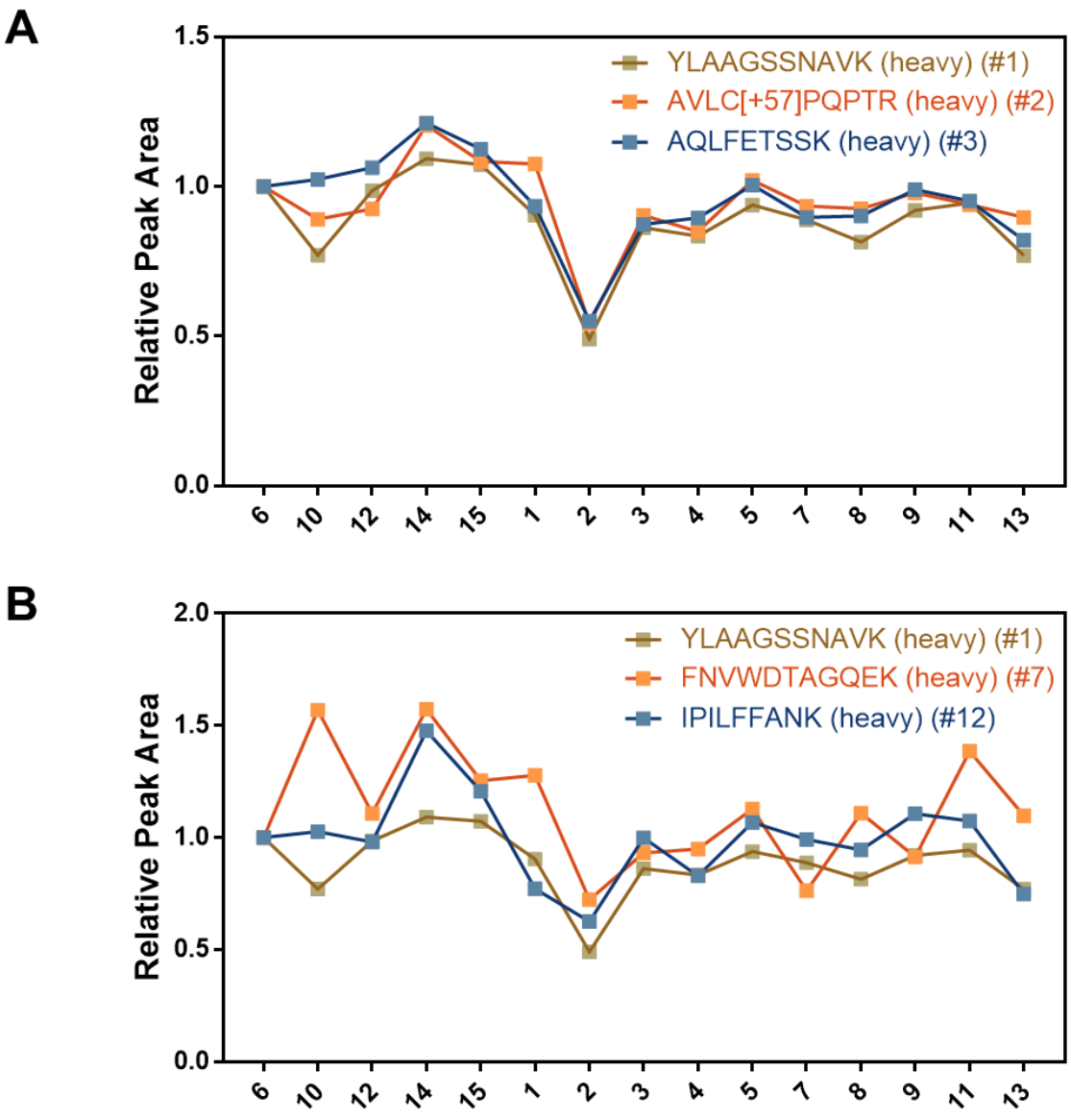


Figure 5.6 A schematic workflow illustrating sample preparation for LC-MRM analysis.

No AD (n = 5) **Intermediate AD (n = 5)** **High AD (n = 5)**

Snap-frozen human brain tissues from medial frontal cortex

↓

Tissue homogenization
Protein extraction

↓

SDS-PAGE followed by band excision of 15-37 kDa
In-gel reduction, alkylation, and tryptic digestion

↓

Peptide desalting and spiked with 2 fmol/µL crude SIL peptides

↓

Scheduled LC-MRM analysis and data processing

Figure 5.7 A heatmap showing the relative quantification of small GTPases in brain tissues obtained from Alzheimer's disease (AD) patients.

(A) A Venn diagram showing the overlap between quantified small GTPase, quantified SIL standards and total targeted small GTPases in the library; **(B)** A heatmap showing the relative quantification (in $\log_2$ scale) of small GTPases in brain tissues obtained from Alzheimer's disease (AD) patients.

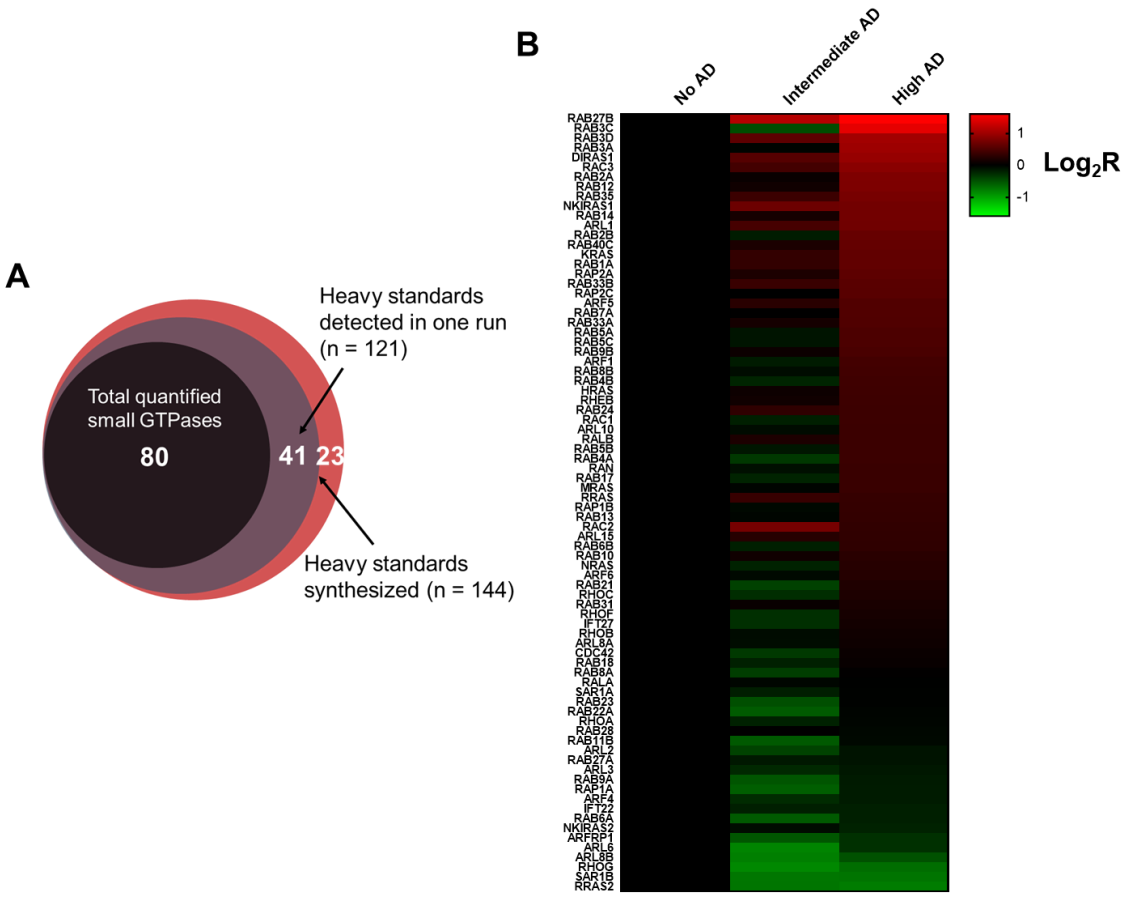


Figure 5.8 Representative MRM quantification results for three synaptic GTPases, Rab4A, Rab4B, and Rab10.

(A) A bar graph illustrating the MRM-based quantification results obtained from three peptides derived from Rab4A; **(B)** A box plot summarizing the quantification results in panel (A); **(C)** A bar graph illustrating the MRM-based quantification results obtained from two peptides derived from Rab4B; **(D)** A box plot summarizing the quantification results in panel (C). **(E)** A bar graph illustrating the MRM-based quantification results obtained from three peptides derived from Rab10; **(F)** A box plot summarizing the quantification results in panel (E). Tukey's multiple comparison test was performed to calculate the *p* values (#, $p > 0.05$; *, $0.01 \leq p < 0.05$).

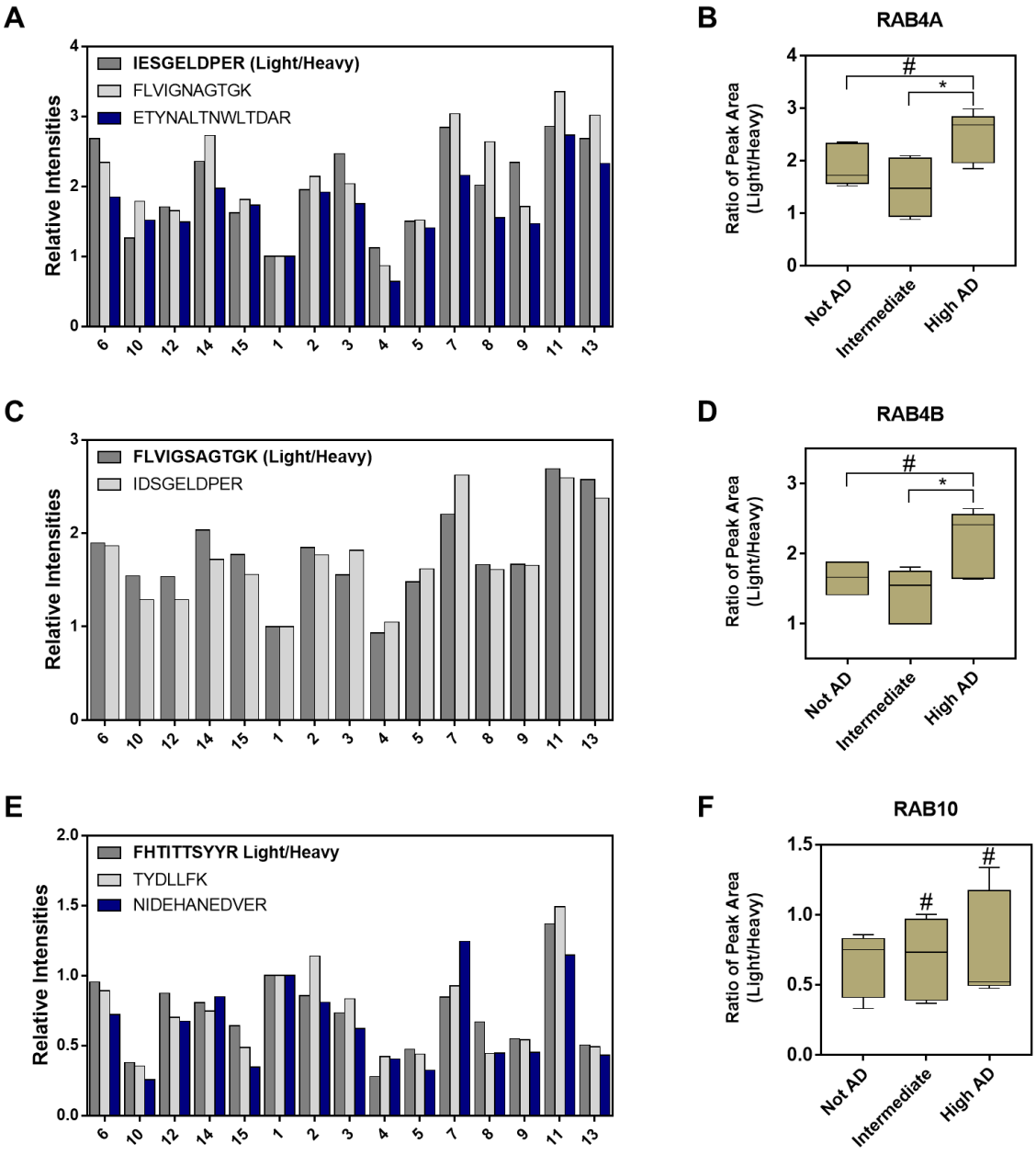

A
IESGELDPER (Light/Heavy)
FLVIGNAGTGK
ETYNALTNWLTDAR
Relative Intensities
6 10 12 14 15 1 2 3 4 5 7 8 9 11 13
B
RAB4A
Ratio of Peak Area (Light/Heavy)
Not AD
Intermediate
High AD
C
FLVIGSAGTGK (Light/Heavy)
IDSGELDPER
Relative Intensities
D
RAB4B
E
FHTITTSYYR Light/Heavy
TYDLLFK
NIDEHANEDVER
F
RAB10

Figure 5.9 Rab27B is up-regulated in higher stages of AD.

(A) A box plot summarizing the quantification of Rab27B levels among the three patient groups; **(B)** Western blot analysis of Rab27B in the 15 brain tissue samples; **(C)** A bar graph showing the quantification results obtained by western blot analysis and MRM analysis; **(D)** Linear regression between the quantification results obtained by western blot analysis and MRM analysis.

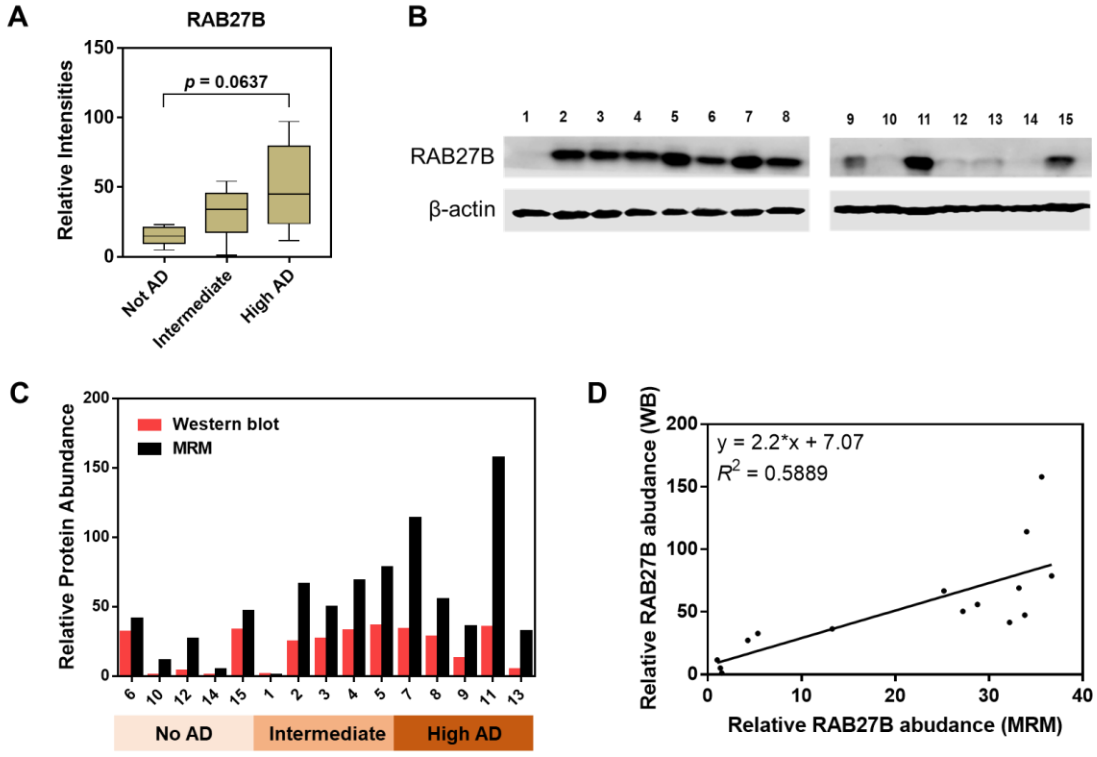
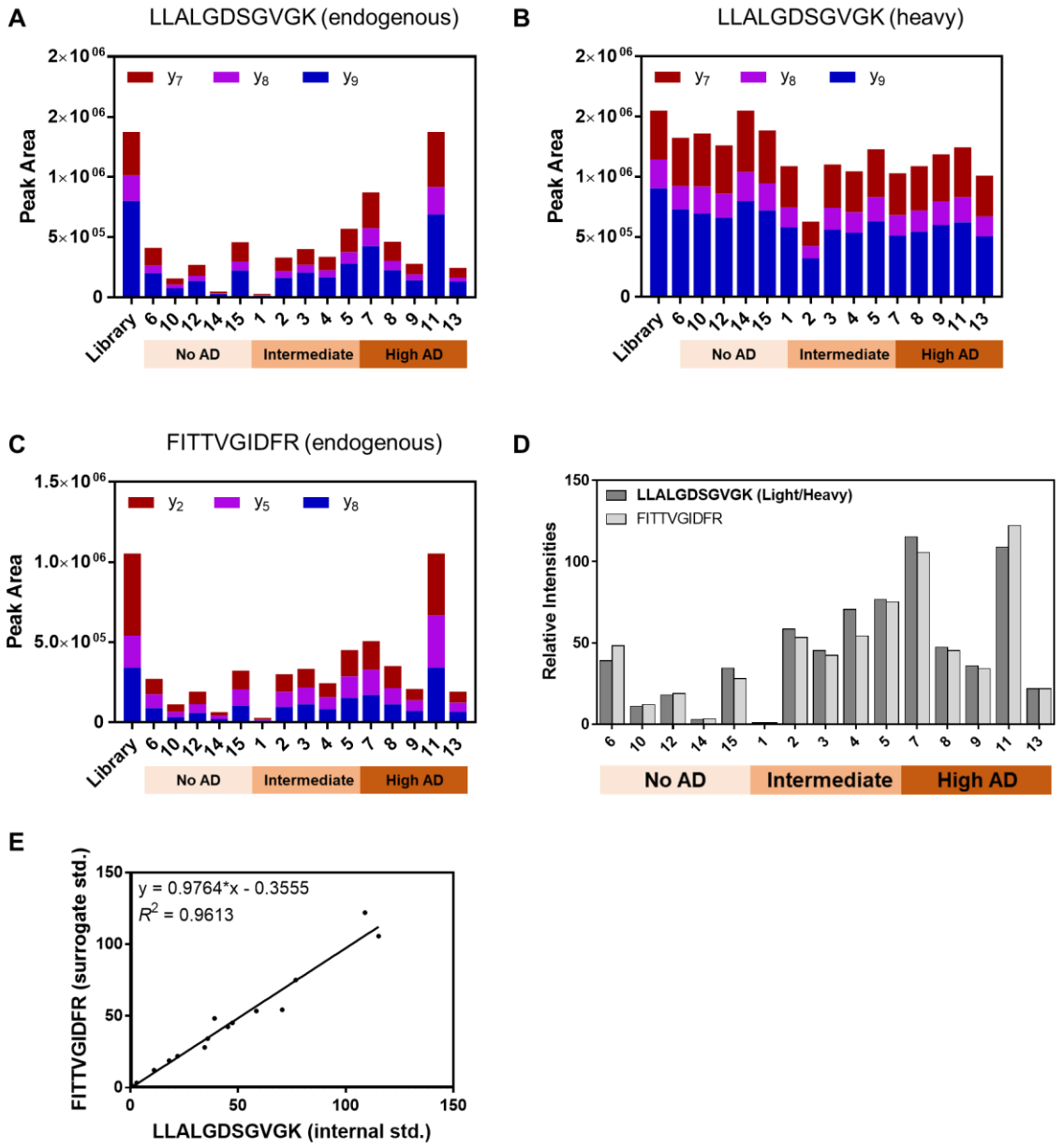


Figure 5.10 Quantification details for Rab27B by two normalization methods.

(A) The relative contribution of each transition to targeted MRM analysis of the endogenous LLALGDSGVGK peptide derived from Rab27B; **(B)** The relative contribution of each transition to targeted MRM analysis of the isotope-labeled LLALGDSGVGK peptide; **(C)** The relative contribution of each transition to targeted MRM analysis of the endogenous FITTVGIDFR peptide; **(D)** Bar charts of the quantification results normalized from IS peptide (LLALGDSGVGK) and SS peptide (FITTVGIDFR); **(E)** Linear regression of the quantification results shown in panel (D).



Chapter 6 Concluding Remarks

In this dissertation, we developed and utilized mass spectrometry-based approaches for targeted quantitative proteomic analysis of small GTPases of the Ras superfamily in cancer cells and tissue samples. By employing multiple-reaction monitoring (MRM) and two different labeling strategies, i.e. stable isotope labeling by amino acids in cell culture (SILAC) and crude synthetic stable isotope-labeled (SIL) peptides, we have successfully achieved high-throughput targeted quantitative profiling of small GTPases by assessing the altered protein expressions of these proteins related to cancer metastasis and drug resistance in cultured cancer cells and development of neurodegenerative diseases in patient-derived tissue samples.

Specifically, the work detailed in Chapter 2 describes the development of a novel targeted proteomic assay for profiling small GTPases in cultured human cells, which involved metabolic labeling by SILAC, fractionation by SDS-PAGE and scheduled MRM analysis. By using this scheduled Ge-LC-MRM method, we assessed the differential expression of small GTPases in paired primary/metastatic melanoma cell lines, which lead to reproducible quantitation of more than 90 small GTPases in two 157-min LC runs. Moreover, we validated quantification results by proteomic analysis by performing Western-blot analysis. Combined with bioinformatic analyses of publicly available patient data and cell-based assays, we uncovered a previously unrecognized role of small GTPase RAB38 in promoting melanoma metastasis *in vitro* and predicting poor melanoma patient prognosis. We further demonstrated that RAB38 promotes melanoma metastasis *in vitro* through the regulation of matrix metalloproteinases, and the increased expression of

RAB38 in metastatic melanoma cells arises from diminished promoter methylation and heightened binding of the MITF transcription factor.

In Chapters 3 and 4, we discuss the application of the MRM-based targeted proteomic approach established in Chapter 2 as a tool to probe the altered expression of small GTPases associated with acquired tamoxifen resistance in breast cancer and colorectal cancer metastasis. In Chapter 3, in total 96 small GTPases were robustly quantified in wild-type MCF-7 and the paired tamoxifen-resistant breast cancer cells, among which down-regulation of RAB31 was analyzed further. By conducting bioinformatic analyses and cell-based assays, we provided evidence that decreased expression of RAB31 predicted poor breast cancer patient survival and modulated development of acquired tamoxifen resistance *in vitro*. In Chapter 4, differential expression of small GTPases in paired primary/metastatic colorectal cancer cell (CRC) lines SW480 and SW620 was assessed, leading to robust quantitation of 83 small GTPases. Among them, we identified SAR1B as a potential suppressor and prognostic marker for CRC metastasis by combining the proteomic data with bioinformatic analyses of publicly available patient data. We also showed that diminished SAR1B expression could stimulate epithelial–mesenchymal transition (EMT), thereby promoting motility and *in vitro* metastasis of SW480 cells.

Although SILAC-based quantitation provides high accuracy and precision, many biological materials such as body fluids and tissue samples are not readily amenable to metabolic labeling. Therefore, in Chapter 5, we explored the analytical performance of scheduled Ge-LC-MRM analysis in conjunction with a different labeling strategy, which relies on the use of crude synthetic stable isotope-labeled (SIL) peptides. Furthermore, we

benchmarked the quantitative performance of crude SIL peptides as both internal standards and surrogate standards, allowing quantification of more peptides by the MRM assay. To investigate the link between small GTPases and progression of Alzheimer's disease (AD), we analyzed, by Ge-LC-MRM analysis, the total lysates of frontal cortex from post-mortem patient AD brain tissues Alzheimer's disease (AD) with varied AD pathology: "no AD" (n = 5), "intermediate AD" (n = 5), and "high AD" (n = 5). To this end, we quantified ~80 small GTPases from the 15 samples with high sensitivity and reproducibility in one 45-min LC run. Among the differentially expressed small GTPases, we observed substantial up-regulation of several synaptic GTPases in tissues with higher levels of AD, including Rab3A/C, Rab4A/B and Rab27B. Future experiments are warranted to provide mechanistic evidence of altered synaptic trafficking contributed by dysregulated small GTPases.

Together, by taking advantage of the above-mentioned methods, high-throughput and targeted quantitative profiling of small GTPases in cultured cancer cells and tissue samples is enabled. The work in this dissertation presented a novel and systematic targeted proteomic approach to investigate the previously unrecognized roles of dysregulated small GTPase expressions in a wide array of biological events relevant to public health, such as cancer progression, AD pathology and cellular response upon exposure to environmental toxicants.

MRM has been the standard for targeted proteomics during the past two decades with superior sensitivity, reproducibility and specificity; however, as with any other proteomic technique, it is highly dependent on sample preparation and instrument optimization and

more susceptible to background interference resulting from sample matrix. Future directions for targeted assay development by using parallel-reaction monitoring (PRM) are therefore highly encouraged. Extending the methods described in this dissertation to other members of GTP-binding proteins such as heterotrimeric G proteins is currently being explored.