MECHANISMS OF GENOME BUFFERING AND CELL FATE COORDINATION IN ADULT TISSUE HOMEOSTASIS

APPROVED BY SUPERVISORY COMMITTEE

Lawrence Lum, Ph.D.

Philipp Scherer, Ph.D.

James Amatruda, M.D./Ph.D.

Ralph Deberardinis, M.D./Ph.D.

DEDICATION

To my mother

MECHANISMS OF GENOME BUFFERING AND CELL FATE COORDINATION IN ADULT TISSUE HOMEOSTASIS

by

RUBINA TULADHAR

DISSERTATION / THESIS

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

August 2016

MECHANISMS OF GENOME BUFFERING AND CELL FATE COORDINATION IN ADULT TISSUE HOMEOSTASIS

Rubina Tuladhar, Ph.D.

The University of Texas Southwestern Medical Center at Dallas, 2016

Supervising Professor: Lawrence Lum, Ph.D.

Self-renewal competency of adult stem cells is essential for tissue homeostasis. The corruption of genes essential for genome preservation or for niche-stem cell interactions frequently results in loss of stem cell viability and disease. The two components of my thesis focus on understanding adult stem cell preservation – the integration of metabolism and intercellular communication mediated by the Wnt family of secreted signaling molecules, and

epigenetic mechanisms that buffer the proteome against insertion/deletion (INDEL)-type genetic mutations.

Wnt-mediated signaling is essential for embryogenesis and the maintenance of adult tissues. Lipidation of Wnt proteins by the acyltransferase Porcupine (Porcn) is crucial for secretory pathway exiting. Using chemically based approaches, I have demonstrated that Porcn active site features conserved across animals enforce ω-7 *cis* fatty acylation of Wnt proteins. Deviant acylation of a Wnt protein using an exogenously supplied *trans* fatty acid cripples its ability to traverse the secretory pathway due to a previously unappreciated stereoselectivity of the Wnt chaperone Wntless (WLS) for fatty acids. My findings provide a mechanistic account of chemical specificity observed in Porcn inhibitors, and delineate a universal mechanism for integrating communal cell fate decision-making with metabolic fitness.

As part of my efforts to generate isogenic cells for the expression of LKB1, a tumor suppressor that regulates Wnt protein production, I encountered the emergence of foreign LKB1 proteins subsequent to the introduction of INDELs by the DNA editing enzyme CRISPR-Cas9. I demonstrate that these novel proteins are the products of: a) the installation of internal ribosomal entry sites (IRES), b) the induction of exon skipping due to compromised exon splicing enhancers (ESEs), and c) the conversion of pseudo-mRNAs to protein-coding mRNAs due to the unwanted elimination of premature termination codons. I propose that these molecular events serve as compensatory mechanisms employed by cells to restore proteome integrity in the face of INDEL-type challenges to the genome posed by pathogens and environmental mutagens. Taken together, these two projects will: a) delineate intervention strategies premised upon the attack of an universally conserved point of intersection between metabolism and cell-to-cell

communication, b) facilitate the personalization of medicine, and c) accelerate tissue engineering initiatives.

TABLE OF CONTENTS

Chapter 2: A universal requirement for protein ω-7 *cis* fatty acylation in coordinated cell fate decision-making in animals

Chapter 3: Mutation in tumor suppressor gene LKB1 increases total protein secretion, including Wnt proteins

Chapter 4: Epigenetic mechanisms of genome buffering revealed by CRISPR-CAS9 failures

Chapter 5: Materials and methods

ACKNOWLEDGEMENTS

I must begin by expressing my sincere gratitude to my mentor, Dr. Lawrence Lum. His enthusiasm for scientific discussion and discovery is truly an inspiration to me. My scientific journey so far would not have been possible without the mentorship and support that I have received from him throughout my time in his lab. I am also grateful to the members of my thesis committee, Dr. James Amatruda, Dr. Philipp Scherer and Dr. Ralph Deberardinis, for their guidance and helpful suggestions.

I would also like to thank all the current and past members of the Lum lab for their support and friendship. I am especially grateful to Ozlem for the continuous insight and encouragement that she has provided me ever since my days as a rotation student in the lab. Finally, I would like to thank my husband, Amit Shrestha, who has given me tremendous moral and emotional support throughout my years in the graduate school. It would have been impossible to have completed this journey without his unwavering optimism and belief in me.

PRIOR PUBLICATIONS

1. **Tuladhar R**, Yarravarapu N, Lum L. Monitoring Wnt protein acylation using an in vitro cyclo-addition reaction. In Wnt signaling- Methods and Protocols. Springer, New York. In Press.
2. **Tuladhar R**, Lum L. Fatty acyl donor selectivity in membrane bound O-acyltransferases and communal cell fate decision-making. Biochem. Soc. Trans. 2015 Apr; 43(2) 235-9.
3. Fan CW, **Tuladhar R**, Lum L. Signaling: Making a leaner Hedgehog. Nat Chem Biol. 2013 Apr; 9(4):217-8.
4. Wang X, Moon J, Dodge ME, Pan X, Zhang L, Hanson JM, **Tuladhar R**, Ma Z, Shi H, Williams NS, Amatruda JF, Carroll TJ, Lum L, Chen C. The development of highly potent inhibitors for porcupine. J Med Chem. 2013 Mar 28;56(6):2700-4.
5. **Tuladhar R**, Lum L. Chemical disruption of Wnt protein fatty acylation as an anti-cancer therapeutic strategy. In: Proceedings of the 104th Annual Meeting of the American Association for Cancer Research; 2013 Apr 6-10; Washington, DC. Philadelphia (PA): AACR; Cancer Res 2013;73.
6. Bhat, S., Sorci-Thomas, M.G., **Tuladhar, R**., Samuel, M. & Thomas, M. (2007). Conformational Adaptation of Apolipoprotein A-I to Discretely Sized Phospholipid Complexes. *Biochemistry, 46* (26), 7811-7821.

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

LIST OF ABBREVIATIONS

ACAT- Acyl-CoA cholesterol acyltransferase

APC- Adenomatous polyposis coli

ATI- Alternate translation initiation

BP1- β propeller 1

CBP- CREB binding protein (CBP)

CHX- Cycloheximide

Ck1α- Casein kinase 1 alpha

crRNA- CRISPR RNA

CTD- C-terminal domain

DGAT- Diacylglycerol acyltransferase

Disp- Dispatched

DMD- Duchenne's muscular dystrophy

DSB- Double strand break

Dvl- Dishevelled

EJC- Exon junction complex

ER- Endoplasmic reticulum

ESEs- Exonic splicing enhancers

ESSs- Exonic splicing silencers

FL- Firefly luciferase

Fzd- Frizzled

GOAT- Ghrelin O-acyl acyltransferase

GSK3β- Glycogen synthase kinase 3 beta

Hh- Hedgehog

HHAT- Hedgehog acyltransferase

HHATL- Hedgehog acyltransferase-Like

INDELs- insertions/deletions

IRES- Internal ribosomal entry site

IRES- Internal ribosomal entry sites

ISEs- Intronic splicing enhancers

ISSs- Intronic splicing silencers

IWP- Inhibitor of Wnt production

IWR- Inhibitor of Wnt response

LPA- Lysophosphatidic acid

LPC- Lysophosphatidylcholine

LPE- Lysophosphatidylethanolamine

LPEAT- Lysophosphatidylethanolamine acyltransferase

LPI-Lysophosphatidylinositol

LPIAT- Lysophosphatidylinositol acyltransferase

LPLAT- Lysophospholipid acyltransferase

LPS- Lysophosphatidylserine

Lrp- Lipoprotein receptor related protein

MBOAT- Membrane bound O-acyltransferase

MEFs- Mouse embryonic fibroblasts

MVB- Multivesicular body

NHEJ- Non-Homologous End Joining

NL- NanoLuc luciferase

NMD- Nonsense mediated decay

NTD- N-terminal domain

ORFs- Open reading frames

PARP- poly(ADP)-ribose polymerase

PJS- Peutz-Jeghers Syndrome

Porcn- Porcupine

PTC- Premature termination codon

Ptch- Patched

RNF43- Ubiquitin E3 ligase ring finger 43

SA/SD- Splicing acceptor/donor

SCD- Stearoyl-CoA Desaturase

sFPR- Frizzled-related protein

SNP- Single nucleotide polymorphism

SR proteins- Serine-Arginine proteins

SRE- Splicing regulatory elements

SS- Signal sequence

Sufu- Suppressor of fused

Swim- Secreted Wingless-interacting molecule

Tnks- Tankyrase (Tnks)

tracrRNA- Transactivating RNA

Wg- Wgless

WIF- Wnt inhibitory protein

WLS- Wntless

Wnt1- Wingless-type mouse mammary tumor virus integration site family member 1

ZNRF3- Zinc and Ring Finger 3

CHAPTER ONE
Introduction and Literature Review

Wnt fatty acylation

The Wnt signaling pathway is an evolutionarily conserved signaling system important for embryonic development and the maintenance of adult tissue homeostasis [1, 2]. Using MMTV retrovirus-mediated insertional mutagenesis strategy, the Int-1 gene was first uncovered as an oncogene in mouse mammary tumors [3]. Int-1 was subsequently found to be identical to the *Drosophila* wingless (Wg) gene that is crucial for the formation of metathoracic imaginal discs and for segment polarity [4]. The amalgamation of Wg and Int-1 led to the coining of the term Wnt1 (wingless-type mouse mammary tumor virus integration site family member 1) and a new field of Wnt biology was established [5].

Three decades after the initial discovery of the first Wnt gene, genome sequencing efforts have now revealed the existence of nineteen Wnt genes in humans [6]. Wnt ligands are secreted proteins that are post-translationally modified with glycosylation and acylation. Wnt proteins are N-glycosylated at multiple sites and the number of glycosylation attachments varies among different Wnts [7]. Mutation of glycosylated residues does not compromise Wnt secretion [8] or Wnt activity [9]. Unlike glycosylation, fatty acylation of Wnt proteins is indispensible for the production of active Wnt ligands. In the absence of the fatty acyl modification, Wnts are trapped in the endoplasmic reticulum (ER) and unable to transverse from the secretory pathway.

Fatty acylation of Wnt proteins is mediated by the evolutionarily conserved ER-resident multi-transmembrane enzyme Porcn [10]. Porcn is the founding member of the membrane bound O-acyltransferase (MBOAT) gene family which consists of 16 polytopic acyltransferases with established lipid and protein substrates (Table 1) [11]. In addition to Wnt proteins, there are two other characterized lipidated extracellular proteins: Hedgehog (Hh) and Ghrelin. The Hedgehog acyltransferase (HHAT) lipidates Hh proteins wih palmitoyl-CoA and the Ghrelin O-acyl transferase (GOAT) acylates ghrelin with octanoyl-CoA [11, 12]. In case of the Wnt proteins, Porcupine (Porcn) acyltransferase was believed to lipidate Wnts at two conserved residues (Cys77 and Ser209 in Wnt3a) based on site-specific mutagenesis and mass spectrometric analysis [13, 14]. However, using a click chemistry based assay (Figure 1.1 A), we found that Wnt3a is acylated at a single serine residue (Figure 1.1 B). The first crystal structure of a Wnt protein (Wnt8 from *Xenopus laevis*) in complex with Frizzled 8 (Fzd8) published by Janda et al. also established the presence of a fatty acyl moiety on the Ser187 in xPorcn (equivalent to Ser209 in Wnt3a) [15]. Furthermore, the crystal structure revealed the engagement of Cys55 in Wnt8 (equivalent to Cys77 in Wnt3a) in a disulfide bond thus decreasing the likelihood of a second acylation site in Wnt proteins.

Function	**Gene**	**Substrate**	**Acyl donor**
Lipid biosynthesis	ACAT1, ACAT2	Cholesterol	Long chain Acyl-CoA
Lipid biosynthesis	DGAT1	Diacylglycerol	Long chain Acyl-CoA
Protein acylation	PORCN	Wnt	Palmitoleoyl-CoA
Protein acylation	HHAT	Hedgehog	Palmitoyl-CoA
?	HHATL/ MBOAT3	?	?
Protein acylation	GOAT/ MBOAT4	Ghrelin	Octanoyl-CoA
Phospholipid remodeling pathway	LPEAT1/ MBOAT1	LPS	Oleoyl-CoA

Phospholipid remodeling pathway	LPLAT5/ MBOAT5	LPS and LPC	Polyunsaturated Fatty Acyl-CoA
Phospholipid remodeling pathway	LPLAT2/MBOAT2	LPE and LPA	Oleoyl-CoA
Phospholipid remodeling pathway	LPIAT1/ MBOAT7	LPI	Arachidonoyl-CoA

Table 1: Diverse cellular functions of MBOATs dictated by fatty acyl donor and substrate selectivity. Abbreviations: ACAT, Acyl-CoA cholesterol acyltransferase; DGAT, Diacylglycerol acyltransferase ; GOAT, Ghrelin O-acyl acyltransferase; HHATL, Hedgehog acyltransferase-Like; LPEAT, Lysophosphatidylethanolamine acyltransferase; LPLAT; Lysophospholipid acyltransferase; LPIAT, Lysophosphatidylinositol acyltransferase; LPE, Lysophosphatidylethanolamine; LPS, Lysophosphatidylserine; LPI, Lysophosphatidylinositol; LPA, Lysophosphatidic acid; LPC, Lysophosphatidylcholine.

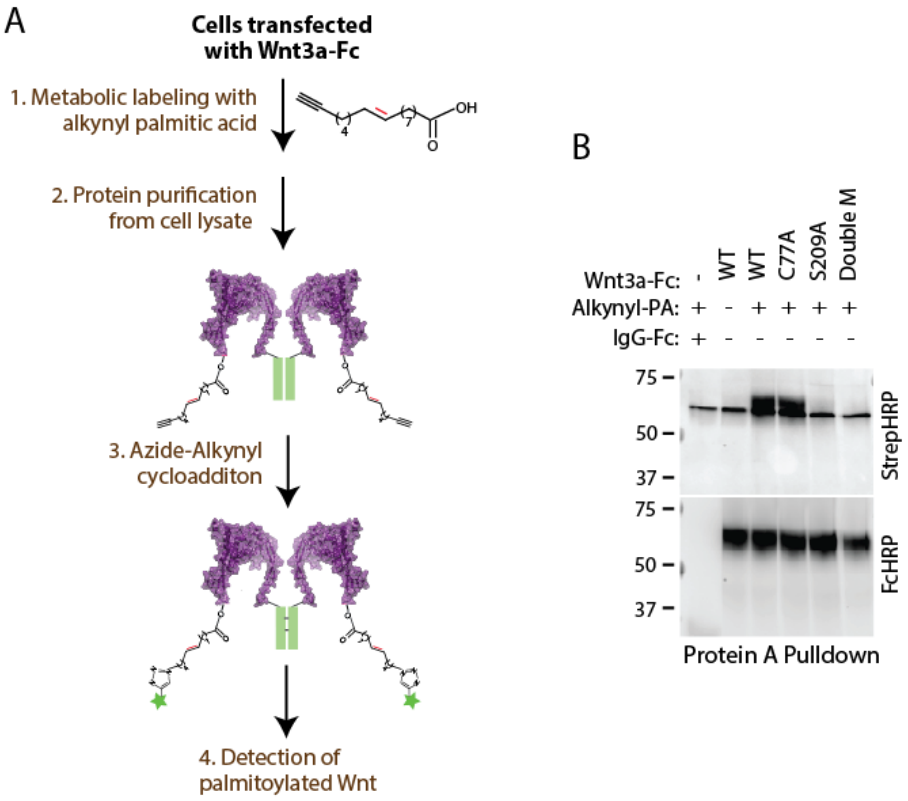


Figure 1. 1: Wnt3a protein is acylated at a highly conserved serine residue. (A) A modified click chemistry assay for monitoring Wnt protein acylation status. (B) The labeling of Wnt3a protein with an alkynyl palmitate probe is abolished for Wnt3aFc Ser209 mutant. On the other

hand, mutation of cysteine 77 residue to alanine did not compromise the fatty acyl labeling of Wnt3a.

Mechanism of Wnt signal transduction

Palmitoleation of Wnt ligands represent a regulatory checkpoint that is crucial for the movement of the Wnt molecules along the secretory pathway. A highly conserved multispan transmembrane protein Wntless (WLS) binds to Wnts in the ER and chaperones Wnts to the cell surface [16-18] (Figure 1.2). WLS is essential for the secretion of all Wnts except for the atypical drosophila WntD, which lacks the serine acylation site and fails to bind to WLS [19]. WLS is thought to bind to Wnts through β-strand stretches also known as N-terminal lipocalin-like domain [20]. Genome-wide RNAi screening studies for genes that regulate Wg secretion in drosophila also revealed p24 proteins as important cargo receptors for Wg anterograde shuttling route [21, 22]. Once the Wnt-WLS complex reaches the plasma membrane, the mechanism of how Wnts are released to the extracellular space had been confounding due to the hydrophobic nature of the Wnt proteins with lipid modification. Several secreted Wnt binding co-factors have been found to facilitate the extracellular solubilization of Wnts. Secreted Wingless-interacting molecule (Swim), a putative member of the lipocalin family of extracellular transport proteins, binds to Wg with nanomolar affinity in a palmitate-dependent manner and helps to maintain the solubility and activity of the Wg proteins [23]. Another protein that binds to Wnts with high affinity and helps to maintain the Wnt activity in water-soluble form is the serum glycoprotein afamin/ α-albumin [24]. Besides binding to co-factors to shield the adducted lipid in an aqueous environment, recent studies have also found Wnt-WLS complex to secrete on exosomes both during drosophila development and in human cells [25]. Exosomes are small secreted vesicles with 40-100 nm diameter that originates from multivesicular body (MVB) and are capable of

carrying miRNA, mRNA, proteins and lipid components [26, 27]. Thus, exosome-mediated secretion of active Wnts represents an alternative method of Wnt secretion independent of lipoprotein carriers. Once the Wnt is released from the cells, due to a conserved endocytosis motif in the third intracellular loop of WLS, WLS is endocytosed from the plasma membrane and the retromer complex to the golgi and the ER and the recycled WLS participates in the secretion of new batch of lipidated Wnts present in the ER [28-32].

In the absence of Wnt ligands, a cytoplasmic complex comprising of scaffolding protein Axin, tumor suppressor Adenomatous polyposis coli (APC) along with protein kinases GSK3β (Glycogen synthase kinase 3 beta) and Ck1α (Casein kinase 1 alpha) recruits and phosphorylate beta-catenin, which leads to its ubiquitination by E3 ligases and subsequent proteasome mediated degradation [33-35] (Figure 1.2 B). With the destruction of beta-catenin in the cytoplasm, its nuclear binding partners TCF/LEF transcription effectors fail to initiate target gene transcription [36, 37]. On the other hand, when a Wnt ligand reaches the surface of a Wnt-responding cell, it binds to the seven transmembrane GPCR like Frizzled receptors [38, 39] along with the low-density lipoprotein related protein (Lrp) co-receptors [40]. The formation of the Wnt ligand and the receptor complex results in the recruitment of the Dishevelled (Dvl) signaling proteins along with Axin and APC at the receptor level, which in turn disassembles the beta-catenin destruction complex [41-44]. The abrogation of beta-catenin degradation allows the cytoplasmic beta-catenin to translocate to the nucleus, bind with TCF/LEF proteins and activate the transcription of beta-catenin target genes [45-47].

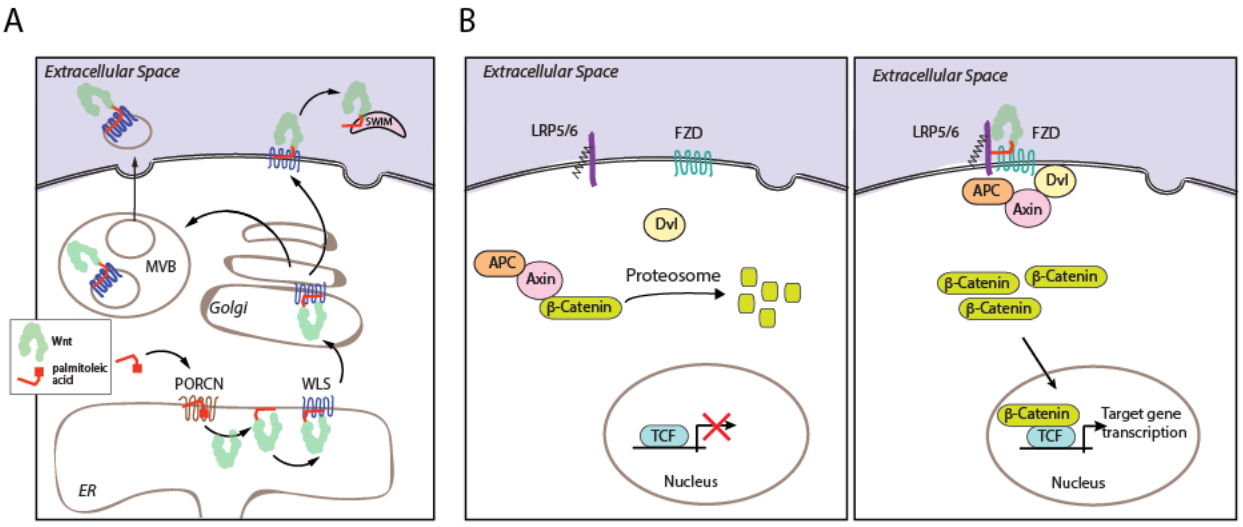


Figure 1.2: The Wnt/beta-catenin signal transduction pathway. (A) Wnt ligand biogenesis. Porcn modifies Wnt ligands with palmitoleic acid in the ER. Acylated Wnt protein is chaperoned by the WLS multipass receptor to the cell surface. The mature Wnt is released into the extracellular space either bound to lipoprotein carriers or secreted in exosomes. (B) In the absence of Wnt ligands, a cytoplasmic destruction complex consisting of Axin, APC, Ck1α (not shown) and Gsk3β (not shown) binds to and phosphorylates beta-catenin, which leads to its proteasome-mediated degradation. Alternatively, when the secreted Wnt protein reaches the cell surface of a Wnt-responding cell, it binds to Fzd receptor and Lrp coreceptor and recruits Dvl, Axin and APC proteins to the membrane, thus dismantling the beta-catenin destruction complex and allowing beta-catenin to bind to TCF/LEF transcription factors in the nucleus thus ultimately initiating the target gene transcription.

Drugging the Wnt pathway

The importance of Wnt signaling in development was first recognized with the discovery of segment polarity gene wingless in *Drosophila*. Wg was shown to be required for proper wing formation [48, 49] and ectopic expression of the Wg gene resulted in the duplication of the embryonic axis [50]. Since then, numerous studies have demonstrated the requirement of Wnt signaling in a broad range of organ systems including the gut, bone skin, brain, eye, spinal cord, lung, liver, kidney, pancreas and others [6, 7, 51]. Beyond development, Wnt signaling is also

crucial for maintaining cellular homeostasis and regeneration in self-renewing adult tissues with rapid turnovers such as hair [52], skin [53], intestine [54] and the hematopoietic systems [55, 56]. Mutations in the Wnt pathway that disrupt this homeostatic balance give rise to the development of pathological conditions, including carcinogenesis. The association of the Wnt pathway and human cancer was first made when tumor suppressor APC (Adenomatous Polyposis Coli) and the Wnt pathway component β-catenin were found to interact with each other [57, 58].

Mutations in APC were already known to cause familial adenomatous polyposis (FAP), a disease characterized by the occurrence of multiple intestinal polyps with high disposition to colon cancer, as well colorectal cancer [59, 60]. In fact, somatic mutations of the APC gene are found in more than 90% of colorectal cancer cases [61, 62]. In addition to APC, several other Wnt pathway components including Axin, beta-catenin and Tcf4 are frequently mutated in colorectal cancer [60, 62, 63], which highlights the significance of this pathway for the maintenance of intestinal tissue homeostasis. Activation of the Wnt signaling pathway due to the loss of function mutation of WTX, which encodes a component of the beta-catenin destruction complex results in Wilms tumor [64]. Increase in the stabilization of beta-catenin is also linked to aggressive fibromatosis and pulmonary fibrosis [65]. De-regulation of the Wnt signaling pathway due to the mutations in the upstream pathway component have been associated to many cancers. Mutations at the ligand level have been linked to polycystic kidney diseases and leukaemia [65]. Loss-of-function mutation in ZNRF3/RNF43 (Zinc and Ring Finger 3/ Ubiquitin E3 ligase ring finger 43) transmembrane ubiquitin ligase gene, which promotes the turnover of the Lrp6 and Frizzled Wnt receptors, is frequently found in cystic pancreatic cancer [66, 67].

Given the involvement of Wnt signaling in tissue homeostasis and a broad range of diseases mentioned above, ways to gain control of this pathway holds tremendous promise in the field of regenerative medicine as well as cancer therapeutics.

Most of the clinically approved drugs fall into two broad categories: small molecules, which typically contain <100 atoms and have molecular mass <1000 Da, and biologics, which include antibodies, peptides and vaccines [68]. For the longest time, Wnt pathway was considered "undruggable" due to the lack of typical pharmacological targets [69]. The existence of 19 Wnt ligands, 10 frizzled (Fzd) receptors, 2 lipoprotein receptor related protein (Lrp) co-receptors, 3 Dishevelled (Dvl) effectors and 4 Tcf/lef transcription factors also gives rise to the redundancy problems when targeting a single component of the pathway [70]. In addition to the beta-catenin dependent *canonical* Wnt signaling, the less established beta-catenin independent *non-canonical* Wnt signaling has its own set of receptors and effectors, which further makes targeting the pathway a daunting task. However, after decades of studying the pathway, we now have promising therapeutics, both small molecule-based and biologics-based, for targeting the Wnt signaling system.

Targeting the Wnt pathway at the ligand level

There are several molecular targeted agents that disrupt the Wnt pathway at different levels of the signaling cascade. Small molecules that disrupt the fatty acylation of the Wnt proteins by directly targeting the Porcn acyltransferase (known as Inhibitors of Wnt production or IWPs) were first discovered in 2009 by Chen et al [71]. Since acylation is likely to be important for maintaining the activities of all 19 Wnts, Porcn inhibitors promises to be effective in treating Wnt ligand mediated diseases regardless of the involvement of a multitude of Wnt

ligands. Inactivating mutations in transmembrane Ubiquitin E3 ligase ring finger 43 (RNF43) that promotes the turnover of Wnt receptors Fzd and Lrp6 [67] are frequently identified in cystic pancreatic tumors [66, 72]. Since loss of RNF43 increases Wnt ligand-dependent signaling, Porcn inhibitors could hold clinical utility in this disease context. LGK974 (also known as Wnt974), a Porcn inhibitor developed by Novartis, is currently being studied in phase 2 clinical trials for patients with metastatic head and neck squamous cell carcinoma caused due to the up-regulation of the Wnt pathway (NCT02649530).

Secreted Frizzled-related protein (sFRP) and Wnt inhibitory protein (WIF) both bind to Wnt proteins [73, 74], which leaves the Fzd receptors unoccupied and the Wnt signaling pathway inactivated. Using the same strategy of occupying the Wnt ligands, OncoMed and Bayer co-developed OMP-54F28 or Ipafricept, which is a fusion protein consisting of a portion of the Fzd8 receptor and Fc domain of a human IgG1 that binds to Wnt ligands thereby preventing the activation of the Wnt signaling cascade. OMP-54F28 is currently in Phase 1b clinical testing for various cancers (OMP-54F28 clinicaltrials.gov). In addition to Wnt ligand binding biologics, Oncomed had also developed monoclonal antibody OMP-18R5 or Vantictumab that interacts with five Fzd receptors through a conserved epitope with the extracellular domain and blocks the Wnt signaling induced by several Wnt ligands. OMP-18R5 has been found to be effective in inhibiting growth of various tumor types in xenograft models [75] and is currently in Phase 1 clinical trial (Vantictumab clinicaltrials.gov).

Inhibiting various cytoplasmic modulators of the Wnt signaling pathway

The chemical screen that led to the discovery of first in class Porcn inhibitors (the IWPs) also yielded another class of compounds IWR (Inhibitor of Wnt Response) that induces the

stabilization of Axin proteins [71]. IWR and another compound called XAV939 (discovered by Novartis [76]) both directly target the Tankyrase (Tnks) enzymes, two members of the poly(ADP)-ribose polymerase (PARP) gene family, which prevents the ubiquitination and degradation of Axin and in turn blocks the beta-catenin mediated transcriptional activation.

Dvl is an important cytoplasmic modulator of the Wnt signaling pathway that interacts with the DIX domain of Axin and recruits it to the plasma membrane thereby preventing beta-catenin degradation [43, 77]. Using NMR spectroscopy, PDZ domain of Dvl was also found to directly interact with Fzd receptors through the conserved motif (KTXXXW) located two amino acids after the seventh transmembrane domain [78]. Chemical compounds that bind to Dvl proteins and interfere with the assembly of the signalosome at the receptor level have the potential to block Wnt signaling. Several compounds such as NSC668036, Fj9 and 3289-8625 that interrupt the binding of Dvl PDZ to the Fzd receptors have been identified through structure-based virtual ligand screening and found to inhibit Wnt signaling in cell culture, xenograft mouse model and *Xenopus* embryos [79-81].

The constitutive degradation of beta-catenin by the destruction complex helps to maintain low levels of cytoplasmic beta-catenin and suppresses Wnt signaling. Any agent that aids in the stabilization of this destruction complex could be useful for inhibiting the Wnt/beta-catenin pathway. The protein kinases Gsk3β and Ck1α present in the destruction complex are essential for phosphorylating and targeting beta-catenin for proteasome-mediated degradation [82]. Pyrvinium, an FDA-approved antihelminth, was identified as a Ck1α activator in a high-throughput screen using *Xenopus* egg extracts [83]. In addition to inhibiting the Wnt pathway by allosteric activation of Ck1α, pyrvinium has also demonstrated to improve cardiac remodeling in

the mouse model of myocardial infarction [84]. Although several compounds that inhibit the Wnt signaling pathway by targeting the cytoplasmic effectors (as mentioned above) have shown promising pre-clinical results, the challenges associated with achieving higher potency and specificity have hindered the development of these compounds for clinical testing.

Targeting the nuclear effectors of the Wnt pathway

Mutations in several upstream components of the Wnt pathway lead to the accumulation of nuclear beta-catenin and binding to TCF/LEF1 transcription co-activators that ultimately activate the transcription of beta-catenin target genes. Inhibiting the pathway activity at the transcription level by blocking beta-catenin-TCF/LEF1 binding represents a point of attack that will likely bypass many well established paths to drug resistance such as mutations in APC. However, due to the high binding affinity between these two proteins (~20nM) and a lack of hydrophobic binding pockets, disrupting the protein-protein interaction with a small molecule remains a challenge [85]. Several high-throughput screens using *in vitro* competition assay, cell-based reporter gene assay as well as *in silico* modeling of compound libraries have yielded small molecule inhibitors that hinders the beta-catenin-TCF/LEF1 complex formation [86-89]. These compounds also maintain Wnt-inhibitory functions in secondary assays such as inhibition of colon cancer cell proliferation and beta-catenin-induced axis duplication of *Xenopus* embryos [87]. However, the use of these compounds in clinical testing has not advanced mainly due to the broad activity profiles and the failure to identify the molecular targets of the chemical compounds. Instead of directly disrupting the beta-catenin and TCF/LEF1 interaction, targeting the beta-catenin transcriptional co-activators is another strategy to gain chemical tractability of the Wnt pathway. ICG-001 and PRI-724 are two small molecule inhibitors that target the

transcriptional co-activator CREB binding protein (CBP) [90]. The use of small molecule inhibitors that target a co-activator protein with multiple molecular partners raises questions about its utility as a specific Wnt antagonist. Despite these concerns PRI-724, developed by Prism Pharma and Eisai Pharmaceuticals, has entered Phase I clinical trial for treating advanced solid tumors and advanced myeloid malignancies (NCT01302405 and NCT01606579).

The role of Stearoyl-CoA Desaturase (SCD) in Wnt biogenesis

Although generally assumed, the lipid adduct identity has only been determined in a single mammalian Wnt protein by mass spectrometry [14]. Despite the high cellular ratio of saturated palmitoyl-CoA to unsaturated palmitoleoyl-CoA [91], the use of the less abundant palmitoleoyl-CoA by Porcn for Wnt acylation suggests a likely mechanism whereby palmitoyl-CoA is first converted into palmitoleoyl-CoA prior to immobilization by Porcn onto Wnt proteins. Since the intracellular levels of monounsaturated fatty acids (MUFA) are controlled by Stearoyl-CoA Desaturase (SCD), the potential involvement of SCD in Wnt biogenesis would seem likely.

SCD is a tetraspan ER-resident enzyme with both the N and C termini located in the cytosol [92]. SCD catalyzes the introduction of a double bond at the cis ω-7 position (also known as *cis*-delta-9 position) in the presence of NADH-dependent flavoprotein cytochrome b5 reductase, electron acceptor cytochrome b5 and oxygen. There are four SCD isoforms found in mice (SCD1, SCD2, SCD3 and SCD4) whereas only two SCD isoforms (SCD1 and SCD5) have been identified in humans. Human SCD1 shares 85% homology with the four murine SCDs and are ubiquitously expressed while the expression of SCD5 is primarily limited to the brain and pancreas [92]. Since the studies involving global knockout of SCD1 gene in a mouse model

suggested protection of the animals against diet-induced and leptin deficiency-induced obesity, hepatic steatosis and insulin resistance [93, 94], several classes of SCD1 inhibitors have been developed as potential metabolic syndrome therapeutics. Using RNAi and commercially available small molecule inhibitors against SCD1, the requirement for *de novo* fatty acid desaturation in Wnt biogenesis and activity was interrogated. Treatment with SCD1 inhibitor as well as RNAi-mediated knockdown of SCD resulted in ablation of Wnt acylation (Figure 1.3 A) thus establishing the essential role of the SCD enzymatic activity in Wnt protein production. SCD1-mediated *de novo* desaturation is not required for the acylation of sonic hedgehog (Shh) and ghrelin proteins (Figure 1.3 B and Figure 1.3 C). Moreover, a MUFA modification likely exists in all Wnt proteins (Figure 1.3 D).

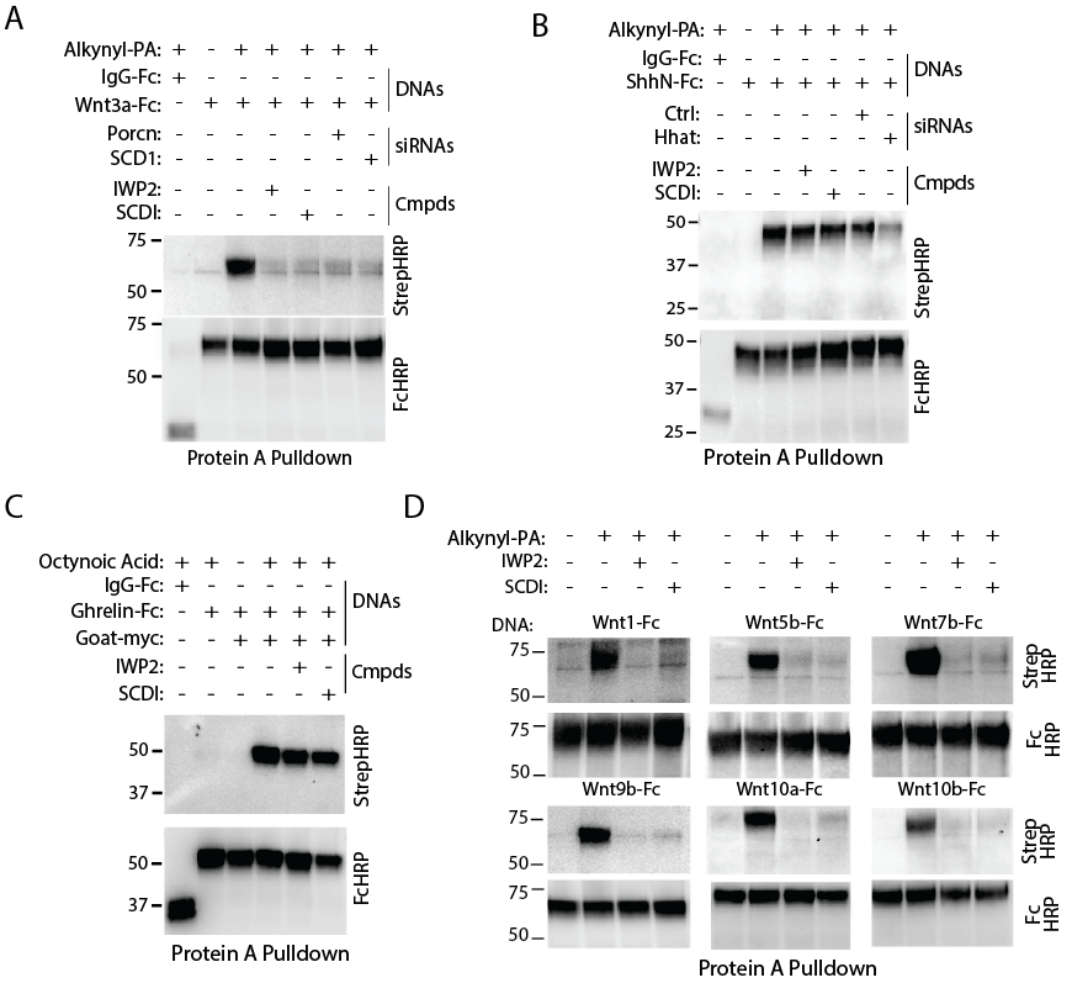


Figure 1.3: SCD is essential for Wnt ligand biogenesis. (A) Stearoyl-CoA desaturase (SCD) is required for Wnt acylation. Wnt3a labeling with 16-carbon alkynyl palmitate is blocked in the presence of Porcn inhibitor (IWP2) and SCD inhibitor (SCDI) at 5μM concentration. RNAi mediated gene knockdown of Porcn and SCD also recapitulated the loss of Wnt3a labeling. (B) and (C) IWP2 and SCDI does not inhibit Shh and Ghrelin fatty acylation. The same click chemistry strategy is used to monitor the fatty acylation of a Shh-Fc and Ghrelin-Fc fusion proteins. (D) Multiple Wnt proteins depend on SCD for desaturated fatty acyl substrates required for acylation. A collection of Wnt fusion proteins analyzed demonstrated the requirement for SCD enzyme for their acylation.

SCD inhibitors: utility for managing Wnt-dependent cancers

With *in vitro* acylation assays such as click chemistry and radioactive metabolic labeling assays [95], we and others have demonstrated the ability of SCD inhibitors to block the incorporation of exogenously supplied palmitate to Wnt proteins. Although these results suggest the importance of *de novo* fatty acid synthesis for Wnt production, consideration should also be given to other sources of palmitoleate that could potentially be used to acylate Wnts. Given that western diet can contain 34% to 36% total energy as fat with approximately one-third MUFA [93], palmitoleate from dietary sources could be a major contributor of fatty acid species for Wnt lipidation. In case of cultured cells, we have seen both *de novo* fatty acid synthesis as well as exogenous fatty acid species from the growth medium as substantial sources of fatty acid pool used for Wnt acylation. Dietary fatty acids, in addition to directly contributing MUFA for Wnt lipidation, also controls the expression of genes associated with fatty acid esterification, including SCD1 [96]. The availability of multiple sources of fatty acyl substrates and the regulation of SCD1 gene expression by various factors such as dietary lipids and hormones hinders the utility of SCD inhibitors as inhibitors of Wnt production. In addition, despite protecting mice from symptoms of metabolic syndrome, SCD inhibitors were also found to increase systemic inflammation and atherosclerosis, thus questioning the utility of these chemicals as viable therapeutic strategy in humans.

Divergent roles of fatty acylation in Hh and Wnt signaling

The Wnt and Hh proteins embody a nearly universal strategy for coordinating cell fate decision-making during development and tissue homeostasis in metazoans. The lipid adducts found on these proteins have largely been attributed roles as molecular brakes that prevent

protein diffusion and instead promote their graded distribution from ligand producing cells. The mature Hh protein is dually lipidated with a palmitate and cholesterol molecule (Figure 1.4 A). Whereas Hhat affixes the palmitoyl adduct onto the N-terminal cysteine residue, the regulatory domain of the Hh precursor protein facilitates a cholesterol-mediated intramolecular attack to release an N-terminal signaling domain [97] (Figure 1.4 A). As a result, the cholesterol molecule is covalently attached to the C-terminus of the newly generated signaling domain. Substituting the cysteine residue targeted for acylation with an alanine modestly reduces Hh protein ability to elicit cellular responses in vivo [12] and does not decrease Hh affinity for the Hh receptor Patched (Ptch) in vitro [98] suggesting acylation is not essential to its engagement of Ptch. Loss of Hhat then may compromise the ability of cholesterol-modified Hh to be lifted from the cell membrane by an enzyme called Dispatched (Disp) or to multimerize thus crippling the long range movement of Hh protein [99-101].

An engineered Wnt molecule with a transmembrane anchor can substitute for the wild-type protein in development suggesting that the monounsaturated fatty acyl adduct found on all 19 Wnt molecules restricts their signaling to adjacent cells and prevents their long range signaling [102]. Additionally distinguishing the role of fatty acylation in Wnt and Hh signaling is the affinity of the Wnt chaperone (WLS) for only lipidated Wnt molecules thus imposing a checkpoint on Wnt ligand biosynthesis that is absent in the case of Hh molecules [16, 18, 103] (Figure 1.4 B). The Frizzled (Fzd) receptors constitute a large family of seven transmembrane proteins that transduce Wnt signals across the membrane [104]. In contrast to the Ptch receptor which appears agnostic to Hh acylation status, the Fzd proteins harbor a pocket for a long chain

fatty acid and contributes to Wnt binding [15]. Thus, the fatty acyl modification is essential at several stages in the Wnt protein life cycle.

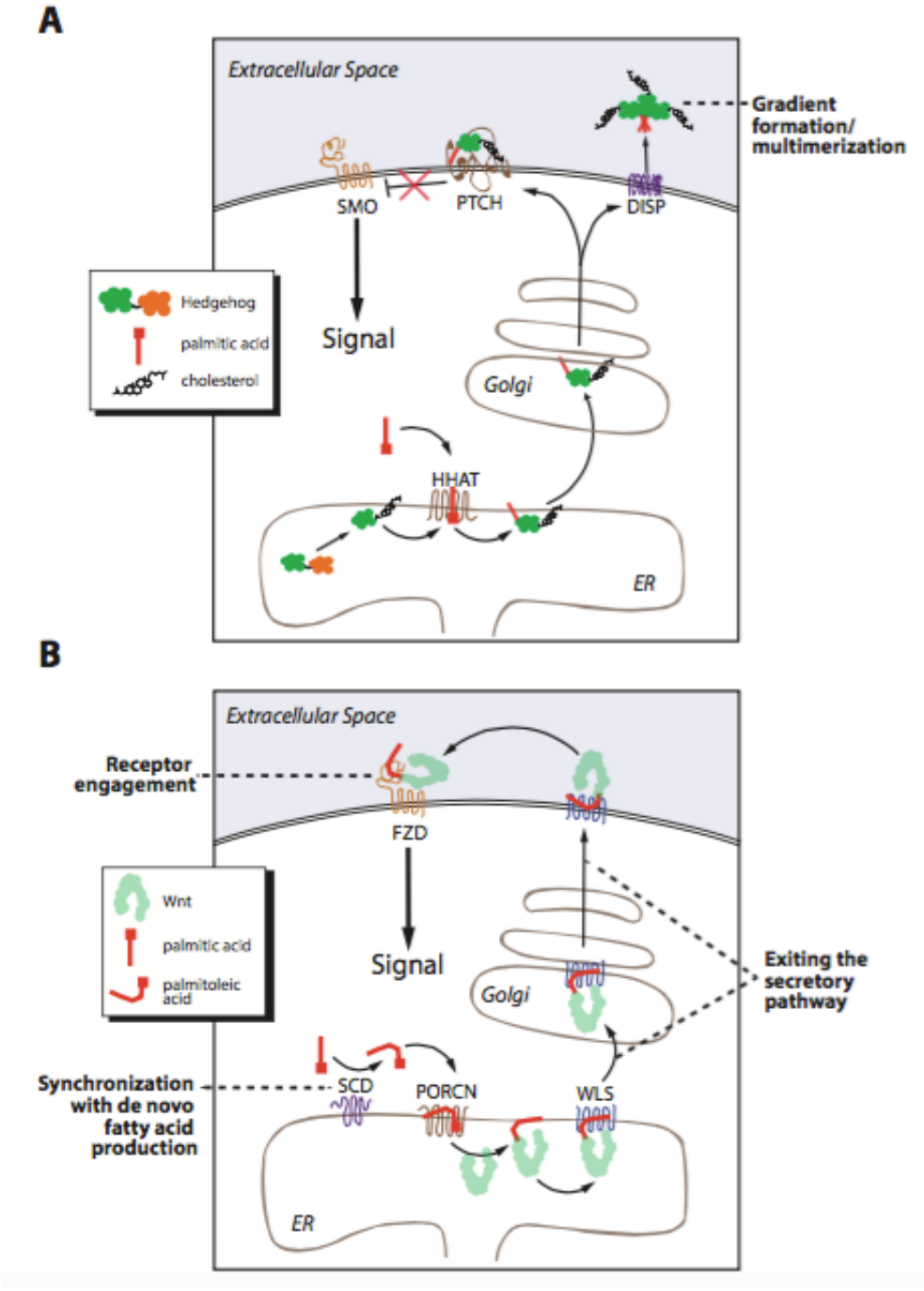


Figure 1.4: The role of fatty acylation in Wnt and Hh signaling. (A) Hedgehog (Hh) molecules are dually lipidated with a cholesterol adduct on the C-terminal domain and a palmitate on the N-terminal domain. Acylated Hh proteins are released from the membrane by the multipass protein Dispatched (Disp) and form multimeric complexes, which enable long range signaling and support formation of graded Hh responses in cells surrounding the ligand producing cells. Hh binding to the Patched (Ptch) receptor lifts the inhibitory action of Ptch on Smoothened (Smo) thus resulting in transcriptional activation of Hh target genes. (B) Porcn utilizes a palmitoleoyl-CoA donor to modify Wnt molecules. Sources of palmitoleoyl-CoA include Stearoyl-CoA Sesaturase (SCD)-mediated desaturation of palmitoyl-CoA. Acylated Wnt protein is chaperoned by the Wntless (WLS) multipass receptor to the cell surface. Secreted Wnt proteins then bind to Frizzled (Fzd) receptors with the acyl moiety inserting into a groove found in the extracelluar cysteine-rich domain.

Coupling of fatty acid metabolism with cell fate outcomes

The hydrophobic nature of the Wnt proteins due to the posttranslational addition of a fatty acid adduct makes the purification of Wnts for crystallization studies extremely difficult. Analysis of the primary sequences of Wnts also does not resemble any known protein folds. Due to these technical reasons, there was no structural information for Wnts until recently Chis Garcia's group crystalized XWnt8 in complex with the cysteine-rich domain (CRD) of mouse Fzd8 at 3.5 A° resolution [15]. This study revealed a unique two-domain structural architecture of Wnts with two interaction sites for Fzd proteins (Figure 1.5). The structure of XWnt8 resembled a "hand" that binds to Fzd8 through its "thumb" (Site 1) and "index finger" (Site 2). The lipid adduct is contained towards the end of the "thumb" region located in the N-terminal domain (NTD) of XWnt3a and is involved in Fzd8 CRD binding thorough a hydrophobic groove. The C-terminal cysteine-rich domain (CTD) of XWnt3a also binds to mFzd8, which further strengthens the Wnt-Fzd interaction. Although not immediately obvious, closer inspection revealed the structure of XWnt8 resembled the fusion of two known structural folds:

the four-helical core structure of XWnt8 NTD superimposed with saposin-like proteins and the long twisted b-hairpin in CTD superimposed with a cytokine-like domain [105]. Saposins are a diverse family of ancient proteins that can interact with lipids and also can serve as lipid presenters for receptors [106]. The remarkable similarities between the structures of XWnt8 and saposins and the importance of lipids in both the cases suggest a potential evolutionary origin of Wnts from the saposins. Speculatively, the Wnts could have utilized the inherent lipid affinity of the saposin-like fold to engage a lipid-interacting Fzd receptor [107]. Therefore, the incorporation of fatty acids in Wnt protein biogenesis could have potentially been an important first step that delineated a universal mechanism for integrating metabolic currency with communal cell fate decision-making.

The participation of SCDs in the production of Wnt proteins [95] may also support a mechanism for coupling cell intrinsic fatty acid metabolism with the coordination of cell fate outcomes during development and tissue homeostasis. Indeed, stem cells and cancer initiating cells appear to rely more on *de novo* fatty acid desaturation than scavenging pathways for maintaining homeostatic levels of desaturated phospholipids that are essential for membrane fluidity [108, 109]. Remarkably, chemical disruption of SCD provokes transcriptional up-regulation of many Wnt family members in induced pluripotent stem cells thus further providing evidence for the coupling of cellular metabolism and tissue homeostatic renewal [108].

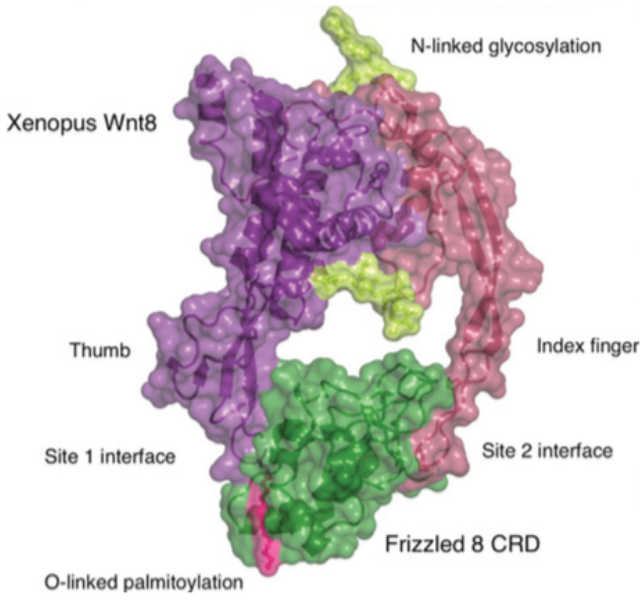


Figure 1.5: The structure of *Xenopus* Wnt8 co-crystallized with mouse Fzd8-CRD. XWnt8 interacts with mFzd8 at two separate binding sites located at the "thumb" region and the "index finger" region. The "thumb" region is comprised of a palmitoyl adduct which binds to the lipid-interacting groove found in mFzd CRD. Derived from Janda, 2015 [105].

CHAPTER TWO

A universal requirement for protein ω-7 *cis* fatty acylation in coordinated cell fate decision-making in animals

Introduction

Multicellular organisms that generate specialized tissues and organs must balance the needs of its constituent cells with those of the collective. Mechanisms that support the integration of metabolic processes and cell-cell communication provide a direct means for synchronizing efforts that meet these demands. For example, the immobilization of lipids onto proteins that partake in cell signaling can dictate protein distribution in the intra- and extracellular milieu, or directly gate protein activity thus influencing cell-to-cell communication.

The Wnt family of secreted proteins control cell fate decision-making in diverse tissues during development and homeostatic renewal. Animals typically express multiple Wnt molecules that require lipidation for their secretory pathway transit and ultimate release into the extracellular milieu [110]. Amongst the three characterized lipidated extracellular proteins, Wnt proteins are unique in their subjection to this production checkpoint. This phenomenon involves WLS, which fails to ferry Wnt proteins to the extracellular space in the absence of Porcn activity. How WLS distinguishes between acylated and naked Wnt proteins remains unknown. The fatty acyl adduct that gets attached to Wnts are also unique since to date, Wnt is the only known secreted protein that is modified with a monounsaturated lipid. Given the lipid adduct identity has only been determined in a single mammalian Wnt protein it remains unclear if it varies depending upon species-specific metabolic considerations or if evolutionary constraints have

eliminated lipid diversity. Indeed, the mechanistic basis for Wnt production dependency on monounsaturated fatty acylation remains elusive.

The recent advances of Porcn inhibitors in clinical testing has galvanized efforts to gain structural insights into the basis of chemical specificity, and more broadly a mechanistic understanding of why loss of lipidation cripples Wnt production [111, 112]. A crystal structure of Porcn will continue to be elusive given the multipass features of Porcn and the additional hydrophobicity contributed by fatty acylation [113]. Here, using a cross-species Porcn interrogation platform and chemical probes that target the Porcn active site, we reveal Porcn harbors features conserved across diverse animals that enforce monounsaturated fatty acid modification of Wnts. Using an exogenous trans fatty acids, we examined whether the basis of the single fatty acyl donor preference is attributed to the stereoselectivity of WLS for *cis* fats thus resulting in the omnipresence of these Porcn active site features in animals. At the same time, our studies provide structural insights into how Porcn inhibitors in clinical testing as anti-cancer agents are likely to achieve their specificity and selective biological activity.

MBOAT substrate and acyl donor specificity

Hhat and Porcn appear to occupy unique roles across multiple species thus signifying a selectivity of each enzyme for their cognate substrates. Indeed, no compensatory mechanisms have so far been identified that enable recovery of Wnt or Hh lipidation upon targeted disruption of either Porcn or Hhat respectively [12, 114, 115]. Clearly contributing to enzyme-substrate recognition is the presence of protein-protein interaction determinants in both MBOATs and substrates [116, 117] (Figure 2.1). At the same time, despite having access to both palmitoyl- and palmitoleoyl-CoA molecules, Wnt and Hh proteins are nevertheless predominantly modified

with a single fatty acyl species [14, 98]. Thus, mechanisms underpinning fatty acyl donor selectivity for MBOATs must also be in place. Whereas Porcn exhibits some predilection for fatty acids with a certain chain length (C10-16) [95, 113], this alone cannot account for Porcn acyl donor selectivity given that Wnt proteins appear to be mostly modified with palmitoleate. Thus, the active site of Porcn must accommodate an acyl donor with a double bond-induced bend that is 9 carbon lengths distal to the coenzyme A moiety. At the same time, the lack of fatty acyl donor selectivity observed *in vitro* and *in vivo* [118] for Hhat when provided with excess fatty acid analogues suggests additional mechanisms enforcing selectivity exist. The relative cytoplasmic abundance of palmitoyl-CoA compared to other acyl-CoA species in addition to active site determinants that limit stearate may dictate Hhat selectivity *in vivo* [119]. Although the mechanistic basis underlying these selectivity issues clearly require further examination, these early studies focused on Porcn and Hhat have already revealed general strategies for achieving fatty acylation selectivity that are likely employed by the MBOAT family at large.

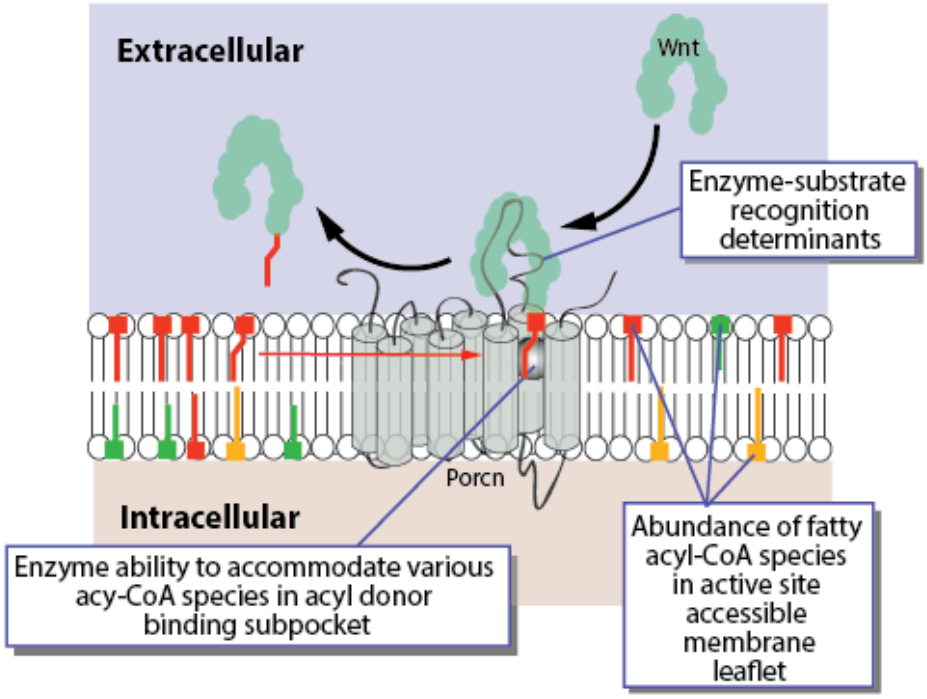


Figure 2.1: Mechanisms underlying MBOAT acyl donor selectivity. Wnt fatty acylation by the Porcn enzyme is modeled here.

Porcn active site features enforces exclusive modification of Wnt proteins with ω-7 fatty acid

Although animals typically harbor multiple Wnt genes, only a single Wnt protein produced from an isolated cell line has been subjected to mass spectrometric analysis in order to identify the adducted lipid [14]. At the same time, inhibitors of stearoyl co-A desaturase (SCDI) which prevent Wnt labeling with exogenously provided palmitate [95] are inconsistent in their activity against Wnt signaling. Although the SCDI is a potent inhibitor of Wnt activity in H23 lung cancer cell line as indicated by the disappearance of phosphorylated Dvl2 protein, it is ineffective to block cell autonomous Wnt signaling in Hela cervical cancer cell line (Figure 2.2). This discrepancy in SCDI efficacies against different cellular context suggests the potential

existence of lipids other than palmitoleic acid that may be incorporated in Wnt proteins.

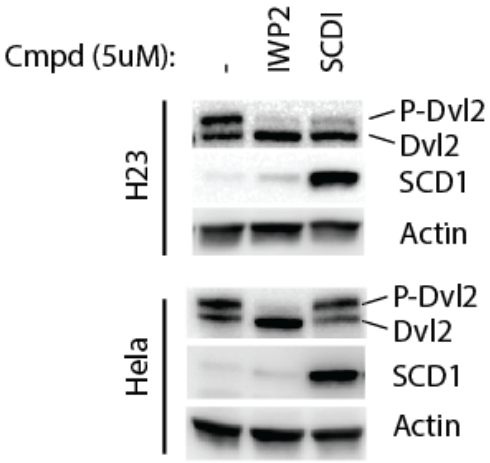


Figure 2.2: SCD inhibitors (SCDI) are inconsistent in their activity to block cell autonomous Wnt signaling activity. Lung cancer cell line H23 and cervical cancer line Hela were both treated with either IWP2 or SCDI at 5uM concentration. Whereas in H23 cells, both the compounds were able to block the Wnt activity as indicated by the disappearance of phosphorylated Dvl2 protein, SCDI was ineffective to inhibit Wnt signaling in Hela cells.

To directly probe the geometry of the acyl donor pocket in the Porcn active site, we labeled cells expressing a Wnt-immunoglobulin G Fc domain (Wnt-Fc) fusion protein with various exogenously supplied lipid probes that differ with respect to desaturation position and length, and that would enable subsequent copper-assisted cycloaddition of a biotin-conjugated azide to the alkyne group (Figure 2.3 A). Whereas previous investigations into fatty acyl donor length preferences are consistent with those presented here showing Porcn favoring medium length fatty acyl chains [113], we also observed Porcn preference for the position of desaturation that is not observed in Hhat (Figure 2.3 B). Indeed, moving the double bond to other positions greatly diminished the ability to label Wnt but not the Hh protein. Taken together with the observations demonstrating flexibility in Porcn chain length preference, this new data suggests that determinants within the Porcn site measure the distance of the double bond relative to the

coenzyme A (CoA) group (Figure 2.3 C).

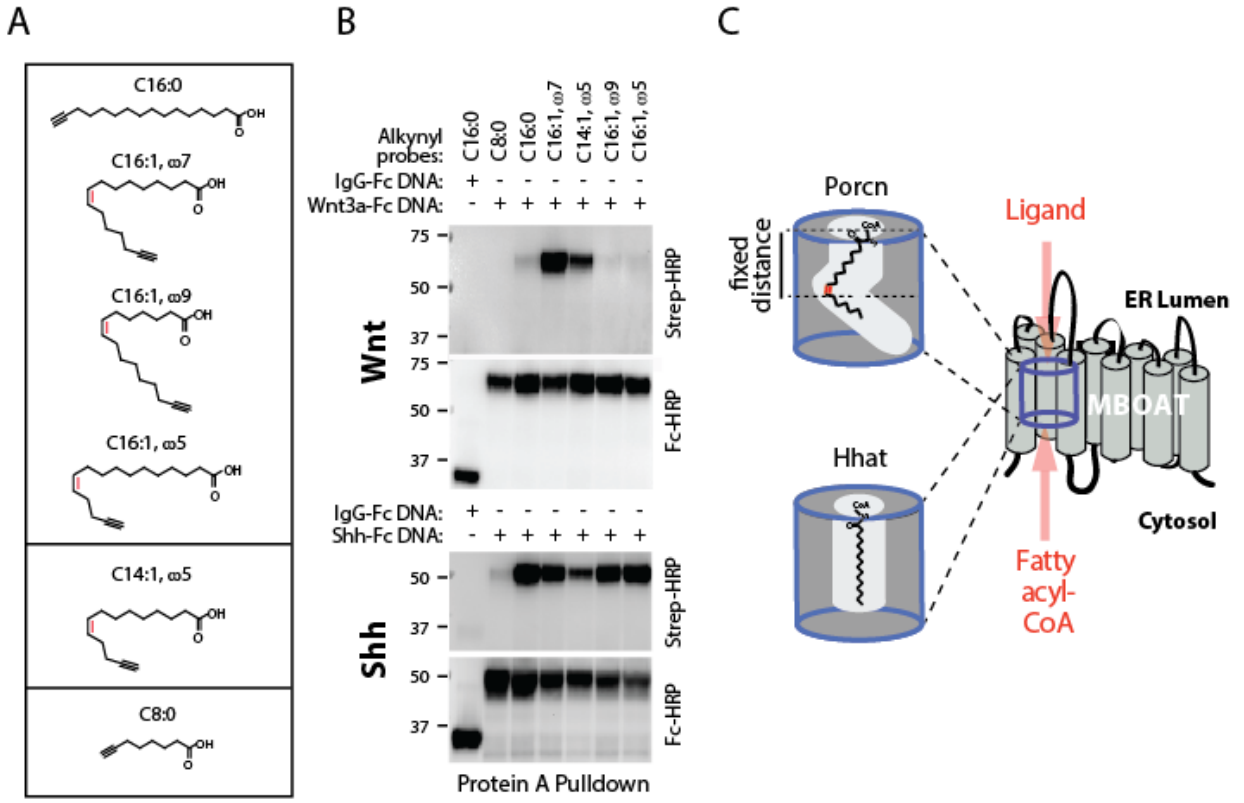


Figure 2.3: Conservation of ω-7 fatty acid labeling of Wnt proteins is attributable to features of the Porcn active site. (A) A collection of ω-alkynylated fatty acids differing in carbon chain length and desaturation position. (B) Characterization of Porcn fatty acyl donor preferences using a click chemistry approach. Cells transfected with cDNA encoding either a Wnt3a- or Shh-IgG Fc fusion protein (Wnt3a-Fc and Shh-Fc, respectively) were treated with various ω-alkynylated fatty acids (illustrated in a). An IgG-Fc protein with a signal sequence serves as a control (IgG-Fc). Wnt3a- , Shh-, or IgG- Fc protein immobilized on Protein A-Sepharose using transfected cell lysate was treated with a biotin-azide reagent to facilitate detection of lipidated protein. (C) Active site models of HHAT and PORCN. The Porcn active site differs in its ability to recognize the position of desaturation within the fatty acyl donor**.**

A universal requirement for ω-7 fatty acids in metazoan Wnt production.

Despite the recognition that Wnt proteins are found in all metazoans, unclear is whether or not this conservation extends to the nature of the covalently attached lipid moiety. An appreciation

of the universality of the lipid adduct in animals should improve mechanistic modeling of Wnt signaling reliance on lipids. For example, the contribution of species-dependent metabolic considerations to the evolution of the quality control machinery governing Wnt production remains a mystery. We assembled a collection of Porcn cDNAs from diverse metazoan phyla that in some cases reach >70% sequence divergence thus providing a means to broadly evaluate Porcn acyl donor preferences within the animal kingdom (Figure 2.4 A). Next, we built upon an assay previously described using Porcn null mouse embryonic fibroblasts expressing a murine Wnt3a protein to first evaluate the ability of various Porcn molecules to induce Wnt signaling [120] (Figure 2.4 B). This approach affords a controlled experimental platform focused on a single Wnt substrate while removing species- or cell line-specific lipid metabolic considerations from the observations. The activation of the Wnt signaling occurs only in the presence of WT Porcn and Wnt3a cDNAs, which reinforces the validity of this *in vitro* reconstitution assay (Figure 2.4 C). We then established that the Porcn proteins from different phyla evaluated were indeed capable of promoting murine Wnt3a activity as measured using either the synthetic Wnt pathway reporter (Figure 2.4 B and 2.4 D) or the Dishevelled 2 (Dvl2) phosphorylation assay described above (Figure 2.4 B and 2.4 E). These observations suggest a minimal Wnt sequence requirement for maintaining an enzyme-substrate relationship. Indeed, Porcn is able to lipidate the disulfide bonded β-hairpin that harbors the conserved acylated serine in Wnt proteins when it is expressed in the context of a structurally distinct protein scaffold [121].

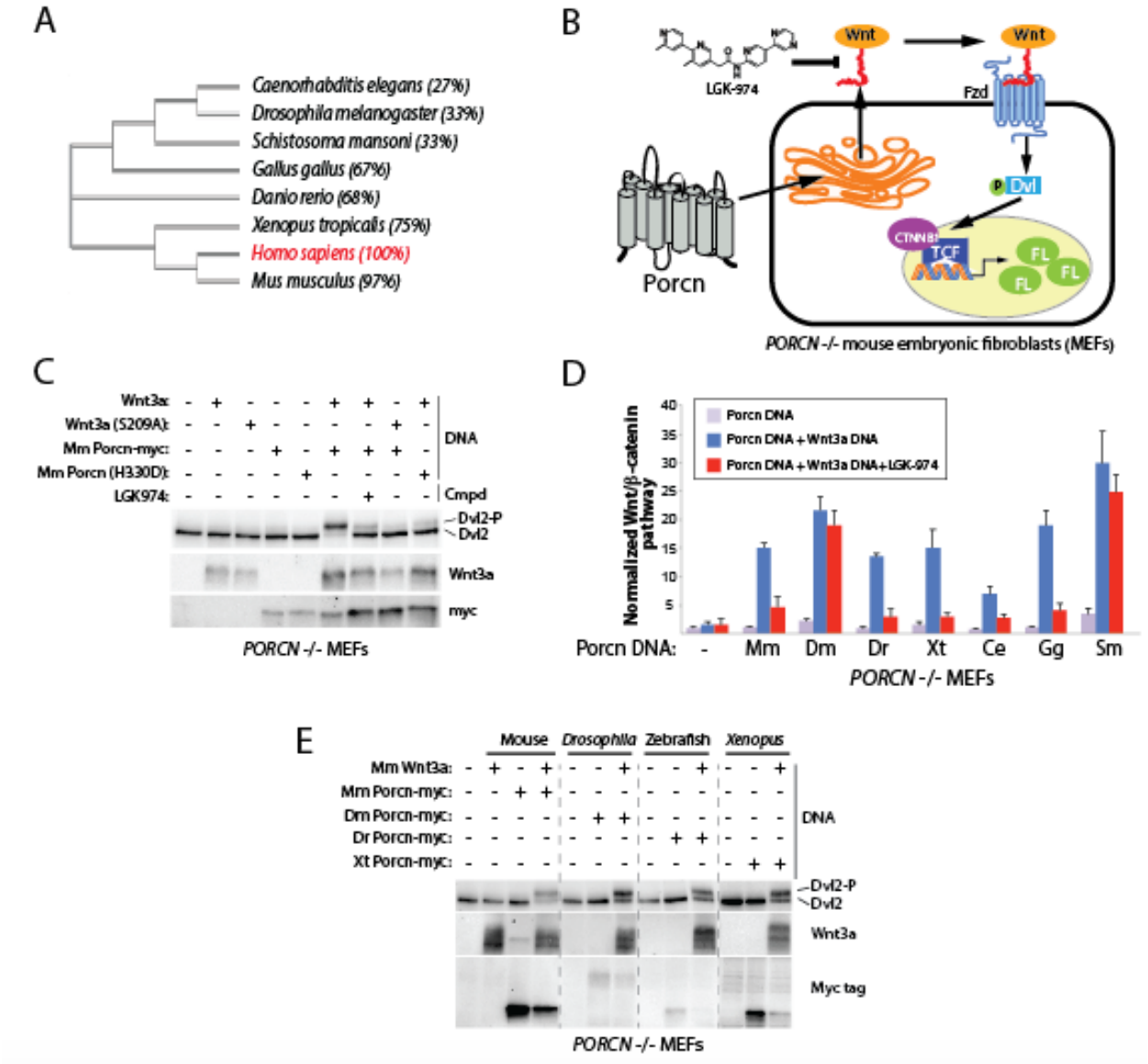


Figure 2.4: An *in vitro* platform for interrogating Porcn fatty acyl donor specificity across metazoan phyla. (A) Cladogram of Porcn protein relatedness across metazoan phyla. Chordata [mouse (*M. musculus*), zebrafish (*D. rerio*), chicken (*G. gallus*), frog (*X. laevis*)], nematoda [roundworm (*C. elegans*)], platyhelminthes [flatworm (*S. mansoni*)], and arthropoda [fruitfly (*D. melanogaster*)]. (B) Phosphorylation of Dishevelled (Dvl2) protein reports Wnt-mediated activation of the Frizzled (Fzd) receptors. Activation of canonical Wnt pathway results in the nucelar accumulation of β-catenin, and induction of TCF/LEF transcription factors-associated target gene expression. LGK974 is a Porcn inhibitor. (C) Restoration of Wnt protein production in Porcn null cells with an exogenously introduced Porcn cDNA.Wnt3a dependent Dvl2 phosphorylation in mouse embryonic fibroblasts (MEFs) devoid of Porcn expression is restored in cells transfected with cDNA encoding a wt but not a catalytically inert (H330D) Porcn protein.

Expression of Wnt3a and mPorcn activates the Wnt-mediated signaling resulting in the phosphorylation of Dvl2 proteins. Wnt3a-S209A lacks the palmiteoylation site. (D) Wnt/b-catenin pathway is activated with the expression of Wnt3a and Porcn derived from mouse, drosophila, zebrafish, xenopus C elegans, chicken, or *Schitosoma*. (E) Porcn from different species (mouse, Drosophila, zebrafish, and xenopus) are able to acylate murine Wnt3a protein in MEFs.

Having established a platform for measuring the activity of Porcn protein from diverse animals using the same cellular backdrop and a shared Wnt substrate, we then evaluated the preference of various Porcn proteins for alkynylated palmitate or palmitoleate**.** Observed for every Porcn protein tested was abundant labeling of the Wnt protein with the palmitoleate but not the palmitate analogue (Figure 2.5**)**. These observations that include Porcn from 4 out of the 9 animal phyla and representative animals that diverged ~1.2 billion years ago (between Nematoda and Chordata) [122] suggest that Porcn universally enforces a monounsaturated fatty acyl donor preference in metazoans.

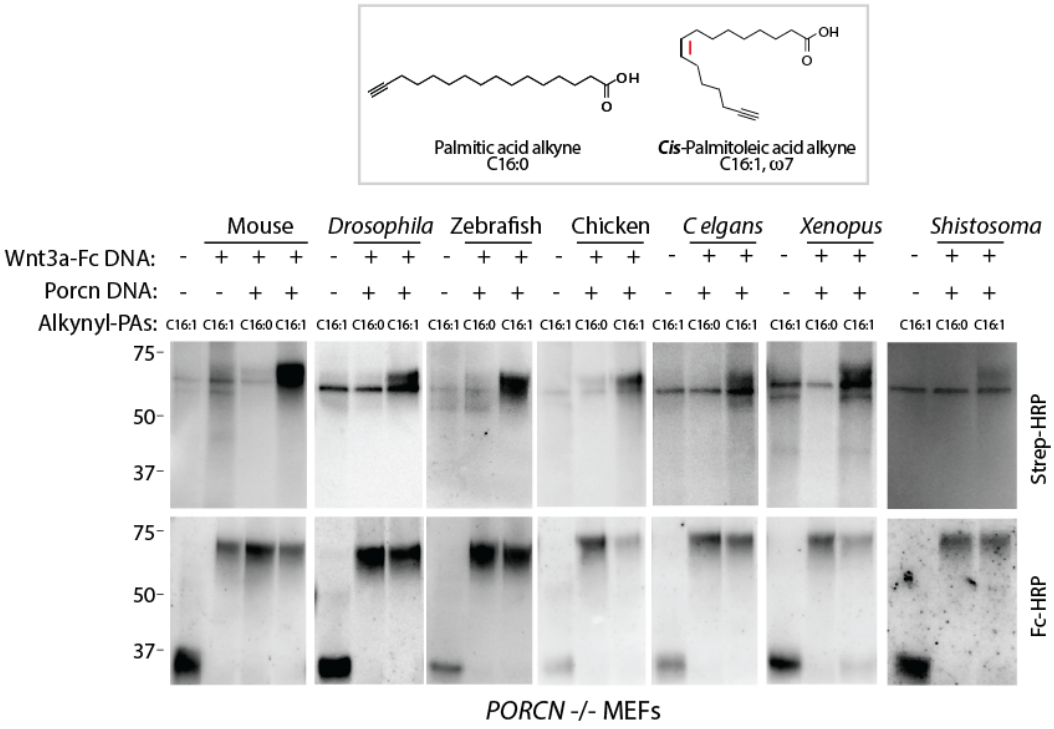


Figure 2.5: A conserved function for cis-palmitoleate in metazoan Wnt signaling. Porcn null MEFs co-expressing Wnt3a- or IgG-Fc and either mouse, *Drosophila*, zebrafish, chicken, *C. elegans*, or *Xenopus* Porcn were treated with alkynyl-palmitic acid (C16:0) or alkynyl-palmitoleic acid (C16:1). Purified Wnt3a-Fc protein was then subjected to the cycloaddition reaction as described above. Porcn across diverse phyla utilizes palmitoleate for acylating Wnt proteins.

Bypassing Porcn fatty acyl donor stereoselectivity using an exogenously provided *trans* fat

A challenge to interrogating the role of the Wnt lipid adduct in Wnt manufacturing and signaling has been the inability to experimentally follow Wnt proteins devoid of a *cis* fatty acid. We attempted to force deviant Wnt acylation with a *trans* ω-7 fatty acyl analogue as a means to overcome this limitation. Whereas Hhat did not distinguish between the C16:1 fatty acid isomers in agreement with its absence of a preference for a specific lipid, Porcn preferred the *cis* fatty acid but nevertheless was able to make limited use of the *trans* molecule (Figure 2.6 A). A SCD inhibitor blunted Wnt acylation with a saturated fatty acyl probe (C16:0) consistent with

previous observations [95] but not with the *cis* or *trans* fat probes suggesting that the *trans* fat labeling was not enabled by cellular isomerization (Figure 2.6 A). Using a pulse-labeling strategy with either *cis/trans* palmitoleic acid analogues (Figure 2.6 B), we observed a crippling effect of *trans* w-7 fatty acylation on Wnt protein cellular release (Figure 2.6 C and 2.6 D). We also noted that LGK974 inhibited *trans* fatty acylation of Wnt protein thus confirming the role of Porcn in this biochemical event. Although previous results have established the importance of lipidation in moving the Wnt protein to the extracellular milieu, these observations are the first that we are aware of to demonstrate the essentiality of the *cis* double bond to Wnt protein manufacturing.

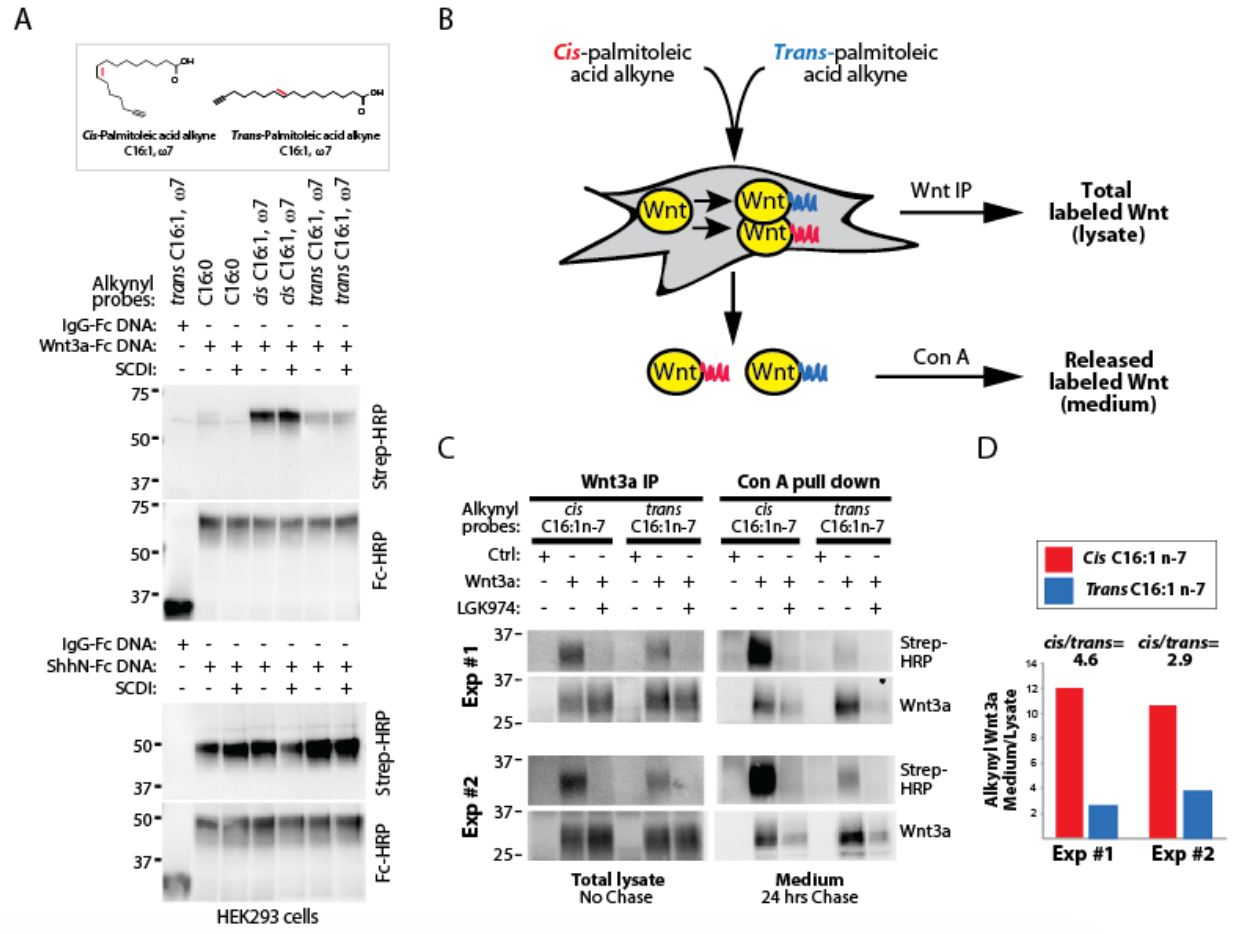


Figure 2.6: Trans fatty-acylated Wnt proteins are secretion incompetent. (A) Porcn exhibits stereoselectivity for its fatty acyl donor. *Cis* and *trans* ω-alkynylated C:16:1 fatty acids were used to label Wnt or Shh (control) as before. Stearoyl-CoA desaturase inihibitor (SCDI) was used here to demonstrate that cellular isomerization of alkynylated probes by SCD does not contribute to protein labeling. (B) A click chemistry based strategy for investigating the influence of palmitoleic acid isomerization on Wnt cellular release. (C) Culture cell medium was collected 24 hrs following pulse labeling of cells with either *cis* or *trans* ω-alkynylated palmitoleic acid. Wnt proteins from the culture medium were enriched using concanavalin A (con A) sepharose beads and subjected to a cycloaddition reaction that appends a biotin group to labeled protein. A baseline labeling efficiency associated with each palmitoleate isomer was determined using a similar analysis of Wnt protein isolated from total lysate. (D) The values for total and released click chemistry-labeled Wnt protein were used to calculate a normalized Wnt secretion value for each palmitoleate isomer. In two separate experiments, Wnt protein release is compromised when it is adducted to a *trans* palmitoleate.

WLS is a lipoprotein receptor that exhibits stereoselectivity for *cis* fatty acylated Wnt molecules

The polytopic WLS protein is essential to Wnt trafficking given its genetically based deletion results in loss of Wnt cellular release [16, 18, 103]. Furthermore, loss of Porcn activity or mutation of the acylated serine in Wnt proteins results in loss of Wnt-WLS interaction suggesting that WLS is a lipoprotein receptor [19, 20]. WLS expression is controlled by Wnt/beta-catenin signaling thus supporting a transcription-based feed forward signaling mechanism [123]. Here, we show that forced expression of Wnt proteins in HEK293 cells including those that do not induce transcriptional responses in these cells is matched by changes in the abundance of WLS protein (Figure 2.7 A). These observations reveal a post-translational regulatory mechanism that provides additional evidence for the coordination of Porcn and WLS cellular activities [16, 19]. Consistent with previous observations, the addition of either IWP-2 or a SCD inhibitor abrogates the interaction between Wnt and WLS confirming an essential role of the lipid adduct in mediating this interaction, and that Wnt molecules are solely modified by ω-7 fatty acids (Figure 2.7 B). At the same time, since both compounds effectively eliminate Wnt

lipidation, unclear is whether or not the *cis* double bond feature found in palmitoleate contributes to Wnt-WLS interaction.

Whereas the heptahelical domain of WLS may contribute to Wnt protein recognition, the N-terminal extracellular domain likely harbors the elements required for stereoselectivity given its resemblance to the lipid-binding lipocalin proteins [20]. Indeed, the secreted Swim protein which promotes long-range *Drosophila* Wnt signaling in a Wnt lipidation-dependent fashion belongs to the lipocalin gene family [124]. We tested the hypothesis that WLS is capable of directly interacting with lipids by utilizing an *in vitro* binding assay that incorporates immobilized *cis* and *trans* palmitoleate and a luciferase tagged WLS protein (Figure 2.7 C). In addition to observing a specific interaction of WLS protein with *cis* palmitoleate, we also noted a preference in WLS for the *cis* molecule as opposed to the *trans* molecule (Figure 2.7 D). Despite our inability to evaluate binding here between WLS and recombinant Wnt engineered with *cis* versus *trans* fats, our data suggests the presence of determinants in WLS that support direct and stereoselective interaction with a monounsaturated fatty acid.

WLS has previously been shown to ferry Wnt proteins from the Golgi to multivesicular bodies where the cargo is then released with cellular membrane particles (exosomes) into the extracellular milieu [25]. However, the extent to which exosomes and the canonical secretory pathway contribute to the secreted Wnt ligand pool remains unclear [102]. We evaluated the relative contribution of exosome-associated Wnt proteins to the overall Wnt protein found in culture medium in our cell culture system (Figure 2.7 E). At least in the case of Wnt5a, we find the majority of the protein found in the culture medium can be isolated using an exosomal purification system. Using the case of Wnt5a to model the production of Wnt proteins in general,

we find that stereoselectivity of WLS for *cis*-lipidated Wnt proteins constitutes a critical quality control measure for the production of functional Wnt proteins (Figure 2.7 F). Likely, this mechanism in addition to the presence of determinants that enforce stereoselectivity of Porcn for fatty acyl donors as shown here supports a buffering mechanism that limits the influence of dietary *trans* fat consumption on the integrity of a signaling pathway important in stem cell homeostasis.

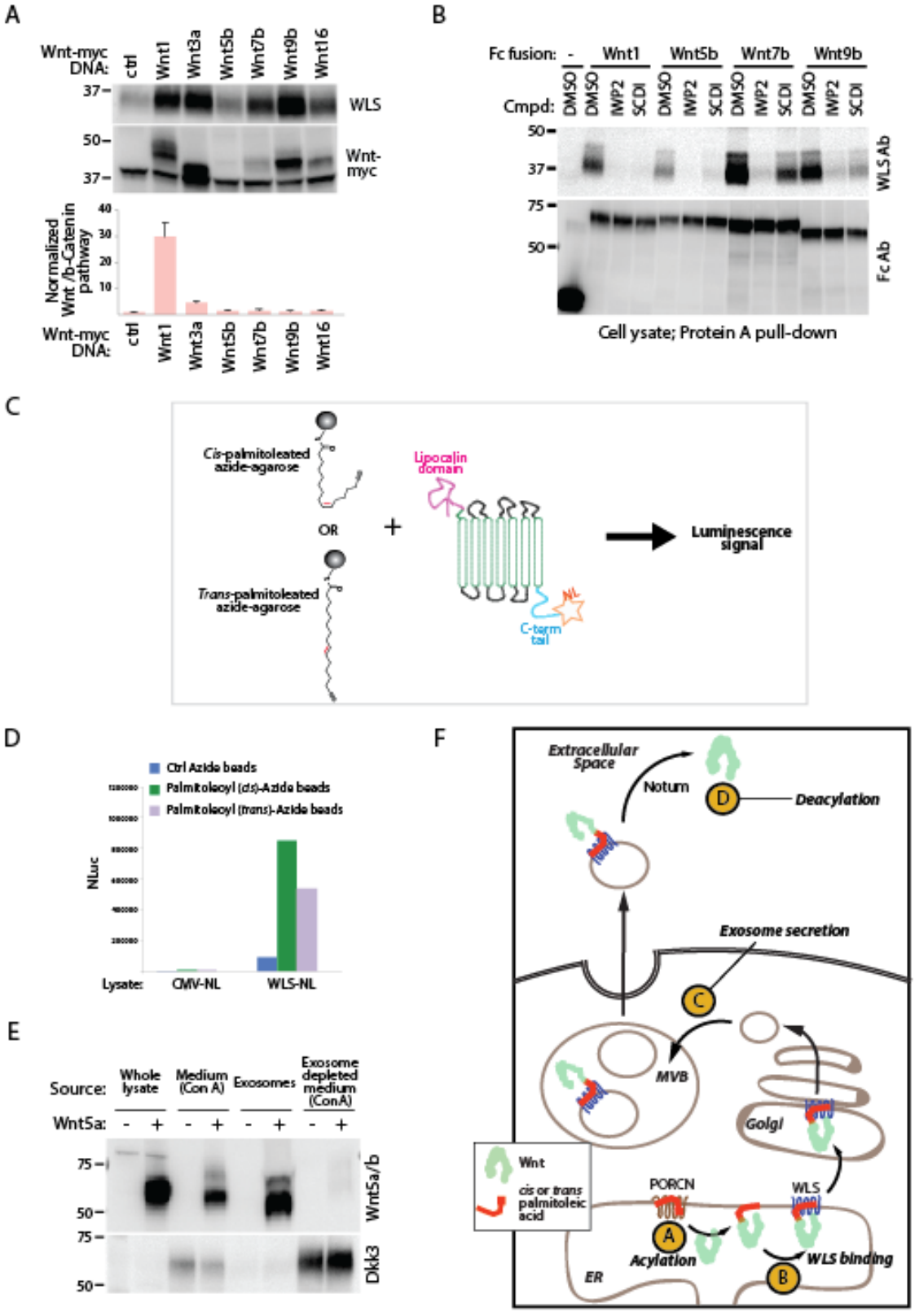

A
Wnt-myc DNA:
ctrl
Wnt1
Wnt3a
Wnt5b
Wnt7b
Wnt9b
Wnt16
37
WLS
50
37
Wnt-myc
Normalized Wnt /b-Catenin pathway
40
30
20
10
B
Fc fusion: -
Wnt1
Wnt5b
Wnt7b
Wnt9b
Cmpd: DMSO
IWP2
SCDI
50
37
75
50
WLS Ab
Fc Ab
Cell ysate; Protein A pull-down
C
Cis-palmitoleated azide-agarose
OR
Trans-palmitoleated azide-agarose
Lipocalin domain
NL
C-term tail
Luminescence signal
D
Ctrl Azide beads
Palmitoleoyl (cis)-Azide beads
Palmitoleoyl (trans)-Azide beads
NLuc
Lysate:
CMV-NL
WLS-NL
E
Source:
Whole lysate
Medium (Con A)
Exosomes
Exosome depleted medium (ConA)
Wnt5a:
75
50
Wnt5a/b
Dkk3
F
Extracellular Space
Notum
D
Deacylation
Exosome secretion
C
MVB
Golgi
Wnt
cis or trans palmitoleic acid
PORCN
WLS
A
Acylation
WLS binding
B
ER

Fig. 2.7. The Wnt chaperone WLS is a stereoselective lipoprotein receptor. (A) Wls protein abundance is gated by Wnt ligands in a b-catenin/TCF-independent fashion. (B) WLS interaction with diverse Wnt proteins is sensitive to chemical disruption of Wnt lipidation. Indicated Wnt proteins fused to the IgG Fc domain were expressed in HEK293 cells and purified using protein A sepharose. Isolated proteins were Western blot analyzed for the presence of Wls and Wnt-Fc proteins. SCDI inhibits stearoyl Co-A desaturase which desaturates palmitic acid at the ω-7 position. (C) An *in vitro* assay for determining WLS preference for lipid stereoselectivity. Lysate from cells expressing WLS fused to Nanoluc (NL) luciferase was incubated with cis- or trans-palmitoleate immobilized on agarose beads. Luciferase signal associated with beads following several washes was then determined. (D) Wnt proteins are predominantly secreted in exosomes. ConA sepharose precipitated protein and purified exosome from medium cultured with HEK293 cells overexpressing Wnt5a was subjected to Wnt5a Western blot analysis in order to compare total secreted and exosome-associated Wnt5a protein. DKK3 serves as a general secreted protein marker. (E) Model of the Wnt protein production apparatus indicating points of lipid stereoselectivity re-enforcement. The Wnt deacylase Notum has previously been shown to favor cis fatty acids as substrates based on structural insights and *in vitro* binding data [125].

Conclusions and Future Directions

The participation of fatty acids in signal transduction poses several challenges for protein engineering including the dearth of salient features found in aliphatic chains that could be exploited for selective protein-lipid recognition. The kink introduced by the *cis* configuration provides a handle that could be leveraged by proteins to distinguish *cis* fatty acids from those that are saturated or that harbor a *trans* double bond [110]. For example, the Wnt deacylase Notum utilizes an arrangement of several residues within its active site to generate a steric gate that precludes the binding of fatty acids without a *cis* double bond in the ω-7 position [126]. In the case of WLS, its stereoselectivity as described here may serve either to buttress the discriminatory prowess of WLS for Wnt ligands such as by increasing affinity for *cis* palmitoleate, or to synchronize Wnt-mediated intercellular communication with the availability of monounsaturated fatty acyl-CoA.

The influence of fatty acyl protein modifications on Wnt signaling is pervasive and extends beyond Wnt lipidation. For example, the ability of Lrp6 to exit the ER is gated by palmitoylation of a juxtamembrane residue [127], and the activity of Porcn itself is controlled by palmitoylation by an unknown intracellular enzyme [113]. This extensive interconnectivity of lipid metabolism and Wnt signaling may have its roots from the coopting of a lipid sensing apparatus in unicellular organisms for the purpose of intercellular communication [105]. At least in humans, palmitoleate is the second most abundant MUFA in adipose tissue and blood [128] and perhaps serves in the Porcn-WLS system as currency for synchronizing metabolic health and tissue renewal. Indeed, loss of SCD in induced pluripotent stem cells elicits a transcriptional response that includes the induction of many Wnt genes suggesting a hardwiring of MUFA abundance with Wnt signaling [108].

Our ability to force Wnt acylation with exogenously provided *trans* palmitoleate suggests that an adjacent saturated bond likely can adopt a *cis*-like configuration albeit with poor efficiency in the Porcn active site. Though not tested here, conceivable a MUFA probe with a *cis* double bond at either the ω-6 or ω-8 positions could be accepted by Porcn. Nevertheless, the ability of Porcn to supply a *trans* fatty acid onto Wnt suggests that in the presence of excess saturated fatty acids, Porcn may also immobilize alternative fatty acyl donors onto Wnt proteins. Based on our observations with a *trans* fatty acylated Wnt molecule, we assume Wnt proteins modified with a saturated fatty acid would not be readily recognized by WLS and thus destined for entrapment in the secretory pathway. Given that exogenously supplied lipids can influence octanoylation of the appetite controlling hormone ghrelin by its acyltransferase GOAT [129], we suspect that dietary contributions of saturated and *trans* fats might influence the efficiency of

Wnt production in adult cells. For example, levels of circulating *trans* fatty acids in the form of triacylglycerols in breast-fed infants are directly correlated with its abundance in breast milk which can reach up to 8% of the total fat content primarily as a consequence of the mother's consumption of prepared foods [130].

Despite their chemical diversity, the majority of Porcn inhibitors likely target the Porcn active site based on their ability to successfully compete with a fluorescently labeled active site probe for Porcn binding [131]. For antagonists such as GNF-1331 (the precursor of LGK974) and IWP-2 that have also been evaluated for their activity against other MBOATs, there is evidence that these molecules exhibit specificity for Porcn [131, 132]. Our observations described here provide insights into how this is achieved - namely determinants that support Porcn selectivity for ω-7 *cis* fatty acyl donors can be exploited to achieve chemical control of Porcn. Additional support for this conclusion may be found in enantiomer-dependent activity of at least one class of Porcn inhibitors [133]. The Porcn-WLS relationship likely extends throughout the animal kingdom based on our evidence for the conservation of determinants that enforce ω-7 lipidation of Wnt proteins, and that mammalian Wnt proteins expressed in *Drosophila* are not released in the absence of Wls [19]. Thus, Porcn inhibitors disrupt an ancient relationship between Porcn and WLS that is centered on their shared ability to recognize a single fatty acid feature displayed on a master regulator of cell fate outcome in diverse animal phyla. The inability of LGK974 to disable a subset of Porcn from the animal sources (Figure 2.4 D) tested further suggests that some non-conserved features found in the Porcn active site could exploited for species-specific disruption of animal development for parasitology goals in a variety of host animals.

CHAPTER THREE

Mutation in tumor suppressor gene *LKB1* increases total protein secretion, including Wnt proteins

Introduction

LKB1/STK11 encodes a serine/threonine kinase that was initially identified as a tumor suppressor gene mutated in Peutz-Jeghers Syndrome (PJS) [134]. PSJ patients with germline inactivating mutations are predisposed to gastrointestinal polyposis and cancer [135, 136]. Subsequent studies have found that somatic mutations in *LKB1* results in several types of cancers including in ~40% of lung cancers [137-139] and ~30% of cervical carcinomas [140, 141]. LKB1 is a master kinase regulating the activity of multiple kinases including two isoforms of AMPK [142]. LKB1 and AMPK negatively regulates mTOR-signaling pathway [143, 144], which controls cellular metabolism and growth [143, 144]. In addition to the AMPK-dependent pathways, LKB1 also controls diverse biological pathways relevant to cancer. From genome-scale loss-of-functions screens in cultured cells to identify novel modulators of cell fate determination pathways, our lab had previously uncovered a role for LKB1 in suppressing the Wnt signaling pathway [145]. LKB1 also mediates the transcriptional repression of the Wnt ligands Wnt5a and Wnt5b through the control of epithelial-to-mesenchymal transition transcription factor Snail [146]. In addition to these previously established roles of LKB1 in regulating the Wnt ligands in a transcriptional and non-transcriptional dependent manner, I have also found LKB1 to be important in fueling rampant production of secreted proteins, including Wnts. In summary, my studies reveal the role of metabolic sensor protein LKB1 as a modulator

of Wnt protein secretion with elevated fatty acid production thus highlighting a previously unappreciated mechanism that couples metabolic derangements to intrinsic and communal cell fate outcomes.

Cancerous cells lacking *LKB1* exhibit deviant Wnt-mediated signaling

RNAi screen using two genome-scale libraries was used in order to identify genetic changes that engage the Hedgehog and Wnt signal transduction pathways [145]. From that screen, LKB1 emerged as a critical mediator in both the Hh and the Wnt pathways (Figure 3.1 A). Loss of *LKB1* results in the upregulation of the Wnt signaling pathway as indicated by increased Dvl2 phosphorylation in mouse embryonic fibroblasts (MEFs) (Figure 3.1 B). Since sporadic cancers, especially from lung and cervix, frequently harbor mutations in *LKB1*, Wnt signaling in panel of cells from those tumor types were evaluated. Lung cancer cell lines lacking *LKB1*, such as H460, A549 and H1819, had enhanced Dvl2 phosphorylation compared to the normal lung cancer cell line HBECs, which is consistent to the results observed in MEFs (Figure 3.1 C). The increased Dvl2 phosphorylation was reversed with a Porcn inhibitor IWP2, but expectedly not with a Tnks inhibitor IWR1. Finally, similar results were garnered from the study of cervical carcinoma cell lines harboring *LKB1* mutations (Figure 3.1 D).

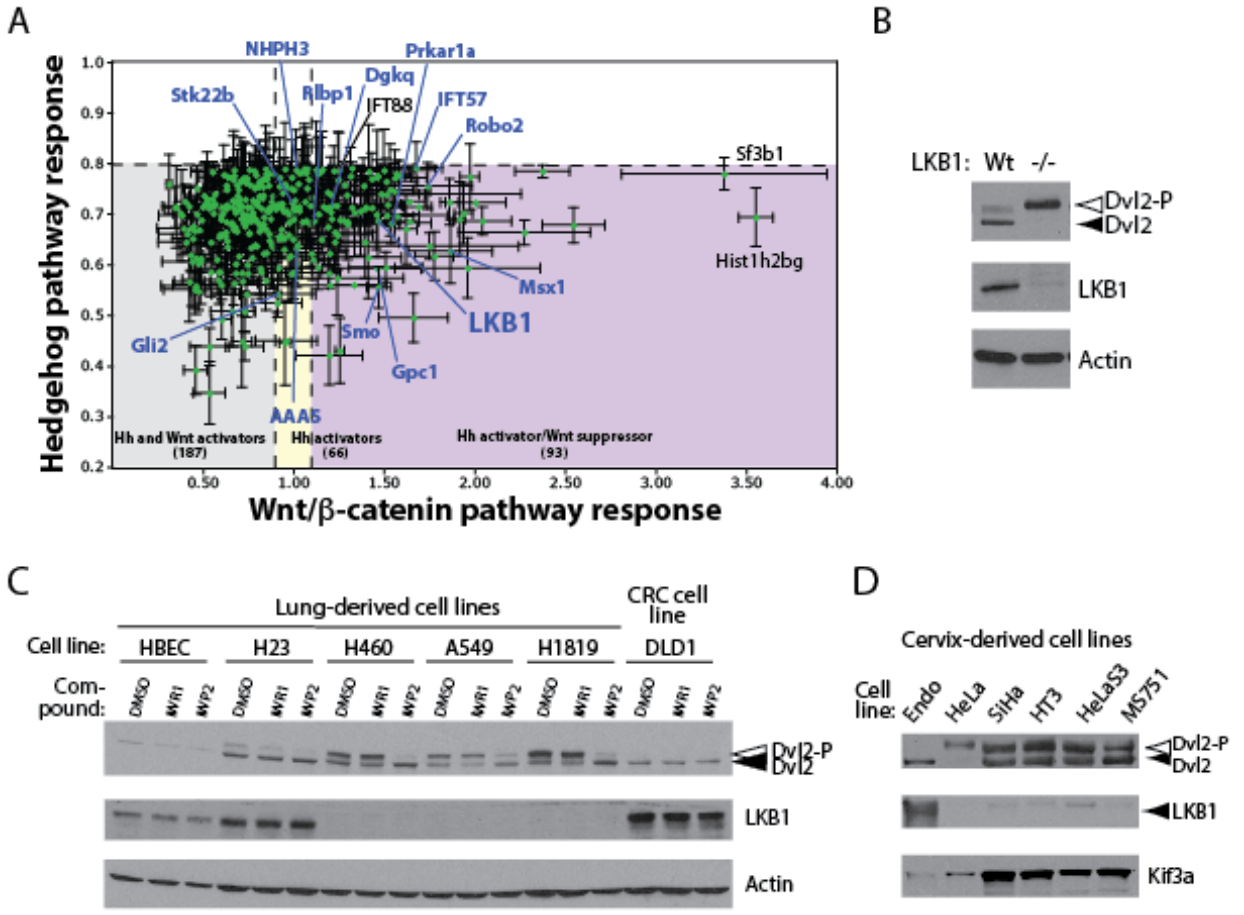


Figure 3.1: Identification of LKB1 as a tumor suppressor in lung and cervical carcinoma. (A) Graphical summary of results from genome-wide siRNA screen in cultured cells for Hh and Wnt pathway components with normalized Hh (Y axis) and Wnt (X axis) reporter signals. Genes, including *LKB1*, in purple quadrant are suppressors of Wnt signaling. (B) *LKB1* null MEFs exhibited elevated Dvl2 phosphorylation. (C) Lung-derived cell lines devoid of LKB1 expression exhibited elevated levels of IWP2-sensitive Dvl2 phosphorylation. IWR1 targets TNKS and serves as a negative control here. (D) Cervical carcinoma cell lines with low or no LKB1 expression exhibit elevated Wnt signaling as indicated by increased Dvl2 phosphorylation levels. IWP2 (Porcn inhibitor) but not IWR2 (Tnks inhibitor) blocked Dvl2 phosphorylation levels relative to the LKB1 expressing Endo cells. Figures 3.1 A, B and D derived from Jacob, 2011 [145].

Induction of non-canonical Wnt pathway components account for the increased Wnt signaling in *LKB1* null MEFs

The molecular mechanism leading to the increased Dvl2 phosphorylation in MEFs lacking *LKB1* is not well understood. In order to systematically identify aberrantly expressed transcripts in MEFs following *LKB1* loss, we performed microarray analysis of WT and *LKB1* null MEFs using the Illumina MouseRef-8 v2 Beadchip array. Transcriptional profiling results revealed an enrichment in the expression of various Wnt pathway components in *LKB1* null cells. The Wnt ligands Wnt5a and Wnt 5b were increased by more than two folds in cells lacking *LKB1* (Figure 3.2 A). This is consistent with findings by Reuben Shaw's group where they demonstrated the loss of *LKB1* induces the expression of the Snail transcription factor, which in turn promotes Wnt5a expression [146]. In addition to the transcriptional induction of these Wnt ligands, the Wnt receptor Ror2, which regulates planar cell polarity (PCP) and signaling independently of influencing beta-catenin activity [147-149], was also upregulated in *LKB1* null MEFs. The binding of Ror2 to Wnt5a via its CRD [149] induces Dvl2 phosphorylation [150]. Furthermore, high-affinity function blocking antibodies targeting the ectodomains (ECDs) of Ror1 and Ror2 blocks Dvl2 phosphorylation in a dose-dependent manner [150]. Thus, the enhanced Dvl2 phosphorylation in cell lines devoid of *LKB1* is likely the consequence of enhanced Wnt5a-Ror2 signaling. The induction of Wnt5a transcript levels was confirmed by quantitative real time PCR of Wnt ligands. Out of 19 Wnt ligands tested, Wnt5a and Wnt5b expression showed the greatest sensitivity to LKB1 status (Figure 3.2 B).

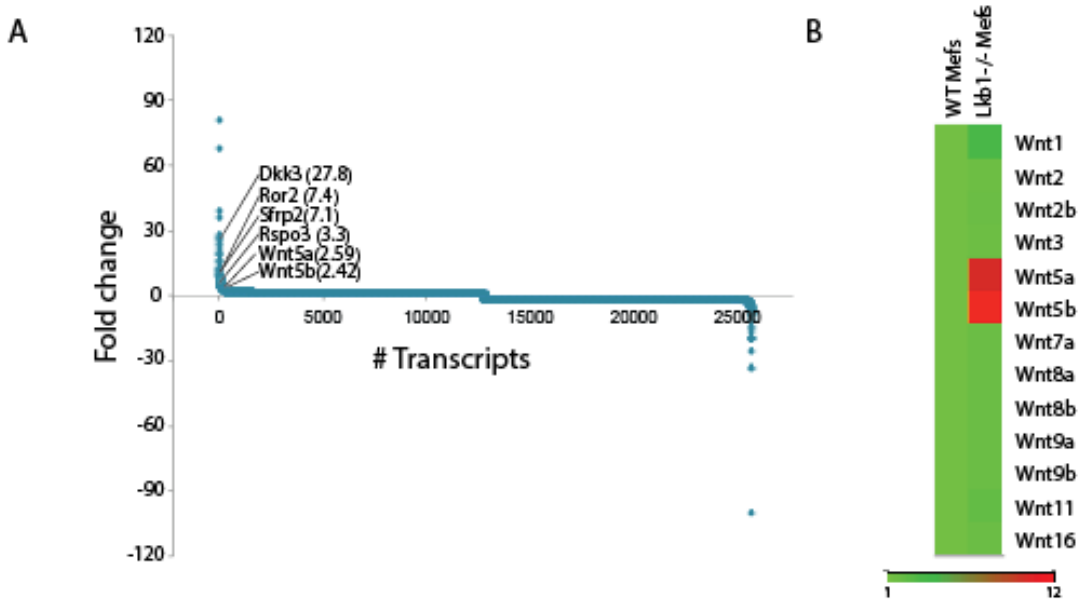


Figure 3.2: Elevated expression of Wnt5a-Ror2 signaling components in MEFs devoid of *LKB1*. (A) Transcriptional profiling reveals induction of Wnt pathway components due to the loss of *LKB1* in MEFs. (B) RT-PCR results of Wnt gene expression in wt and *LKB1* null MEFs. Only results that exceed a baseline signal for expression are shown.

In accordance to the enrichment of Wnt5a/b transcripts in *LKB1*-/- MEFs, loss of LKB1 also increased the Wnt5a/b protein levels (Figure 3.3 A). LKB1-mediated suppression of Wnt5a/b was independent of AMPK activation since the accumulation of Wnt5a/b proteins was not observed in *AMPK*-/- MEFs. *LKB1* null MEFs also exhibited elevated Ror2 expression. Moreover, the Porcn inhibitor LGK974 greatly reduced the abundance of Wnt5a/b and blocked the phosphorylation of Ror2 receptor. Given the ability of a Porcn inhibitor to reverse deviant Wnt signaling in these cultured cells devoid of LKB2 expression, LGK974 could potentially be used as an anti-cancer agent against LKB1-associated tumorigenesis (Figure 3.3 B).

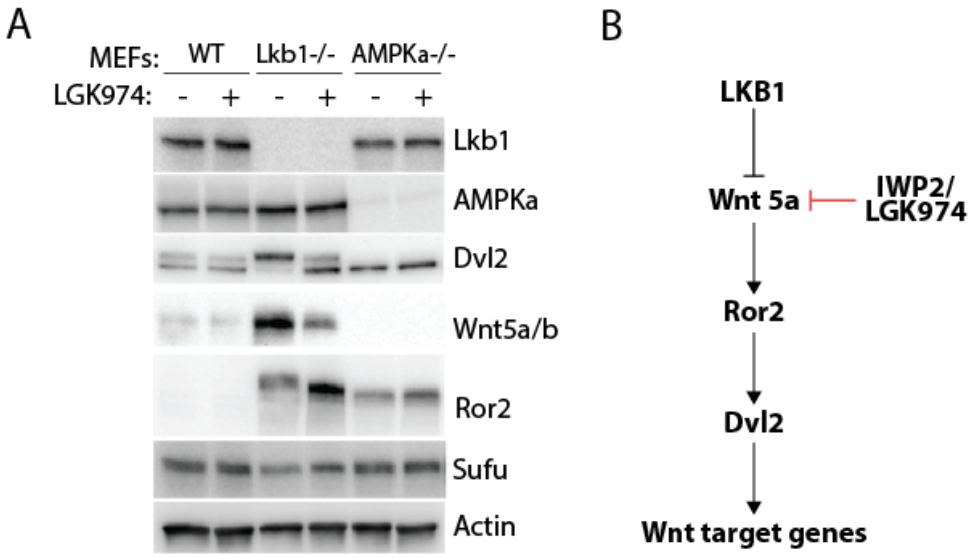


Figure 3.3: AMPK-independent induction of noncanonical Wnt signaling pathway components in *LKB1*-/- cells. (A) *LKB1* null MEFs have increased Wnt5a and Ror2 protein expression. LGK974 reverses Dvl2 and ROR2 phosphorylation induced by loss of *LKB1*. (B) A model delineating the mechanism of increased Wnt signaling pathway in *LKB1*-null cells.

Increased secreted protein biosynthesis as a consequence of *LKB1* loss

Given the role of LKB1 in regulating metabolic processes including lipid production [151], the increase in Wnt-dependent signaling in *LKB1* null MEFs may be a consequence of increased amounts of palmitoleoyl-CoA species, which is necessary to acylate Wnt proteins. Indeed, total cellular fatty acid content, including palmitate and palmitoleate, are elevated in *LKB1* null cells (Figure 3.4 A). To quantify the rate of Wnt production in WT and *LKB1* null MEFs, a luminescence-based assay for measuring Wnt secretion was used (Figure 3.4 B). The *Gaussia* luciferase (GL) signal from the culture medium reports the total levels of secreted Wnt proteins whereas *Cypridina* luciferase (CL) is a control reporter. *LKB1*-null MEFs have elevated Wnt secretion (Figure 3.4 C). The secretion of Wnt proteins is acylation dependent as the presence of the Porcn inhibitor IWP2 largely eliminates the reporter signal from the culture

medium. Thus, LKB1 suppresses Wnt production independently of the transcription control of Wnt genes. This enhancement of Wnt secretion in *LKB1* null MEFs is independent of AMPK activity as *AMPK*-/- MEFs fail to recapitulate the increased secretion of Wnt proteins (Figure 3.4 D). To test the strength of this observation beyond the context of fibroblasts, I leveraged CRISPR/Cas9 gene editing technology to engineer isogenic human pancreatic cancer cell line MIA that differ only with respect to *LKB1* status. While using CRISPR-Cas9 system to achieve *LKB1* knockout, I discovered unanticipated mechanisms that compromise the CRISPR-associated genome editing efficiency, which I will be discussing in detail in the next chapter (Chapter 4). The successfully CRISPR-Cas9-edited *LKB1* mutant Mia clones had significantly higher rate of Wnt secretion compared to the WT Mia clones, which corroborates our initial finding in MEFs (Figure 3.4 E). Therefore, indifferent to the cellular background, loss of *LKB1* results in elevated Wnt ligand synthesis. Furthermore, I wanted to analyze whether the increased secretion of the Wnt proteins was a specific consequence of LKB1 loss or LKB1 has broader implication in suppressing total protein secretion. WT and *LKB1*-/- MEFs were pulse-labeled with [^{35}S]-Methionine for 30 minutes and the radiolabelled proteins secreted in the culture medium were tracked over the course of 24 hours. Overall, loss of *LKB1* resulted in an increased total protein secretion (Figure 3.4 F).

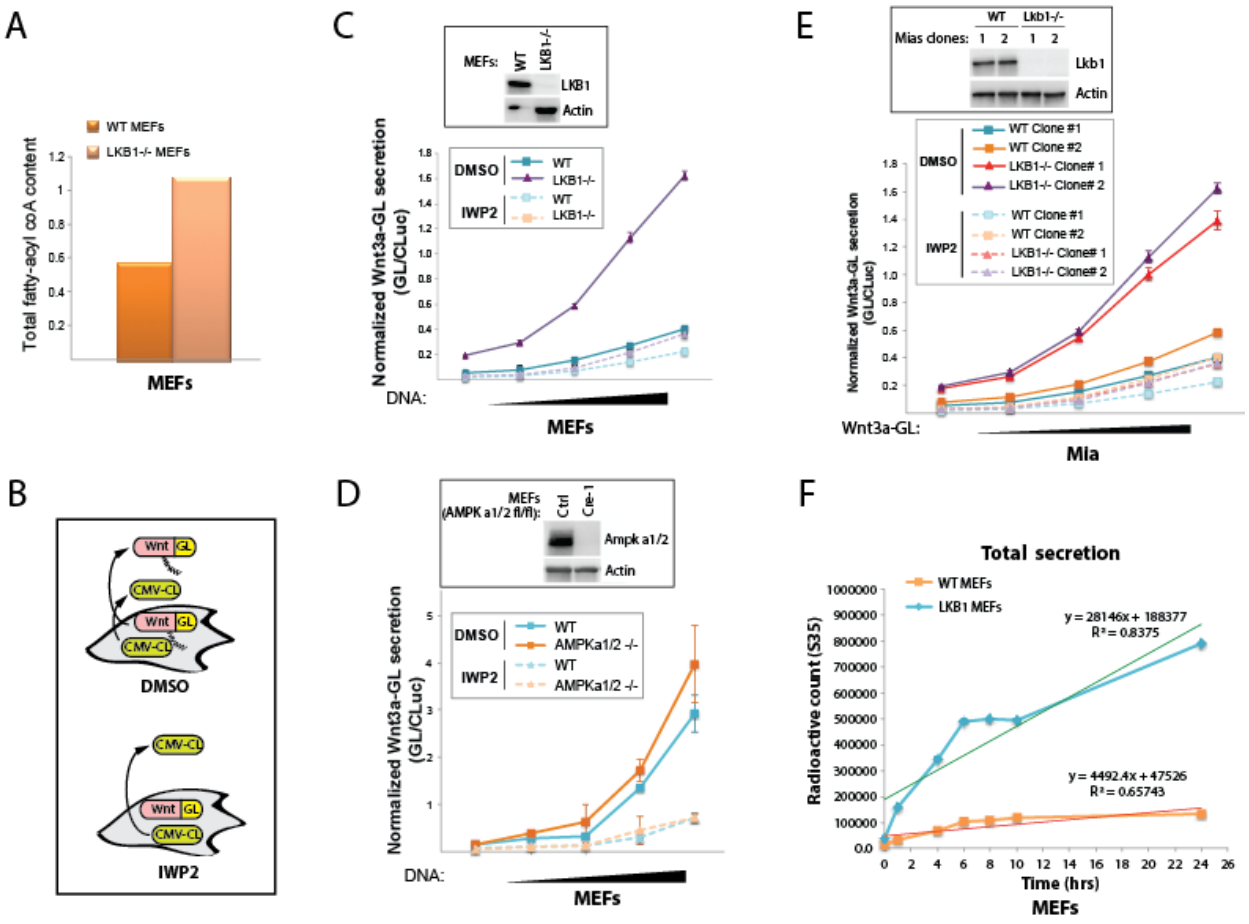


Figure 3.4: LKB1 suppresses global protein secretion (A) *LKB1* null cells have elevated total fatty-acyl CoA content. LC-ESI MS/MS (Liquid chromatography-Electrospray ionization mass spectrometry) was used to quantify fatty acyl CoA cellular content (Dr. Mike Bennett's group, Children's Hospital of Philadelphia). (B) A luminescence-based assay for monitoring Wnt protein secretion. Wnt protein level in the medium was determined using a Wnt-*Gaussia* luciferase fused plasmid (Wnt-GL). (C) *LKB1* null cells have increased secretion of lipidated Wnt proteins. Reporter activities in Wt and *LKB1*-null MEFs transfected with increasing amounts of reporters described in "B" treated with and without IWP2. (D) *AMPK*α 1/2-null cells, unlike *LKB1*-null cells, do not have elevated Wnt protein secretion. (E) CRISPR-Cas9 gene editing technology was used to generate two clones of *LKB1* mutants (which no longer express LKB1 protein as shown in the inset) in Mia pancreatic cancer cell line. Consistent to the findings in the MEFs, loss of *LKB1* in Mia cells also resulted in elevated Wnt protein biosynthesis. F. Pulse-chase experiment using Radiolabelled [^{35}S]-methionine revealed increase in global protein secretion in the absence of *LKB1*.

Conclusion and future direction

Genome-wide RNAi screen performed in the Lum lab first revealed the role of LKB1 in regulating the Hedgehog and the Wnt signal transduction pathways [145]. LKB1, a bona fide ciliary resident protein [152], inhibits the disassembly of the primary cilia, which in turns heightens the Hh signaling. Besides its role in ciliary maintenance, LKB1 mediates the level of Hh response through the control of the transcriptional repressor Gli3R. The same RNAi screen also identified LKB1 as a negative regulator of the Wnt signaling pathway. Subsequent studies by our group and others [146] have shown that LKB1 controls the transcription of non-canonical Wnt components including Wnt5a/5b and Ror receptors in fibroblasts. Since *LKB1* is mutated in the Peutz-Jehgers syndrome (PJS), further confirmation of the role of the Wnt signaling in intestinal polyposis found in PJS patients is yet to be conducted. The establishment of the Wnt signaling in PJS would potentially increase the utility of Wnt inhibitors as promising therapeutics in this disease context.

While the LKB1-mediated transcriptional modulation of Wnt signaling in fibroblasts is validated, unknown is whether this signaling axis persists in the epithelial cells. In order to delineate the LKB1-Wnt signaling cascade in epithelial cells, I used CRISPR-Cas9 technology to generate *LKB1* knockout in Mia pancreatic cancer cell line. The unexpected observations from my gene deletion efforts revealed novel compensatory mechanisms employed by cells to counter the detrimental consequences of INDEL mutations, which is covered in detail in the next chapter. Successfully CRISPR-edited Mia *LKB1* null clones had elevated secretion of Wnt ligands, which was also the case in *LKB1* null fibroblasts. Further characterization of the Wnt signaling pathway in the *LKB1* -/- Mia cells and other epithelial cell lines with *LKB1* deletion remains to be

performed. In addition to the biochemical analysis, I will evaluate the anti-cancer efficacy of the Wnt inhibitor LGK974 on *LKB1*-deficiency induced tumorigenesis in a mouse model of endometrial cancer in collaboration with Dr. Noelle Williams (Department of Biochemistry, UTSW). I will be using a well-described mouse model of endometrial cancer with Sprr2f-Cre driver for conditional *LKB1* deletion within the endometrial epithelium [153, 154]. Overall, the results of my studies will determine the viability of small molecule modulators of Wnt signaling pathway as molecularly targeted therapeutics for cancers associated with *LKB1* mutation.

CHAPTER FOUR

Epigenetic mechanisms of genome buffering revealed by CRISPR-CAS9 failures

Introduction

Peutz-Jeghers patients with *LKB1* mutation have an increased disposition for developing cancerous malignancies. The lifetime risk for developing some form of cancer in PJS patients can be as high as 93%, with 11-36% developing pancreatic carcinoma [155-157]. *LKB1* is also sporadically mutated in 20% of pancreatic ductal adenocarcinomas [158]. Due to the frequent association of *LKB1* mutations and pancreatic cancer, and our interest in understanding how loss of *LKB1* in epithelial cells give rise to increased Wnt signaling, I pursued an *LKB1* knockout strategy in a pancreatic cancer cell line (Mia cells) using CRISPR-Cas9 technology. CRISPR-Cas9-mediated genome editing has revolutionized our approach to understanding gene function, and promises to transform the management of genetically based diseases.

Cas9-mediated double strand breaks (DSB) activate the error-prone Non-Homologous End Joining (NHEJ) pathway which frequently results in small nucleotide insertions or deletions (INDELs) [159]. The Cas9 enzyme is commonly leveraged as a tool for achieving gene deletions in a variety of cell types. The introduction of INDELs can bring about a coding frame shift and thus the installation of a pre-mature termination codon (PTC) within the gene. If this PTC is located within 50-55 nucleotides upstream of the last exon-exon junction (referred to as the "50-55 nucleotide" rule), a nonsense mediated decay (NMD) complex will be assembled during the pioneer round of translation and the INDEL-containing mRNA will be subsequently destroyed

[160]. However, PTCs introduced less than 50 bp from the final exon-exon junction fail to elicit NMD and thus a C-terminally truncated protein may emerge in lieu of a complete protein loss [161, 162]. To limit the frequency of this outcome, the majority of sgRNAs used in gene deletion strategies target sequences proximal to the translation initiation [163].

Codon usage bias exerts a powerful influence on the mRNA life cycle and impacts the efficiency of transcription and splicing, and mRNA stability [164]. At the same time, it can dictate the local rate of translation elongation and thus the folding of nascent polypeptides [165]. Despite the understanding that genetic mutations encoded by mRNAs can affect protein expression outcomes in ways that are independent of any protein sequence alterations, the impact of INDELs such as those created by CRISPR-Cas9 or other genome-editing systems on gene expression are nevertheless mostly calculated using codon change predictions. The influence of INDELs on the life cycle of targeted mRNA and the protein that it encodes remains mostly uninvestigated.

In my efforts to leverage CRISPR technology to achieve a cell line devoid of LKB1 expression, I unexpectedly observed a frequent induction of novel forms of LKB1 in CRISPR-Cas9-edited cells. In this chapter, I describe a series of studies that reveal how these foreign proteins are generated by the cellular machinery that normally promotes fidelity in mRNA splicing and translation when it encounters INDEL-type mutations such as those introduced by CRISPR-Cas9. I then reveal using a commonly used genome scale screening library and commercially available CRISPR-edited cell lines the impact of my observations on the practice of genome editing that rely on INDEL-based mutagenesis. Overall, my studies reveal major blind spots associated with the interpretation of CRISPR-Cas9 associated experimentation and at

the same time highlight new strategies for leveraging this system to achieve research and therapeutic goals.

Induction of foreign proteins as a consequence of CRISPR introduced insertion-deletion changes (INDELs)

The introduction of genomic INDELs by CRISPR-Cas9 gene editing frequently generates PTCs which when recognized by a ribosome initiates NMD and the subsequent destruction of mRNA. Typically, sgRNAs are designed to target the five prime region with the understanding that PTCs that emerge too close to the true termination codon would not induce NMD [166]. True termination codons do not elicit a NMD response due to the lack of exon junction complex (EJC) downstream of a translation termination codon [167]. Adhering to this common practice in the design of my sgRNAs, two sgRNAs targeting the first exon (canonical AUG encoding) of *LKB1* were selected. Cas9 expression DNA along with a single sgRNA targeting one of the two *LKB1* sequences (CGCCGCAAGCGGGCCAAGCTC or GGTGGTGGACCCGCAGCAGC) (Figure 4.1 A) were transfected into Mia cells and then treated with puromycin to select for the transfected cells. Western blot analysis of nearly 200 isolated clones revealed some that appear to have loss of wt LKB1 protein but also contained novel proteins recognized by two LKB1-specific antibodies (Figure 4.1 B). Genomic sequencing of selected CRISPR-Cas9 targeted clones confirmed the introduction of INDELs in the sgRNA-targeted sequence (Figure 4.1 C), but failed to account for the presence of novel LKB1 proteins that migrated either more rapidly or slowly on SDS-PAGE. The molecular mechanisms underlying this INDEL-associated induction of foreign proteins are discussed below.

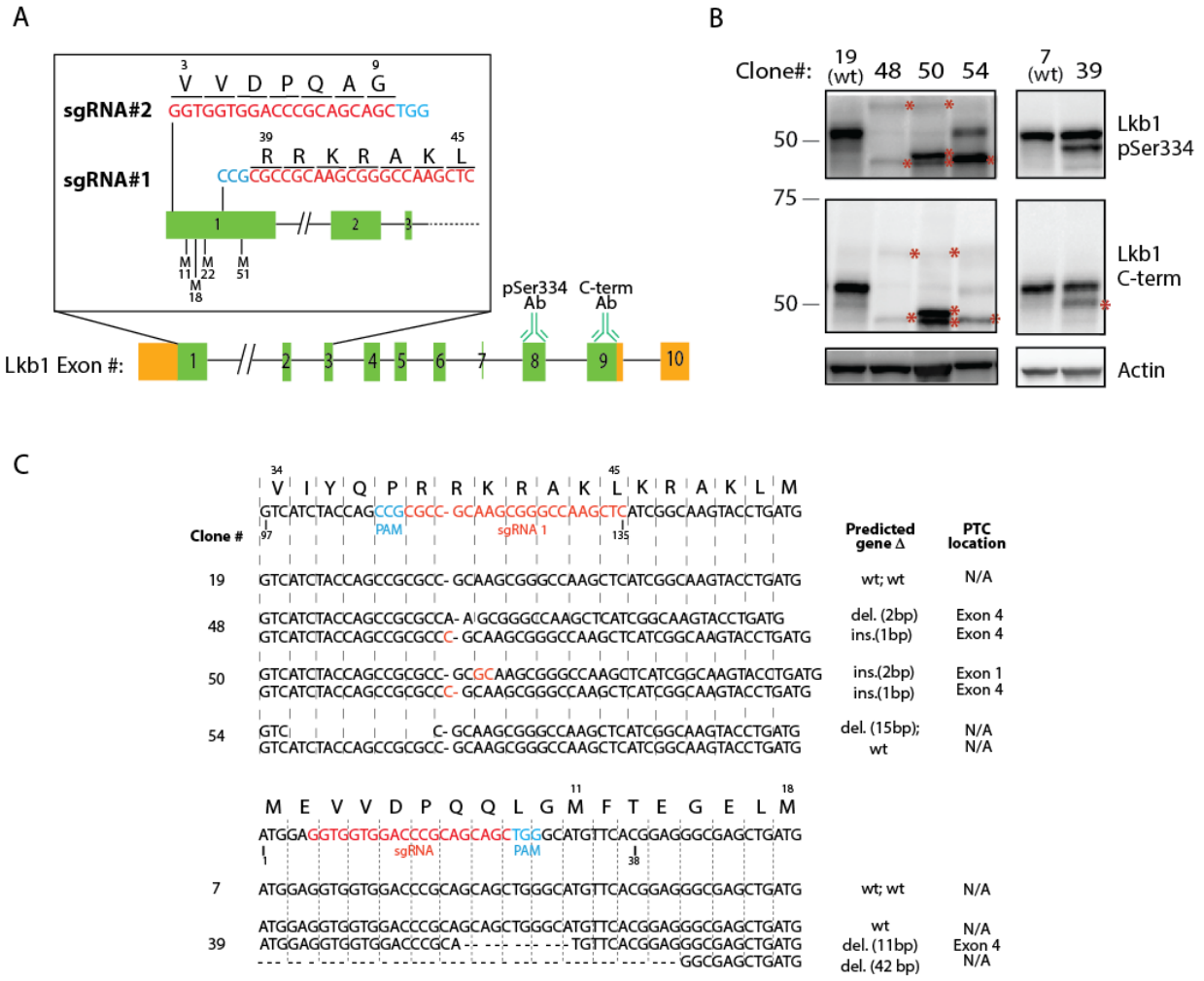


Figure 4.1: Induced expression of novel proteins from a CRISPR-Cas9 genome editing. (A) Genomic structure of the *LKB1* gene and the exonic sequences targeted by sgRNAs. (B) CRISPR-Cas9-edited MIA clones identified by Western blot analysis of nearly 200 clones. Asterisks indicate novel proteins recognized by Lkb1-specific antibodies (see A). (C) Genomic sequencing results of CRISPR-Cas9 edited clones. Gene alteration and anticipated location of pre-mature termination codon (PTC) due to frame shifting (if relevant) is indicated.

Inadvertent conversion of a pseudo-mRNA to a protein-coding mRNA by CRISPR-Cas9

In some CRISPR-edited clones (Clones 48 and 50, for example) a protein with a molecular mass larger than wt LKB1 was recognized by both the LKB1 pSS34 and the C-terminal antibodies. Having not found a molecular basis for the appearance of these novel proteins from

genomic sequencing, we employed RT-PCR analysis to examine for the presence of changes in mRNA products in the CRISPR-edited clones that might account for their existence (Figure 4.2 A). We observed in clones with larger LKB1 protein, a corresponding appearance of a 741bp mRNA species when using primers targeting LKB1 5' UTR and exon 4 for RT-PCR. Sequencing of the novel mRNA variant revealed the presence of a transcript that includes a 131 bp sequence not found in unedited cells (henceforth "cryptic exon"). We confirmed this insertion lies between exon 1 and exon 2 using an RT-PCR approach and primers that generate amplicons from different portions of LKB1 mRNA (Figure 4.2 B).

The *LKB1* cryptic exon is flanked by the canonical dinucleotide splice acceptor (AG) and splice donor (GT) sites [168, 169], suggesting it is a true exon and that one or more naturally occurring splice variants incorporate it. Indeed, a transcript harboring this exon likely does not produce a protein due to the presence of a PTC that has been reported [170]. However, with the CRISPR-introduced INDELS (1bp ins or 2 bps del), the PTC is eliminated resulting in a conversion of the pseudo-mRNA to a protein encoding transcript that supports translational termination at the native terminal codon. We termed the protein product from this long mRNA species as "super" LKB1.

We evaluated for the presence of the *LKB1* pseudo-mRNA in Mia cells by suppressing NMD using the protein translational inhibitor cycloheximide (CHX), which blocks the pioneer round of mRNA translation and thus the cell's ability to sense PTCs and engage the NMD apparatus [171, 172]. In the presence of CHX, the large molecular weight mRNA found in two of the CRISPR-Cas9 edited cells becomes apparent by RT-PCR (Figure 4.2 C). The unexpected emergence of the Super LKB1 protein in our study thus stems from our incomplete inventory of the *LKB1*

splice variants and highlights a general challenge associated with genome editing technology such as CRISPR-Cas9 that creates INDEL-type mutations. Furthermore, pseudo-messenger RNAs that are not predicted to produce proteins due to the presence of PTCs constitute as much as 10% of all transcripts, yet the significance of their wide-spread existence remains unclear [173].

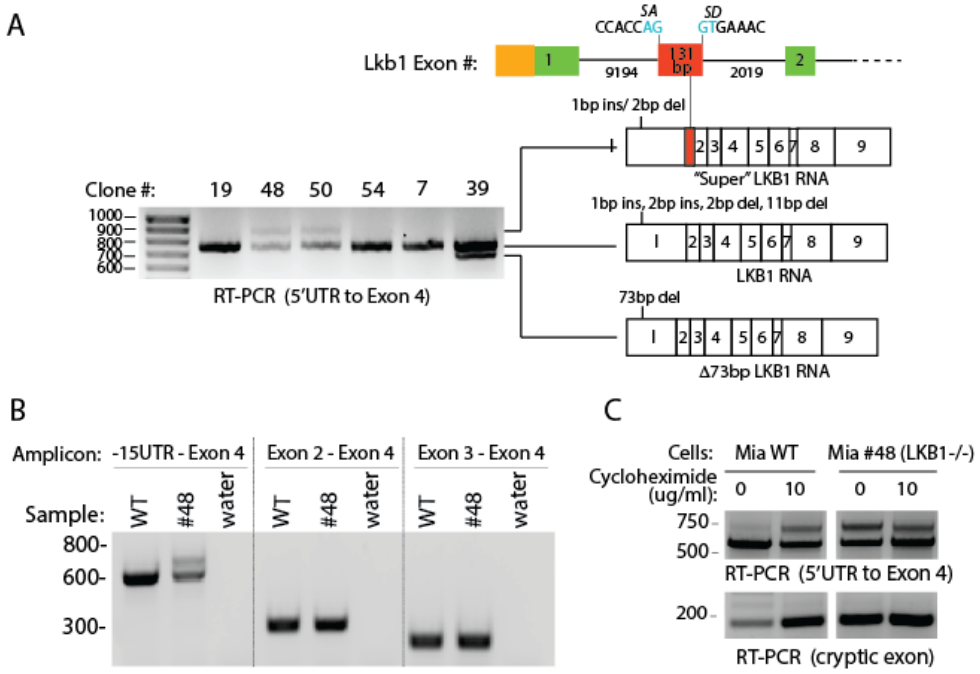


Figure 4.2: CRISPR-Cas9-mediated conversion of a pseudo-mRNA to a protein encoding mRNA. (A) An RT-PCR-based analysis of CRISPR-Cas9 edited clones reveals anticipated and unanticipated mRNA alterations. Primers mapping to the 5'UTR and Exon 4 in *LKB1* were used to generate amplicons from cDNA for sequencing. Clones 48 and 50, which express a larger molecular weight LKB1 protein ("super" Lkb1 protein; see Figure 4.1 B) harbor an mRNA species that includes an additional exon (termed "cryptic exon"). Clone 39 transcribes an *LKB1* mRNA species lacking 73bp as anticipated by genomic sequencing. (B) RT-PCR analysis of Clone 48 using indicated Lkb1 sequence-targeted primers confirms the CRISPR-Cas9 introduced mRNA alteration localizes between Exon 1 and 2. (C) MIA cells express a pseudo-mRNA that includes the "cryptic" exon. RT-PCR using primers targeting the cryptic exon reveal an Lkb1 mRNA species that is sensitive to the NMD inhibitor cycloheximide. In Clone 48, which harbors a CRISPR-Cas9 introduced single bp insertion, this splice variant is no longer sensitive to cycloheximide treatment.

Alternative translation initiation induced by CRISPR-Cas9

Our CRISPR-Cas9-based attempt to achieve a gene knockout also unexpectedly yielded LKB1 proteins that migrated faster than the wt LKB1 protein on SDS-PAGE (Figure 4.1 B). These smaller LKB1 proteins could have arisen from the proteosomal processing of a larger protein product such as Super LKB1. To test this hypothesis, we targeted the Super *LKB1* mRNA with an siRNA that engages the cryptic exon and found little effect on the expression of the short form of LKB1 even though the abundance of Super Lkb1 protein was effectively decreased (Figure 4.3 A). From this result and the absence of changes in mRNA splicing to account for the smaller LKB1 proteins that nevertheless retained reactivity to an antibody recognizing the C-terminal region of LKB1, we reasoned that alternative translational initiation (ATI) might contribute to the emergence of these new forms of LKB1 protein.

To determine if the mutations in *LKB1* found in our CRISPR-edited MIA clones contribute to the appearance of the foreign Lkb1 proteins, we characterized using SDS-PAGE the protein product generated from cDNAs representing a single mutated *LKB1* allele (in this case from Clone 48). Given that Clone 48 has two mutant *LKB1* alleles and that we have identified two major LKB1 splice variants in MIA cells (Super LKB1 and wt LKB1), we engineered and tested a total of four different cDNAs (see Fig. 4.3 B). LKB1 pseudo mRNA-derived cDNAs harboring either the 1bp insertion or 2bp deletion as found in the Clone 48 *LKB1* loci encode open reading frames (ORFs) that terminate at the native termination codon, and thus not surprisingly yielded proteins that co-migrated with Super LKB1 (Fig. 4.3 B). On the other hand, we were surprised to observe proteins generated from the wt LKB1 ORF harboring the same 1bp ins/2bp del alterations that were recognized by the C-terminus antibody given that they should encode

proteins that are prematurely terminated and that are precluded from expressing the antibody epitope (see Fig. 4.3 B). Having seemingly reconstituted all of the foreign LKB1 proteins observed in Clone 48 using engineered cDNAs as measured by Western blot analysis with the C-terminus antibody, we assumed that the short forms of LKB1 are generated from alternative translation initiation (ATI). Indeed, a cDNA encoding *LKB1* initiated from Met51, the first in frame methionine that follows the CRISPR-edited sequence co-migrates with the short LKB1 proteins. Furthermore, the Lkb1 cDNA constructs used in these reconstitution studies lack the native UTR sequence suggesting that these ATI events are due to the creation of internal ribosomal entry sites (IRESs) by CRISPR-introduced INDELS.

Using the same cDNA reconstitution approach, we next tested the ability to generate the foreign LKB1 protein observed in Clone 39. We engineered a UTR containing *LKB1* cDNA harboring the 73bp deletion (31 bp deletion of 5' UTR and 42 bp deletion of exon 1) as found in a mutant allele of clone 39. Similar to the previous results, we were able to recapitulate the expression of the short form of LKB1 observed in Clone 39 from the LKB1 cDNA with 73bp deletion (Figure 4.3 C). Moreover, a cDNA encoding *LKB1* initiated from Met18, the first in frame methionine that follows the CRISPR-edited sequence co-migrates with the short LKB1 protein of clone 39.

In summary, the mechanisms of NMD dictate the gene editing outcomes of CRISPR-Cas9 technology (Figure 4.3 D). The introduction of INDELs can result in the induction of novel proteins stemming from the alternative translation initiation mechanisms and/or stabilization of pseduoRNA transcripts. These unsuccessful gene deletion attempts provided new insights about

the mechanisms of translational control and highlighted unanticipated consequences associated with CRISPR-Cas9 technology.

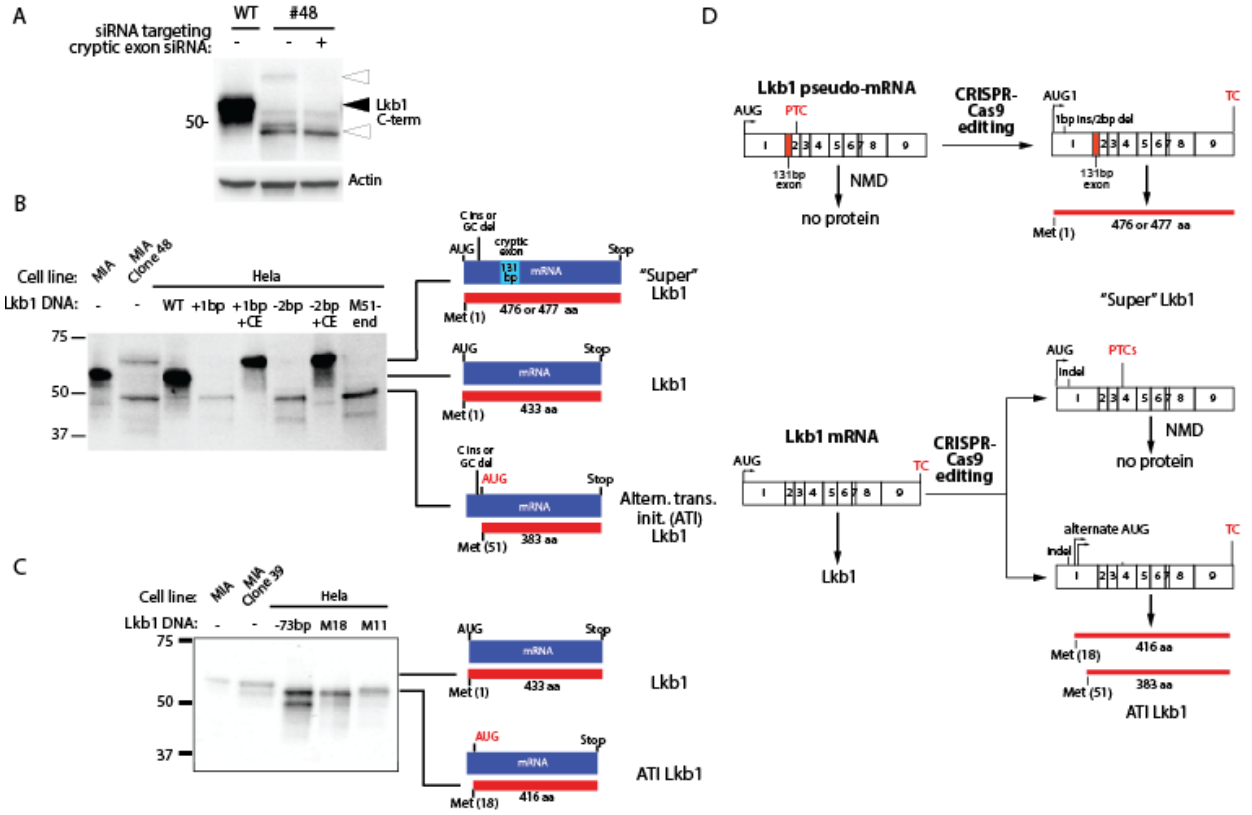


Figure 4.3: Evidence for alternative translation initiation (ATI) induced by CRISPR-Cas9 editing. (A) Short forms of LKB1 found in CRISPR-Cas9-edited cells are not derived from proteolytic processing of Super LKb1. An siRNA targeting the *LKB1* cryptic exon fails to reduce the expression of the short from of LKB1 in Clone 48. (B) A cDNA expression strategy for understanding allele-specific CRISPR-Cas9 introduced INDELs on protein expression. Lkb1 and Super Lkb1 cDNA expression constructs harboring genomic alterations found in *LKB1* of Clone 48 were introduced into Hela cells that lack endogenous Lkb1 expression. The 1bp and 2bp insertion and deletion, respectively, in the Super *LKB1* cDNA resulted in proteins that co-migrate with the large foreign protein observed in Clone 48. On the other hand, the same mutations in Lkb1 cDNA gave rise to proteins that co-migrated with the short foreign protein found in Clone 48. (C) An allele lacking the canonical AUG produces an N-terminally truncated protein. A similar strategy as described in "B" was used to interrogate an allele lacking a 73bp genomic DNA fragment that encompasses a portion of the 5'UTR region and 42bp of exon 1 (Clone 39). cDNAs expressing LKB1 proteins initiating at various AUGs proximal to the first AUG were used to generate a protein ladder for determining a predicted alternative translational

initiation site used in Clone 39. (D) Contribution of NMD and ATI to protein expression outcomes in CRISPR-Cas9 edited cells.

Cap-independent alternative translation initiation (ATI) as a result of CRISPR-generated internal ribosomal entry sites (IRESs)

Favoring target sequences proximal to the AUG sequence is thought to facilitate the introduction of PTCs away from the last exon-exon junction, which would fail to induce mRNA degradation by NMD [166]. Indeed, a survey of genome editing approaches using CRISPR-Cas9 technology for the goal of deleting gene function reveals a heavy reliance on the use of sgRNAs that target constitutive 5' coding exons [163]. However, AUG-proximal PTCs have also been reported to inefficiently trigger NMD primarily due to ribosomal re-initiation [166].

Two observations garnered so far suggest our ATI protein products are due to the introduction of IRESs: a) cDNAs encoding mutant *LKB1* alleles which harbor PTCs nevertheless are able to reconstitute the expression of an N-terminally truncated protein that co-migrates with the short form of LKB1 found in CRISPR-edited cells suggesting that alternative splicing does not contribute to the appearance of these foreign proteins, and b) the PTCs in the mutant *LKB1* mRNAs are found 3' to the predicted site of ATI thus precluding the possibility of ribosomal re-initiation that can occur in cases of aborted translation in regions in proximity of the AUG [174].

In order to further investigate my hypothesis that CRISPR introduces IRESs into the targeted gene sequence, I first identified an antibody that would enable me to recognize a short polypeptide that immediately follows the canonical AUG so that I can determine the effects of various INDELs on ribosomal scanning – a phenomenon associated with 5' cap-dependent translational initiation (Figure 4.4 A). The relative position of other potential initiation sites and

the CRISPR-edited *LKB1* sequence are also mapped. mRNAs that are generated by the spliceosome are associated with EJC protein complexes that then interact with PTCs to engage NMD. We utilized cDNAs encoding mutant *LKB1* alleles (in this case the 2bp deletion mutant allele) in order to bypass the effects of NMD. We confirmed that the "canonical" AUG is indeed still recognized in a mutant *LKB1* cDNA that harbors this deletion given the presence of a ~17kd LKB1 protein that can be isolated using the N-terminal antibody but not by a C-terminal antibody recognizing an epitope encoded in the last exon (Figure 4.4 B).

To generate additional evidence that the CRISPR-introduced INDELs contribute to cap-independent protein expression, we engineered a luciferased-based bicistronic reporter that is similar to those used in other studies to identify IRESs (Figure 4.4 C). A comparison of *LKB1* exon1 sequences from wt and mutant alleles inserted between the termination codon of the first reporter (Firefly luciferase; FL) and a second reporter (NanoLuc; NL), revealed increased 5'cap independent expression for constructs harboring the mutant *LKB1* sequence (Figure 4.4 D). In total, this evidence supports the ability of CRISPR-introduced INDELs to downstream generate internal ribosomal entry points thereby facilitating the production of N-terminally truncated proteins.

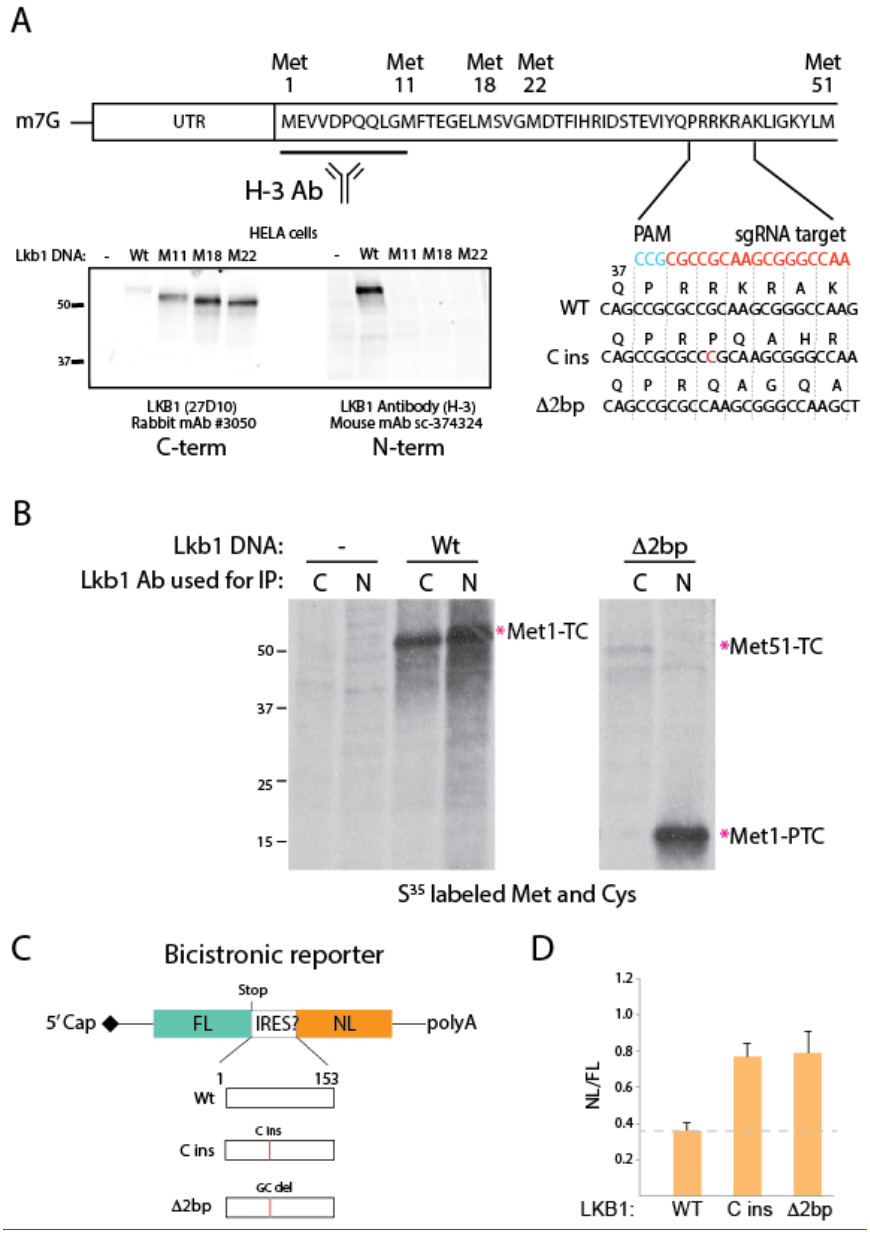


Figure 4.4: Cap-independent alternative translation initiation (ATI) as a result of CRISPR-generated novel internal ribosomal entry sites (IRES). (A) Relative location of potential translation initiation sites, epitope for the H-3 Lkb1 antibody, and CRISPR-introduced INDELs from Clone 48. Epitope mapping of the LKB1 N-terminal antibody was accomplished using lysate derived from cells transfected with *LKB1* cDNA expression constructs with varying initiating AUGs. (B) Preservation of canonical AUG-dependent translation initiation revealed by suppression of NMD. cDNA expression constructs encoding either wt or the Δ2bp Lkb1 alleles were transfected into HEK293 cells followed by pulse labeling for 45 min using S35-Met/Cys.

Lkb1 proteins were then immunoprecipitated using either the Lkb1 C- or N-terminal specific antibody. (C) A bicistronic luciferase-based reporter strategy for determining the effects of CRISPR-Cas9 introduced INDELs on 5' cap-independent translation initiation. FL=firefly luciferase, NL = NanoLuc luciferase. LKB1 Exon1 sequence from wt or mutant alleles were inserted between the termination codon of the FL sequence and the initiation site for the NL sequence. D. Bicistronic reporter constructs were transfected into HEK293 cells and NL/FL ratios were plotted.

IRESs created by secondary structure changes likely account for CRISPR-Cas9-induced ATI.

I compared the effects of missense mutations introduced at the same position as a CRISPR-induced single base pair insertion (C119) I had previously associated with the induction of an N-terminally truncated Lkb1 protein in MIA cells (Figure 4.5 A). Whereas the Lkb1 cDNA with the C119 insertion had given rise to the ATI product that likely initiates at Met51, the 3bp missense mutation did not (Figure 4.5 B), suggesting that structural changes local to the site of the INDEL are likely to account for internal ribosomal initiation.

Although once considered primarily a mechanism used by host cells infected with viral genomes to promote the translation of viral proteins, the advent of technologies that allow high resolution mapping of translation initiation sites such as ribosomal profiling has revealed a significant contribution of cap-independent translation initiation to the human proteome [175, 176]. Yet the majority of these events involve initiation in the UTR regions, which likely harbor secondary structure features that enable ribosomal entry [177]. The frequency in which CRISPR-Cas9-introduced INDELs in *LKB1* coding sequence promoted ATI likely reveals the importance of secondary structure in designating IRESs, a factor that significantly contributes to cap-independent translation initiation as determined by a massive screening effort using a bicistronic reporter akin to the one we have designed (Figure 4.5 C and Figure 4.4 C). This assertion is

supported by the absence of other factors that contribute to internal ribosomal entry including the presence of poly(U) and 18S ribosomal RNA complementary sequences preceding the site of ATI (see Fig. 4.5C).

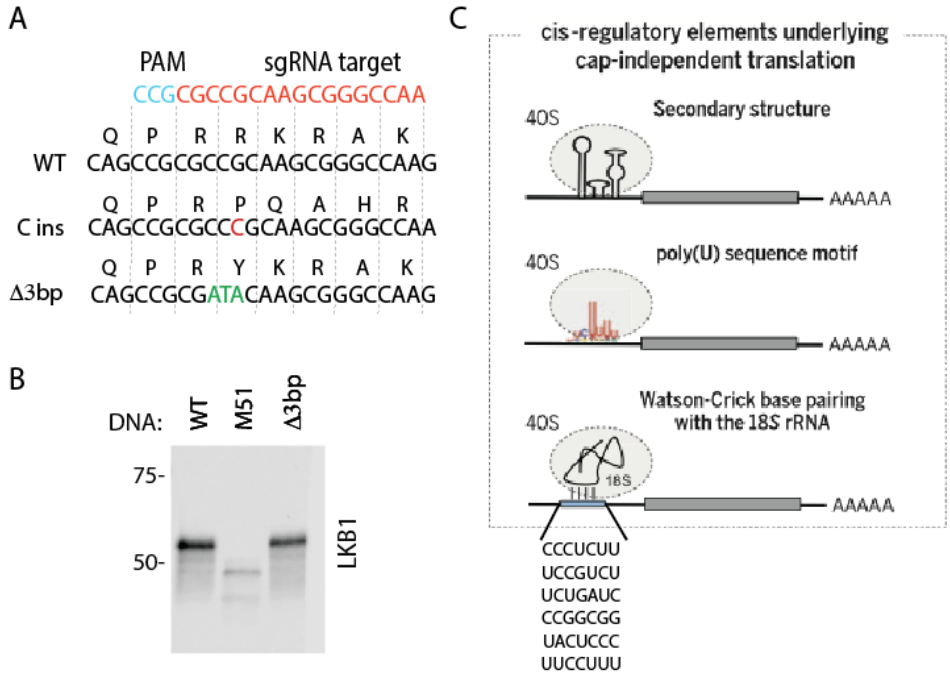


Figure 4.5: CRISPR-induced ATI likely stem from secondary structure changes in mRNA. A. Alignment of the C119 ins from CRISPR-edited MIA Clone 48 and a 3bp substitution in the same position introduced into the Lkb1 cDNA. B. A 3bp substitution introduced in the same position as INDELs that promoted Met51 initiated Lkb1 protein fails to promote ATI. Lysate from Hela cells transfected with cDNAs encoding wt or the Met51 Lkb1 protein, or the 3bp substitution mutation were subjected to Western blot analysis using the C-terminal Lkb1 antibody. C. mRNA features that contribute to cap-independent translation initiation identified from a systematic screen of 55,000 sequences using a bicistronic reporter platform (modified from [177]).

Induction of foreign proteins in commercially available CRISPR-Cas9 edited cell lines

Our CRISPR-Cas9 experiments in Mia cells revealed two unanticipated molecular response to the introduction of INDELs: ATI and pseudo-mRNA conversion. To investigate the frequency of these events in CRISPR-Cas9 edited cells generated using commonly used resources, we

examined a panel of purported knockout cell lines that are commercially available. Haploid human cells (Hap1 cells) with frameshift mutations in genes important in the Wnt signaling pathway were purchased from Horizon genomics. In each of the knockout cell lines purchased, the CRISPR-introduced INDELs are anticipated to create a PTC atleast in some known transcript variants that should earmark the mRNAs for NMD-mediated degradation (Figure 4.6 A). However, Western blot analysis of cell lysates generated from the *CTNNB1* and *LRP6* "knockout" cells using two different antibodies revealed expression of novel faster migrating proteins in SDS-PAGE (Figure 4.6 B). Yet, RT-PCR analysis of these same cells showed no changes in mRNA splicing suggesting the shorter CTNNB1 and LRP6 proteins are likely the products of alternative translation initiation events (Figure 4.6 C). In summary, the CRISPR-Cas9-mediated induction of foreign proteins through ATI is not cell-line or copy-number dependent since this phenomenon was observed in both the diploid Mia cells and the haploid Hap1 cells.

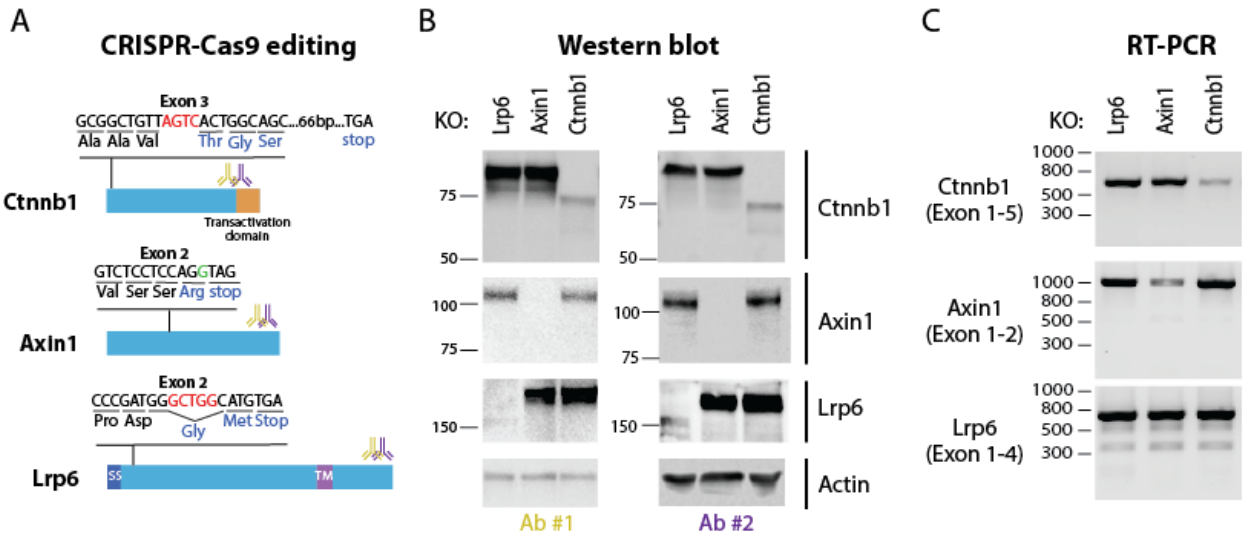


Figure 4.6: CRISPR-Cas9-induced alternative translation initiation in commercially available engineered cell lines. (A) Predicted frameshift alteration in CRISPR-Cas9 edited HAP1 clones (Horizon Discovery) relative to the recognition site of antibodies used in "B". Green=insertion; Red=deletion. (B) Western blot analysis of CRISPR-Cas9-targeted gene product using two distinct antibodies. (C) RT-PCR analysis and summary of sequencing results of CRISPR-Cas9-edited exon using primers recognizing indicated flanking exons.

Aberrant splicing due to compromised exon splicing enhancers (ESEs)

Having encountered unanticipated challenges for achieving a gene knock-out using sgRNAs that recognize AUG proximal sequences, we expanded our sgRNA targeting range to include exons from different positions within a gene this time using the tumor suppressor Supressor of Fused (SUFU) gene as a case study. SUFU acts as a nuclear exporter for the Gli family of transcriptional regulators, and functions as a key brake on Hh signaling [178, 179]. Somatic mutations in *SUFU* occur in a subset of patients with medulloblastoma [179]. pCas-guide plasmids containing sgRNAs targeting exon 2, 3 or 8 were transfected in RMS13 cells and approximately 60 clones from each transfection pool was analyzed by Western blot analysis for changes in SUFU expression status (Figure 4.7 A). Unlike in the case of LKB1, all four sgRNAs targeting SUFU were successful at seemingly achieving *SUFU* knockouts (Figure 4.7 B). However, given our previous observations with LKB1, we were nevertheless wary of potential unanticipated post-transcriptional phenomena. We first performed RT-PCR analysis on CRISPR-edited *SUFU* clones R1 and R7 generated using sgRNAs 3a and 3b respectively (Figure 4.7 C). To our surprise, despite both of these clones were seemingly devoid of SUFU expression, one of these (Clone R1) transcribed an additional mRNA species that would be predicted to be shorter than the native *SUFU* transcript. Sequencing of the smaller amplicon revealed the absence of exon 3, the exon targeted by the sgRNA 3a. Taken together with the evidence for on-target gene

editing (Figure 4.7 D), we assume the introduction of the 26 bp deletion by CRISPR-Cas9 compromised an exon splicing regulatory feature found in exon 3.

We noted that Clone R7 which was subjected to editing using an sgRNA immediately adjacent to sgRNA 3a used to generate Clone R1 failed to induce exon skipping of the targeted exon (Figure 4.7 C). Numerous *cis*-acting elements exist within exons and introns that regulate alternative splicing. These splicing regulatory elements (SREs) on the basis of their locations and activities are categorized as exonic splicing enhancers (ESEs), intronic splicing enhancers (ISEs), exonic splicing silencers (ESSs) and intronic splicing silencers (ISSs) [180, 181]. ESEs are prevalent in all exons, including constitutive and alternate exons [182] and serve as a binding site for Serine-Arginine (SR) protein family, which functions to activate splicing [183]. Using RESCUE-ESE Web server (http://genes.mit.edu/burgelab/rescue-ese/) that identifies 238 candidate ESE hexanucleotides, we analyzed for the presence of putative ESEs in the sequences targeted by our *SUFU* sgRNAs 3a and 3b. Whereas the 26bp deletion in the CRISPR-edited R1 clone eliminated a putative ESE sequence (AACATG; Figure 4.7 D), there were no predicted ESEs in the genomic sequence altered in the R7 clone. Therefore, disruption of an ESE by sgRNA 3a but not 3b appears to result in the skipping of the targeted exon.

We examined the remaining CRISPR-edited cell lines generated using sgRNAs targeting sequences in Exon 2 and 8 in order to determine if our observations above were specific to Exon 3. Using an RT-PCR strategy as before, we observed in both Exon 2 and 8 targeted cell lines, the appearance of a smaller molecular weight amplicon generated using primers flanking the CRISPR-targeted exon (Fig. 4.7 E and 4.7 G). Sequencing of the cDNA amplicons confirmed skipping of the targeted exons in both of representative clones (Figure 4.7 F and 4.7 H). We

noted in these instances that an ESE as determined using RESCUE-ESE Web server was present in these targeted sequences (Figure 4.7 F and 4.7 H). Thus, three out of four sgRNAs targeting SUFU resulted in the inadvertent exclusion of the sgRNA targeted exon. In the case of SUFU, the resulting transcripts due to exon skipping nevertheless produced a PTC due to a shift in codon reading frame. At the same time, the prevalence of these events suggests that in many instances in which the reading frame is preserved following an exon-skipping event, the introduced INDEL would be bypassed thus rendering genomic sequenced-based predictions of protein expression inaccurate.

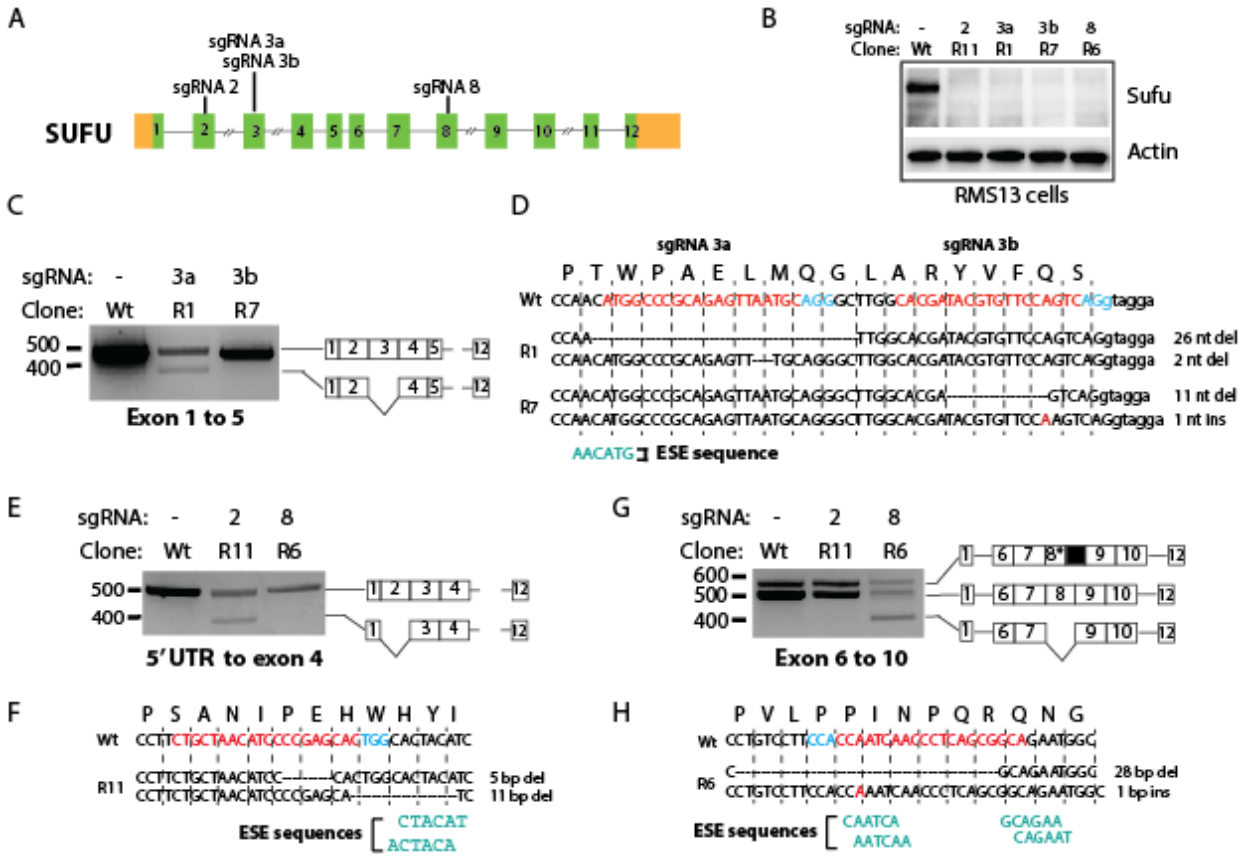


Figure 4.7: Skipping of CRISPR-Cas9 targeted exons is associated with the disruption of predicted exon splicing enhancers (ESEs). (A) SUFU exonic structure and sequences targeted by sgRNAs. (B) Western blot analysis of lysates derived from the CRISPR-Cas9-edited RMS13 clones using indicated sgRNAs. (C) RT-PCR analysis and summary of cDNA sequencing results of R1 and R7 clones using primers targeting SUFU exon 1 and 5. (D) Genomic sequencing of the SUFU exon 3 from the R1 and R7 clones targeted by sgRNAs 3a and 3b. A predicted ESE site (AACATG) found in the sequences targeted by sgRNA 3a was identified using RESCUE-ESE Web Server (http://genes.mit.edu/burgelab/rescue-ese/). (E) RT-PCR and summary of cDNA sequencing results from the R11 and R6 clones using a primer pair that generates an amplicon extending from the 5'UTR to Exon 4. (F) Genomic sequencing of the SUFU Exon 2 region targeted by sgRNA 2 in Clone R11 and R6 and predicted ESE sequences. (G) RT-PCR and summary of cDNA sequencing results from the R11 and R6 clones using a primer pair that generates an amplicon extending from Exon 6 to Exon 10. (H) Genomic sequencing of the SUFU Exon 8 region targeted by sgRNA 8 in Clones R11 and R6 and predicted ESE sequences.

Exon skipping in commercially available CRISPR-Cas9-edited cell lines produces internally truncated proteins

We had previously observed induction of foreign proteins by ATI in commercially available CRISPR-edited Hap1 cell lines (Figure 4.6 B). To determine if exon skipping is a common phenomenon in commercially engineered cell lines using CRISPR-Cas9, we purposefully selected Hap1 knockout cell lines generated using sgRNAs targeting a symmetrical exon (exons whose length is an exact multiple of 3) in order to evaluate the frequency of obtaining an internally truncated protein if the targeted exon is indeed skipped (Figure 4.8 A). Western blot analysis of lysates derived from each of these cell lines were examined using two distinct antibodies (Figure 4.8 B). In two out of five cell lines examined (those targeting TOP1 and SIRT1), we detected a novel protein that migrated faster than the wild-type protein. RT-PCR analysis of the same cell line panel revealed a co-incidental appearance of smaller molecular amplicons generated using primers recognizing exons flanking the sgRNA targeted exon (Figure 4.8 C). In addition, a cell line that appears to lack expression of any protein product subsequent

to gene editing exhibited altered splicing with the appearance of both larger and smaller amplicons (the Vps35 cell line). Sequencing of amplicons corresponding to the non-native mRNA species found in this cell line panel reveals exon skipping of the targeted exon in the TOP1 and SIRT1 KO cell lines and stabilization of pseudo mRNAs in the case of VSP35. Thus, in ten commercially available CRISPR-edited cell lines with purported eliminated expression of a targeted gene, five exhibited altered mRNA splicing that either resulted in the emergence of a truncated protein product in lieu of a complete loss in protein expression, or the stabilization of a pseudo mRNA presumably due to suppression of NMD.

Exon skipping induced by CRISPR-Cas9 is independent of persistent sgRNA expression.

The commercially available CRISPR-edited cell lines as well as those we generated targeting *LKB1* in MIA cells both utilize DNA constructs to introduce the desired sgRNA. To investigate the possibility that persistent expression of an sgRNA promotes exon skipping events, we transfected a RNA guide complex consisting of a CRISPR RNA (crRNA) targeting exon 5 of *LKB1* and a transactivating RNA (tracrRNA) that contains Cas9 nuclease-recruiting sequence into HAP1 cells (Dharmacon's Edit-R CRISPR-Cas9 platform; Figure 4.8 D). The transient expression of synthetic RNAs eliminates the need for cloning the sgRNA sequence in a vector backbone and also abolishes any unwanted effects that may arise due to the constitutive expression of an sgRNA construct. Unlike our previous attempt to knockout LKB1 using sgRNAs targeting exon 1, targeting exon 5 of *LKB1* generated a cell line (Clone H2) with complete loss of LKB1 expression that was consistent with evidence for on-target gene editing by genomic sequencing (Figure 4.8 E, and Figure 4.8 F). However, RT-PCR analysis of the H2 clone again using primers flanking the CRISPR-targeted exon followed by sequencing of the

generated amplicons revealed yet again the appearance of a transcript devoid of the targeted exon (Figure 4.8 G). Given the impermanence of the crRNA/tracRNA complexes, we assume the induction of exon skipping observed in CRISPR-Cas9 experimentation is not due to persistent expression of sgRNAs. In total, our analysis of a small panel of commercially available CRISPR-edited cell lines with purported gene knock-outs revealed a significant percent exhibited unanticipated effects at the RNA and/or protein expression level (Figure 4.8 H).

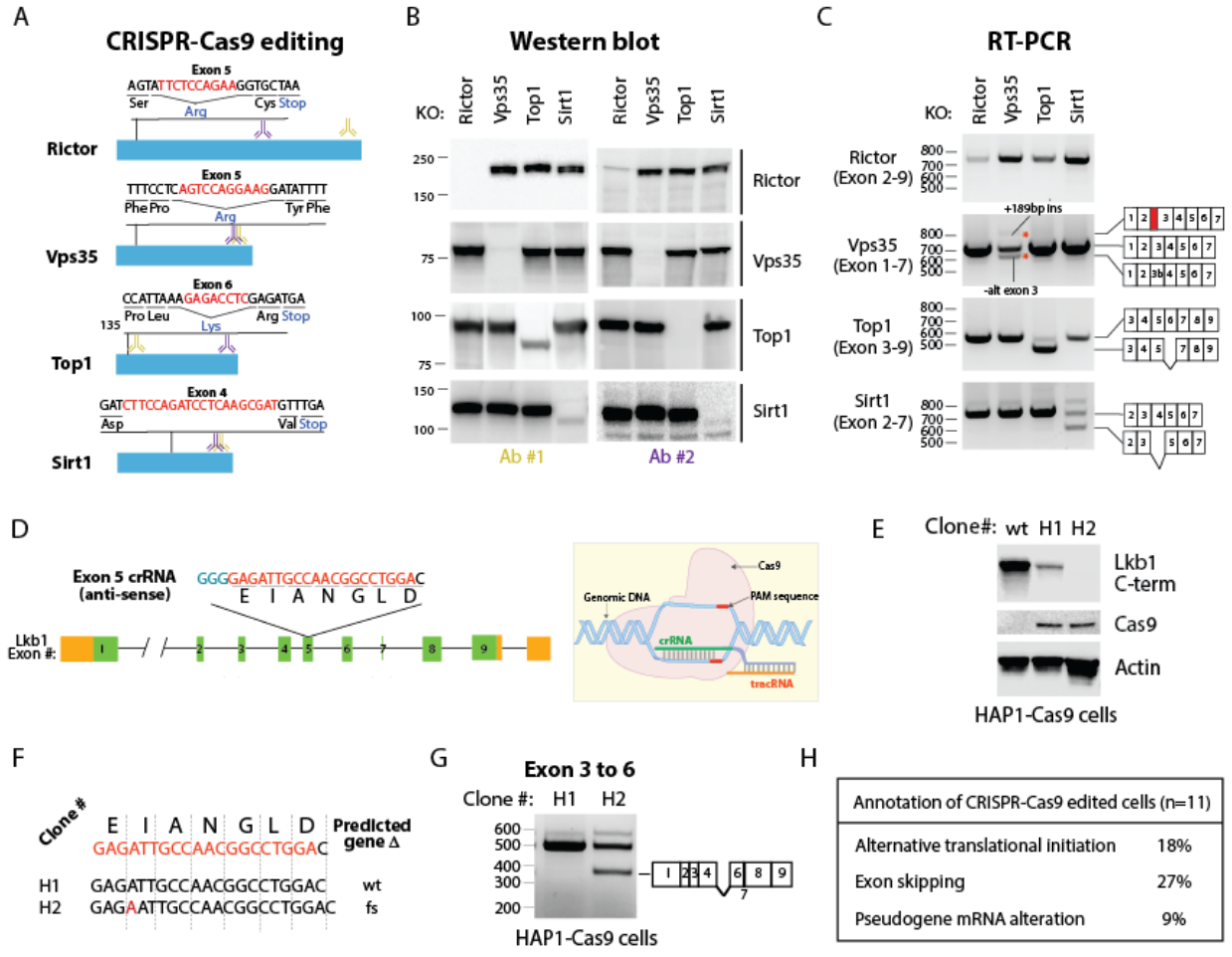


Figure 4.8: CRISPR-Cas9-induced exon skipping is not dependent upon persistent sgRNA expression. (A) Predicted frameshift alteration in CRISPR-Cas9 edited HAP1 clones (Horizon

Discovery) relative to the recognition site of antibodies used in "B". Red=deletion. (B) Western blot analysis of CRISPR-Cas9-targeted gene product using two distinct antibodies. (C) RT-PCR analysis and summary of sequencing results of CRISPR-Cas9-edited exon using primers recognizing indicated flanking exons. (D) A crRNA based strategy to target exon 5 of *LKB1*. (E) Western blot analysis of lysate derived from the wt Hap1 clone H1 and the CRISPR-Cas9-edited clone H2. (F) Genomic sequencing results from clones analyzed in "E". (G) RT-PCR analysis of targeted exon using primers recognizing flanking exons reveals the emergence of a novel *LKB1* transcript. Sequencing of cDNAs corresponding to the novel transcript reveals skipping of the targeted exon 5. (H) Frequency of CRISPR-Cas9-mediated induction of novel proteins in experimentally generated and commercially purchased Hap1 cell lines.

CRISPR-induced mislocalization of proteins

We had previously observed the unexpected induction of novel Lrp6 and Top1 proteins in commercially available "knock-out" cell lines (Figures 4.6 and 4.8). In order to analyze characteristics of these truncated proteins, we first examined the protein domain features that were compromised due to exclusion of the exonic regions. The faster migrating Lrp6 protein is likely generated by ATI after bypassing the CRISPR-Cas9-mediated Δ4bp deletion in exon 2 (Figure 4.9 A). The N-terminally truncated Lrp6 protein therefore likely lacks the signal sequence (SS) encoded by the first nineteen amino acids [184]. The elimination of the SS should result in the cytoplasmic delivery of the novel protein as suggested by the absence secretory-pathway associated glycosylation (Figure 4.9 B). Given that many protein localization signals such as signal sequences are found on the N-terminal region of proteins, the emphasis on the attack of the 5'end of genes using CRISPR-editing enzymes suggest that ATI events will likely result in not only N-terminally truncated proteins but also those that are mislocalized.

The novel Top1 protein was likely generated after the skipping of the CRISPR-targeted exon 6 (Figure 4.9 C). The exclusion of exon 6 eliminates a major Top1 sumoylation site (Lys117), which is essential for protein localization [185]. In our internally truncated Top1 protein, we

observed an increased cytoplasmic pool of the truncated Top1 protein compared to the wild-type Top1 protein (Figure 4.9 D). Given the importance of Top1 protein in DNA relaxation, the foreign Top1 protein coupled with its deviant subcellular distribution likely will have unpredictable contributions to the cell biology of this cell line. These examples of unexpected foreign protein production coupled with their aberrant subcellular site of accumulation illustrated by the LRP6 and TOP1 cell lines reveals the limitations of using only genomic sequencing data for genotype-phenotype correlations.

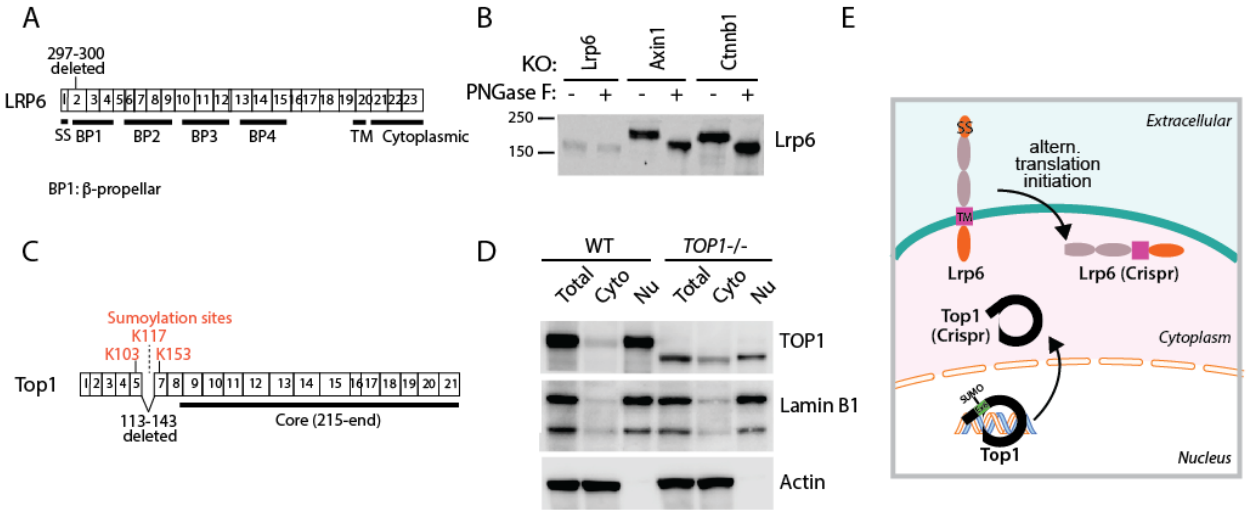


Figure 4.9: Truncated proteins generated by CRISPR-Cas9 gene editing have altered subcellular localization. (A) Exonic sequences encoding protein landmarks in LRP6. (B) ATI in a CRISPR-edited HAP1 cell line results in the production of an intracellular form of LRP6. An IRES introduced by an IDEL (see Fig. 4.7) likely gave rise to a protein lacking a signal sequence and that is not subject to secretory pathway-associated glycosylation. PNGAse F is a deglycosidase. (C) Exonic sequences encoding protein landmarks in TOP1. (D) Cell fractionation of wt and the putative TOP1 null clone (see Fig. 4.5) reveals distorted distribution of an internally truncated TOP1 protein compared to wt protein. (E) Model of CRISPR-induced protein localization events.

Prevalence of exonic exclusion in a therapeutically relevant CRISPR-Cas9-sgRNA delivery approach

Mutations associated with congenital diseases found in symmetric exons could be eliminated from incorporation into the holo-transcript to yield a functional protein by inducing exon skipping using various approaches [186, 187]. Duchenne's muscular dystrophy (DMD) is caused by any number of mutations in dystrophin, a multiexonic gene that forms the testing ground for many exon-skipping technologies. Whereas anti-sense oligos targeting splice acceptor/donor (SA/SD) sequences found on the ends of gene exons have moved forward in clinical testing as agents for inducing exon skipping in DMD patients [188, 189], the advent of CRISPR technology affords an alternative approach based on permanent alteration of the patient's genomic DNA [190-192]. Exon skipping with CRISPR has so far been achieved by using viral mediated delivery of sgRNA pairs that effectively excise the compromised exon. Our observations suggest that: a) single sgRNAs targeting ESEs can be potentially as effective as the use of sgRNA pairs thus decreasing the off-targeting potential associated with two sgRNAs, b) use of a single sgRNA should minimize the DNA damage response that can promote cell senescence subsequent to CRISPR attack, and c) including ESE targeting as a therapeutic option should greatly expand the number of viable target sequences available for inducing exon skipping. At the same time, unwanted exon skipping events due to the inadvertent disruption of an ESE or other exon localized splicing regulatory features, could compromise targeted exon skipping agendas.

To test the frequency of a virus-mediated sgRNA delivery approach for inadvertently inducing exon skipping, we established an *in vitro* assay for measuring exon skipping events that

incorporates random sgRNAs isolated from the Gecko V2 library, a CRISPR screening resource commonly used for *in vitro* genome-scale screens. Lentivirus carrying RNA encoding a single sgRNA targeting an exon that is flanked by at least one protein-encoding exon was delivered to MIA cells and the pool of infected cells (selected for by drug resistance) were analyzed by RT-PCR (Figure 4.10 A). Examples of the types of outcomes encountered by RT-PCR for experiments in which genomic sequencing results confirm genomic editing at the target sequence are shown (Figure 4.10 B). A random sequencing of two of these amplicons confirmed elimination of the targeted exon in novel amplicons detected in the CRISPR-edited cell line (Figure 4.10 C). We confirmed using cycloheximide treatment to suppress translation and thus NMD that these shortened amplicons that are ultimately generated from *de novo* transcripts are induced by CRISPR editing (Figure 4.10 D). In total, from experiments in which gene editing can be observed in the sgRNA target sequence, we observed ~50% of sgRNAs induced alterations in mRNA splicing (Figure 4.10 E). We assume these exon skipping events are largely driven by alteration of exonic sequences and not from perturbations within the SA/SD sequences due to the following arguments: a) CRISPR typically introduces small INDELs [193], b) experiments considered in our analysis contained lesions that are predominantly localized within the sgRNA targeted sequence (Appendix A) and c) the average distance of sgRNA targeted sequences from those that induced or did not induce exon skipping were comparable (Appendix B).

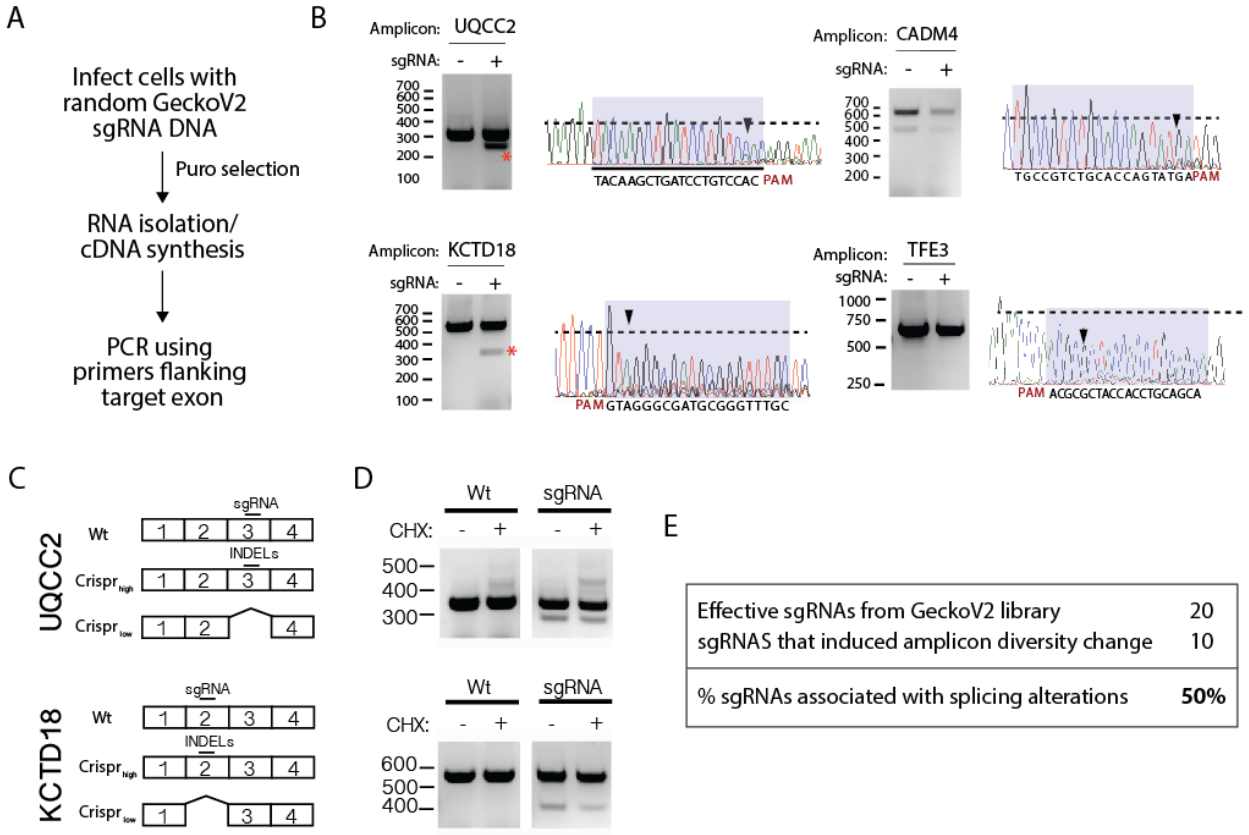


Figure 4.10: Prevalence of exonic exclusion in a therapeutically relevant CRISPR-Cas9-sgRNA delivery approach. (A) A set of sgRNAs representative of those typically found in genome-scale screening libraries for evaluating the frequency of exon skipping was generated by subcloning random sgRNA expression constructs from the GeckoV2 sgRNA library. cDNA isolated from cells infected with a single sgRNA delivered by lentivirus was subjected to PCR analysis using primers recognizing exons flanking the targeted exon. (B) Examples of RT-PCR and genomic sequencing results to determine for the presence of sgRNA-induced exon skipping events and successful Crispr-Cas9 editing, respectively. (C) Summary of cDNA sequencing results for two pools of Mia cells treated with indicated sgRNAs. (D) Evidence for *de novo* splicing events induced by Crispr-Cas9 editing. Suppression of NMD with cycloheximide did not induce alternative splice variants unique to CRISPR-edited cells. (E) Prevalence of exon skipping phenomena associated with lentivirally delivered sgRNAs.

Anticipating INDEL-induced changes in gene expression

My study using CRISPR-Cas9 initially focused on understanding the role of the *LKB1* tumor suppressor gene in cancer has instead returned a deeper understanding of how cells process

mRNA that harbor INDELs. First, I observed frequent exon skipping in CRISPR-targeted exons that can either result in the introduction of a PTC and NMD engagement or in the case of symmetric exons targeted a novel protein with an internally truncated sequence (Figure 4.11 A). Given the prevalence of ESEs and ESSs and likely other splicing regulatory features not yet described found in exonic sequences [194], we anticipate that exon skipping events will pose a significant challenge to INDEL-mediated mutagenesis strategies. On the hand, in transcripts that harbor INDELs but did not undergo exon skipping, we observed ATI that result from the creation of IRESs (Figure 4.11 B). The end products in these instances are N-terminally truncated proteins. Finally, 10% of the transcriptome are thought to be pseudo mRNAs [173]and the overall transcriptome with an average of 7 alternative splicing events per gene [195] suggest that the blue print available for predicting CRISPR-Cas9 consequences is mostly incomplete (Figure 4.11 C). Given the enormity of the problem associated with this burgeoning technology, we need computational firepower in order to leverage these new observations into constructive strategies for achieving therapeutic goals.

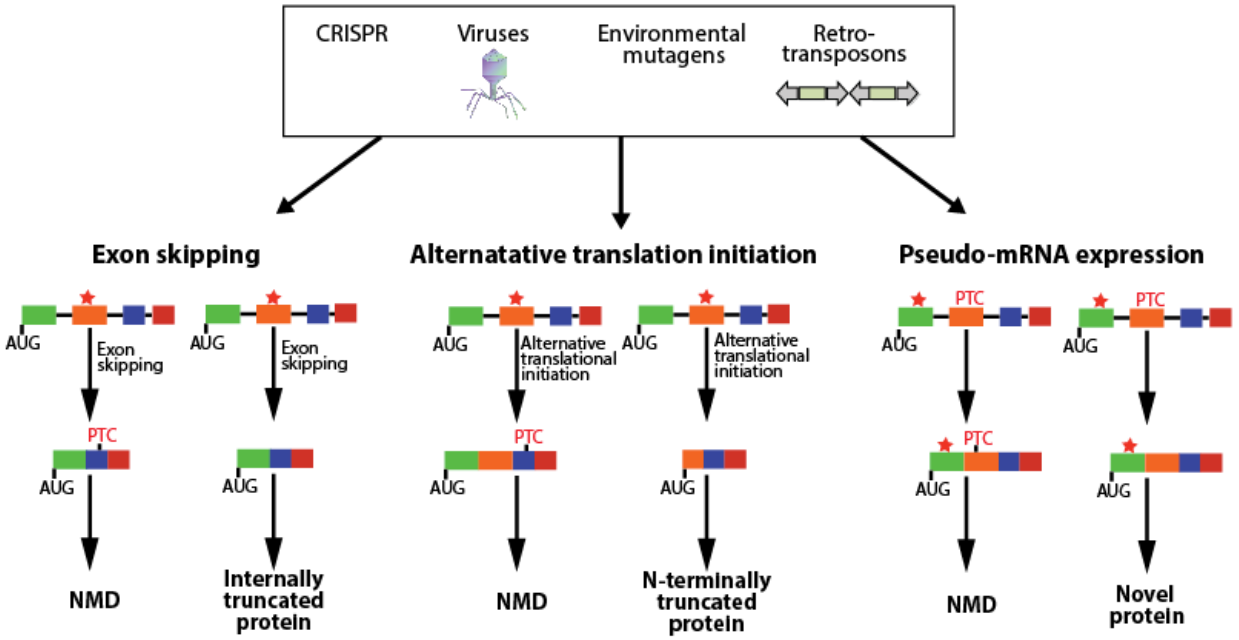


Figure 4.11: Cellular responses to INDELs revealed by CRISPR-Cas9 genome editing. INDELs introduced either by environmental mutagens (viruses, UV radiation etc) or DNA cutting enzymes (CRISP-Cas9, Cpf1, etc) may be processed by the splicing and translation machinery to yield protein products not anticipated by merely considering codon alterations.

In collaboration with Dr. Tae Hyun Hwang and Dr. Yunku Yeu (Department of Clinical Science, UTSW). we have developed a computational tool that we have named the CRISPinator (Figure 4.12). A primary motivation for developing this resource is to facilitate the identification of guide RNAs for various genome-editing enzyme systems with the potential for compromising ESE/ESS sequences and that might elicit exon skipping either purposefully for protein engineering in the case of symmetric exons targeted or gene deletion agendas in the case of asymmetric exons. Indeed, sgRNAs that target ESE sequences could be an effective approach to inducing the skipping of mutated exons in the case of inherited diseases. At the same time, the CRISPinatoR could be used to anticipate the emergence of novel proteins due to INDELs

introduced into pseudo mRNAs. Implicit in both motivations is the potential for improving the understanding of phenotypes in cell lines associated with INDEL-type mutagenesis. At this time, we do not have a computational approach for predicting when IRESs will emerge and thus potential ATI events will have to be empirically determined using cDNA reconstitution assays such as those described here (Figure 4.12 B).

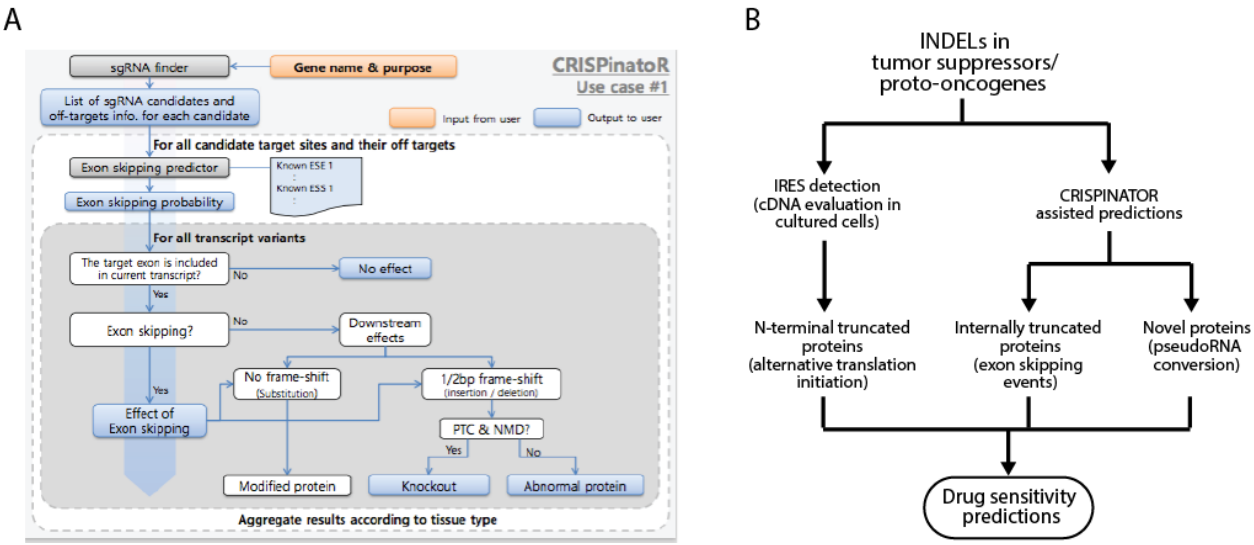


Figure 4.12: CRISPinatoR: an *in silico* platform for predicting gene products associated with INDEL mutagenized coding sequences. (A) Decision-making workflow for a beta version of the CRISPinatoR. (B) A personalized medicine-focused strategy for predicting cancer-relevant INDEL effects on gene expression using the CRISPinatoR and IRES detection in cDNA transfected cultured cells.

Discussion/ Conclusions

The observations described here have broad implications for INDEL-associated mutagenesis either in the context of carcinogenesis or directed genome editing efforts. First, inherent to INDEL-type mutagenesis of protein coding DNA is the generation of new codons due to frame-shifting. In addition to predicted changes in protein expression from assessing the

known transcriptome, these events will impact an estimated 10% of the genome designated as pseudo-mRNAs – mRNAs that harbor PTCs thus likely suppressed by NMD from contributing to the proteome [173]. Our incomplete inventory of these pseudo-mRNAs suggests that frame-shifting or exon skipping induced by CRISPR-Cas9 would occasionally give rise to novel proteins (such as Super Lkb1, for example). At the same time, the purpose of generating mRNAs that are subject to NMD is unclear and the introduction of INDELs may compromise other cellular functions performed by such RNAs in addition to those attributable to any novel protein expression.

An estimated 92–94% of human genes undergo alternative splicing [196]. The ease with which I was able to induce exon skipping using CRISPR-Cas9 opens new avenues for achieving both basic science and regenerative medicine goals. First, with the aid of the CRISPRinatoR, we should be able to develop sgRNAs that intentionally promote exon skipping and thus facilitate the study of protein isoforms that may be found in the host proteome perhaps in low abundance or novel proteins with value in protein engineering agendas. High-throughput structure function studies should also be achievable with genome-scale guide RNA libraries targeting sequences harboring ESE/ESE elements.

The frequent creation of IRESs by CRISPR-Cas9 mutagenesis observed in my studies suggest that targeted introduction of INDELs that do not impact exon splicing could be used to bypass deleterious mutations found proximal to the native translational initiation site or for structure function studies of important proteins. With improvements in our understanding of the rules for generating IRESs, genome-scale guide RNA libraries could be also generated for high-

throughput protein structure-function annotation.

Although the field of genomics in the past has primarily focused on single base pair substitutions as a means to delineate actionable gene alterations, a study that analyzed DNA sequencing data from 79 persons selected from diverse ethnic backgrounds mapped 1.96 million INDELs of which 2123 are anticipated to affect gene expression [197]. My observations relating to how cellular mRNA splicing and protein translational machinery perceive and deal with INDELs should improve our ability to anticipate gene expression changes and thereby better match available therapeutic agents with the underlying genetic basis of disease. I speculate that the challenges posed to eukaryotic cells by CRISPR-Cas9 mutagenesis have been visited before in the form of environmental mutagens such as viruses (Figure 4.10). Thus, targeted genome-editing efforts either for regenerative medicine or anti-cancer goals must circumnavigate the genome buffering mechanisms that likely have evolved to minimize the contribution of mRNAs harboring INDELs to the proteome.

Chapter 5: Materials and methods

Reagents

Cell lines

Wt and Lkb1 null mouse embryonic fibroblasts were received from Dr. Nabeel Bardeesy (Harvard Medical School). Porcn null mouse embryonic fibroblasts were provided by Dr. Charles Murtaugh (University of Utah). CRISPR-edited Hap1 knockout commercial cells were purchased from Horizon Genomics and are as follows:

Lrp6-/- Hap1 cells:	Cat# HZGHC000598c007
Ctnnb1-/- Hap1 cells:	Cat# HZGHC001032c004
Axin1-/- Hap1 cells:	Cat# HZGHC000355c011
Tbk1-/- Hap1 cells:	Cat# HZGHC000031c012
Sirt1-/- Hap1 cells:	Cat# HZGHC001008c008
Bap1-/- Hap1 cells:	Cat# HZGHC003319c004
Rictor-/- Hap1 cells:	Cat# HZGHC001020c010
Vps35-/- Hap1 cells:	Cat# HZGHC000863c012
Top1-/- Hap1 cells:	Cat# HZGHC001210c003
Pten -/- Hap1 cells:	Cat# HZGHC000624c009

Antibodies

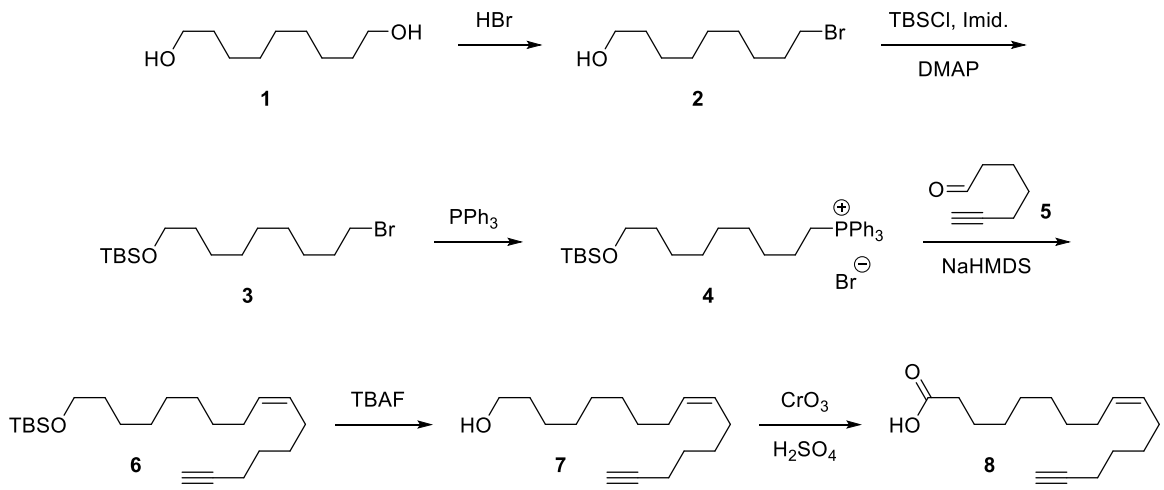
Antibodies were purchased from the following sources: Cell Signaling Technology (Dvl2, Myc, Wnt3a, Lkb1 C-term, Lkb1 p-Ser334, Lrp6, Axin1, Sirt1, Rictor, Sufu), Millipore

(WLS/Grp177, Ctnnb1), Bethyl Laboratories (Sirt1, Top1, Vps35, Rictor), Santa Cruz Biotechnology (Lkb1 N-term, Lrp6, Fc-HRP), Genetex (Vps35), and Biolegend (Strep-HRP).

Chemicals

IWP-2 was synthesized as previously described [71]. LGK974 (also known as Wnt974) was purchased from Active BioChem and cycloheximide was purchased from Sigma. SCD inhibitor A939572 was purchased from Biofine International. *Cis* and *trans* ω-alkynyl fatty acid probes were synthesized by Dr. Chuo Chen's group and is described as follows:

***Cis* ω-alkynyl palmitoleic acid synthesis:**



9-Bromononan-1-ol (2). To a solution of 1,9-nonanediol (**1**) (3.20 g, 20 mmol) in toluene (50 mL) was added 48% hydrogen bromide (28 mL, 160 mmol) at 0 °C and then heated to reflux. After stirring overnight, the solution was cooled to 23 °C, diluted with ethyl acetate (50 mL), and washed with saturated sodium chloride (50 mL). The aqueous layer was extracted by ethyl acetate (30 mL ×2). The combined organic layers were dried over anhydrous sodium sulfate, concentrated and purified by silica gel column chromatography (25% ethyl acetate/hexanes) to provide **2** as a yellow solid (3.40 g, 76% yield). ^{1}H NMR (400 MHz, $CDCl_3$): δ 3.64 (t, J = 6.4 Hz, 2H), 3.40 (t, J = 6.8 Hz, 2H), 1.86-1.82 (m, 2H), 1.57-1.54 (m, 2H), 1.44-1.30 (m, 10H).

9-(*tert*-Butyldimethylsiloxy)nonanyl bromide (3).To a solution of **2** (3.40 g, 15.2 mmol) in dichloromethane (60 mL) were added imidazole (2.68 g, 39.5 mmol), 4-dimethylaminopyridine (185 mg, 1.5 mmol) and *tert*-butyldimethylsilyl chloride (2.98 g, 19.8 mmol) at 23 °C. After stirring overnight, the solution was washed with brine (20 mL), dried over anhydrous sodium

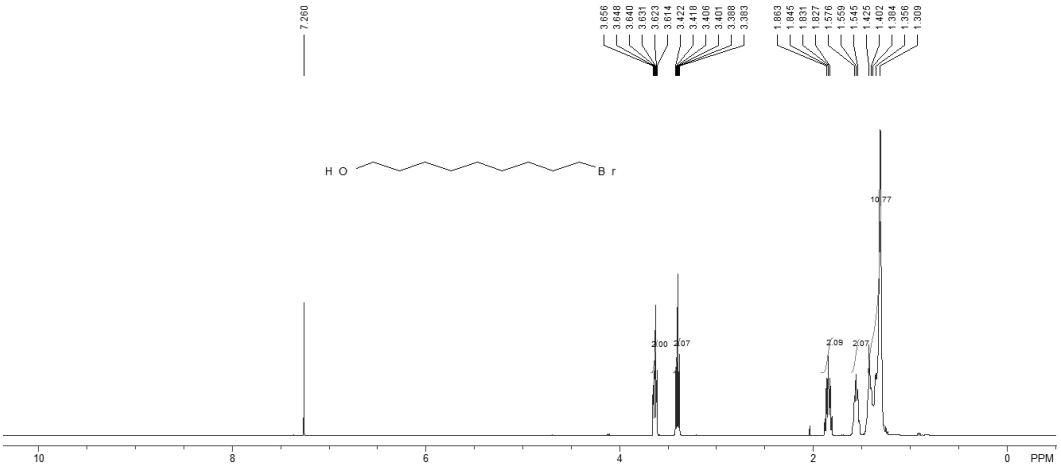
sulfate, filtered, concentrated and purified by silica gel column chromatography (25% ethyl acetate/hexanes) to provide **3** as a colorless oil (5.00 g, 98% yield). ^{1}H NMR (400 MHz, $CDCl_3$): δ 3.59 (t, *J* = 6.4 Hz, 2H), 3.40 (t, *J* = 6.8 Hz, 2H), 1.86-1.82 (m, 2H), 1.54-1.29 (m,12H), 0.89 (s, 9H), 0.04 (s, 6H).

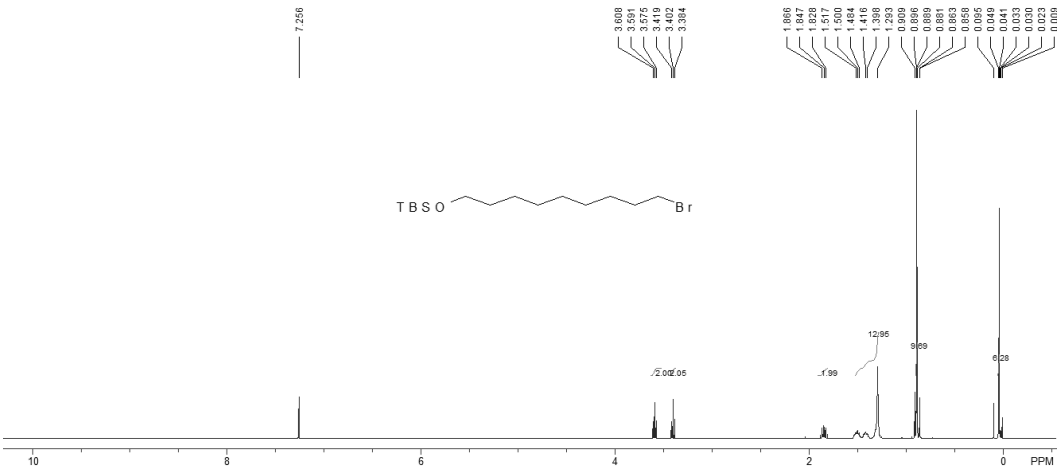
9-(*tert*-Butyldimethylsiloxy)nonanyltriphenylphosphonium bromide (4).To a solution of **3** (3.0 g, 8.89 mmol) in acetonitrile (40 mL) was added triphenylphosphine (2.56 g, 9.78 mmol) and then heated to reflux. After stirring for two days, the solution was cooled to 23 °C and concentrated to give crude **4**, which was used directly for the next step without purification.

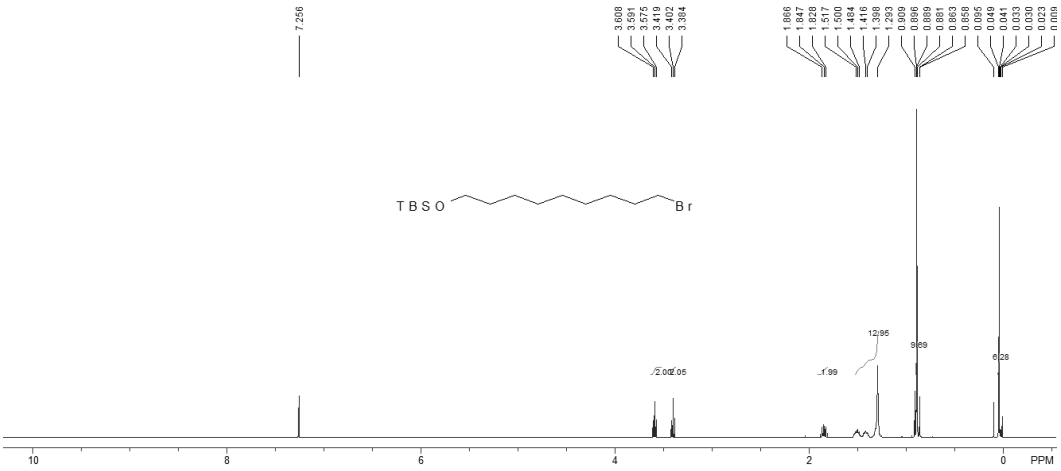
(*Z*)-*tert*-Butyl(hexadec-9-en-15-ynyloxy)dimethylsilane (6). To a solution of **4** (1.3 g, 2.16 mmol) in tetrahydrofuran (15 mL) was added sodium bis(trimethylsilyl)amide (1.0 M in tetrahydrofurn, 2.0 mL) at 0 °C followed by aldehyde **5** (183 mg, 1.67 mmol) in tetrahydrofuran (5 mL) dropwise at –78 °C. After stirring at 23 °C overnight, the reaction was quenched with saturated ammonium chloride solution and extracted with ethyl acetate (10 mL ×3). The combined organic layers were dried over anhydrous sodium sulfate, filtered and concentrated to give crude **6**, which was used directly for the next step without purification.

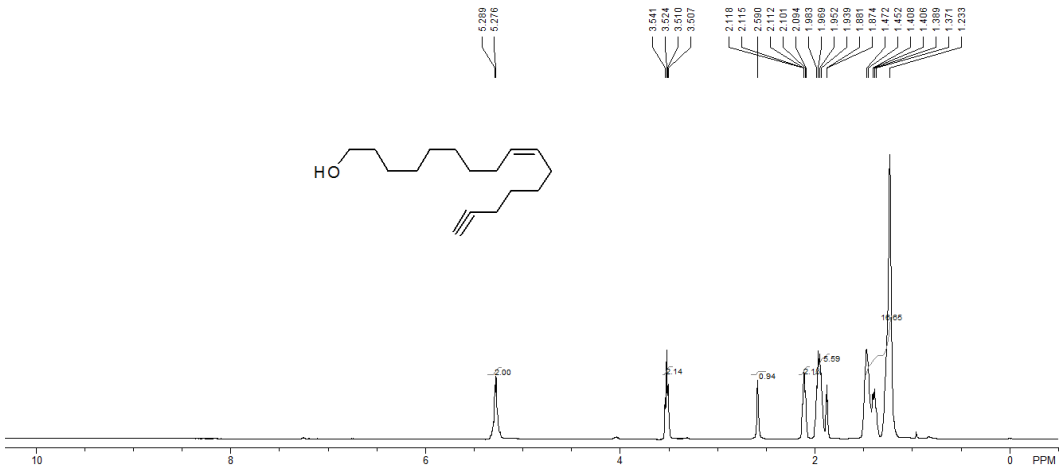
(*Z*)-Hexadec-9-en-15-yn-1-ol (7). To a solution of crude **6** in tetrahydrofuran (10 mL) was added tetra-*n*-butylammonium fluoride (1.0 M in tetrahydrofuran, 4.0 mL) at 23 °C. After stirring for 4 h, the reaction was quenched with water and extracted with ethyl acetate (10 mL ×3). The combined organic layers were dried over anhydrous sodium sulfate, filtered, concentrated and purified by silica gel column chromatography (15% ethyl acetate/hexanes) to provide **7** as a colorless oil (200 mg, 50% yield by two steps). ^{1}H NMR (400 MHz, $CDCl_3$): δ 5.29-5.26 (m, 2H), 3.52 (t, *J* = 6.8 Hz, 2H), 2.59 (s, 1H), 2.11-1.87 (m,7H), 1.47-1.23 (m, 16 H).

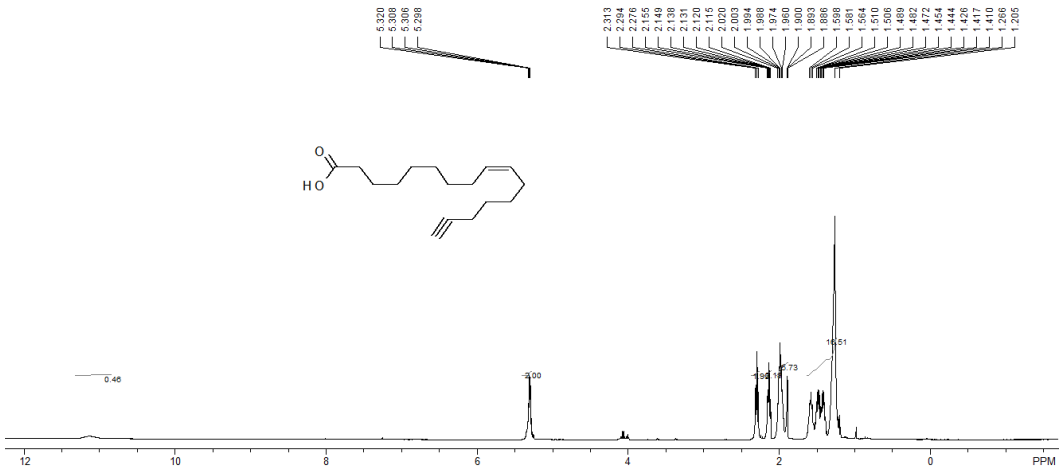
(*Z*)-Hexadec-9-en-15-ynoic acid (8). To a solution of **7** (200 mg, 0.846 mmol) in acetone (3 mL) was added chromium(VI) oxide (210 mg, 2.1 mmol) in 20 % aqueous sulfuric acid (0.7 mL) dropwise at 0 °C. After stirring for 30 min, the reaction mixture was diluted with water (10 mL) and extracted by ethyl acetate (10 mL ×3). The combined organic layers were dried over anhydrous sodium sulfate, filtered, concentrated and purified by silica gel column chromatography (25% ethyl acetate/hexanes) to provide **8** as a colorless oil (130 mg, 62% yield). ^{1}H NMR (400 MHz, $CDCl_3$): δ 11.1 (brs, 1 H), 5.32-5.28 (m, 2 H), 2.29 (t, *J* = 7.2 Hz, 2H), 2.13 (dt, *J* = 7.2, 2.4 Hz, 2H), 2.02-1.88 (m, 5H),1.15-1.26 (m, 14 H).


7.260
3.656
3.648
3.640
3.631
3.623
3.614
3.423
3.418
3.406
3.401
3.388
3.383
1.863
1.845
1.831
1.827
1.576
1.559
1.545
1.425
1.402
1.384
1.366
1.309
HO
Br
10.77
2.00
2.07
2.09
2.07
10
8
6
4
2
0
PPM

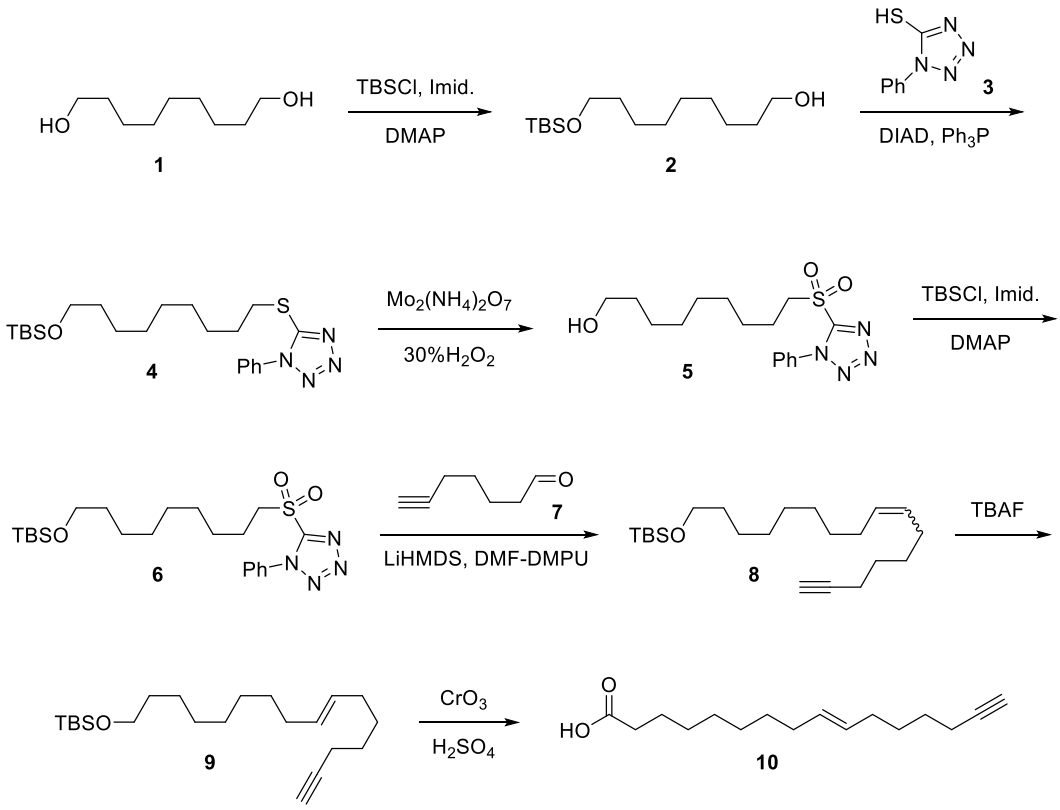

7.256
3.608
3.591
3.575
3.419
3.402
3.384
1.866
1.847
1.828
1.517
1.500
1.484
1.416
1.398
1.293
0.909
0.896
0.889
0.881
0.863
0.858
0.095
0.049
0.041
0.033
0.030
0.023
0.009
TBSO
Br
12.95
9.69
6.28
2.00
2.05
1.99
10
8
6
4
2
0
PPM


7.256
3.608
3.591
3.575
3.419
3.402
3.384
1.866
1.847
1.828
1.517
1.500
1.484
1.416
1.398
1.293
0.909
0.896
0.889
0.881
0.863
0.858
0.095
0.049
0.041
0.033
0.030
0.023
0.009
TBSO
Br
12.95
9.69
6.28
2.00
2.05
1.99
10
8
6
4
2
0
PPM


5.289
5.276
3.541
3.524
3.510
3.507
2.118
2.115
2.590
2.112
2.101
2.084
2.063
1.969
1.952
1.939
1.881
1.874
1.472
1.452
1.408
1.406
1.389
1.371
1.233
HO
2.00
2.14
0.94
2.16
5.59
16.65
10
8
6
4
2
0
PPM


5.320
5.308
5.306
5.298
2.313
2.294
2.276
2.155
2.149
2.138
2.131
2.120
2.116
2.020
2.003
1.994
1.988
1.974
1.960
1.900
1.893
1.886
1.588
1.581
1.564
1.510
1.506
1.489
1.482
1.472
1.454
1.444
1.426
1.417
1.410
1.266
1.205
O
HO
0.46
2.00
3.73
16.51
12
10
8
6
4
2
0
PPM

***Trans* ω-alkynyl palmitoleic acid synthesis:**



9-(*tert*-Butyldimethylsilyloxy)nonan-1-ol (2). To a solution of nonane-1,9-diol **1** (3.2 g, 20 mmol) in tetrahydrofuran (200 mL) was added imidazole (2.72 g, 40 mmol) and *tert*-butyldimethylsilyl chloride (3.01 g, 20 mmol) at 23 °C. After stirring overnight, the solvent was removed and the residue was diluted with ethyl acetate (50 mL), washed with saturated sodium chloride. The organic layer was dried over anhydrous sodium sulfate, concentrated and purified by silica gel column chromatography (15% ethyl acetate/hexanes) to provide **2** as an oil (2.9 g, 53% yield).

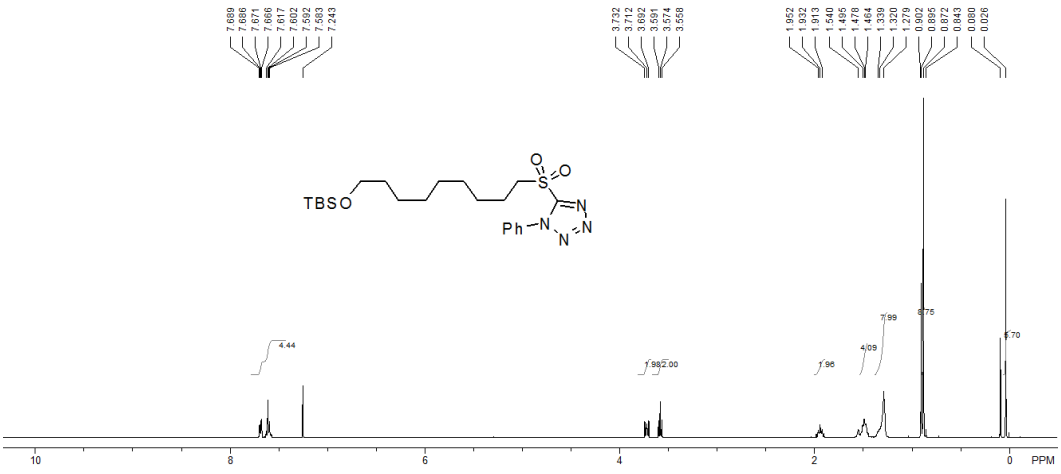
9-(1-Phenyl-1*H*-tetrazol-5-ylsulfonyl)nonan-1-ol (5). To a solution of **2** (2.0 g, 7.28 mmol) in tetrahydrofuran (80 mL) was added 1-phenyl-1*H*-tetrazole-5-thiol (1.94 g, 10.93 mmol), triphenylphosphine (2.86 g, 10.93 mmol) and diisopropyl azodicarboxylate (2.14 mL, 10.93 mmol) at 23 °C. After stirring for 2 h, the reaction mixture was diluted with ethanol (60 mL) and cooled to 0 °C. To this solution was added a bright yellow solution of a mixture of 30% aqueous hydrogen peroxide (12 mL, 109.3 mmol) and ammonium molybdate (1.61 g, 1.09 mmol) dropwise. After stirring at 23 °C overnight, the reaction was quenched with water and extracted with dichloromethane (80 mL ×3). The combined organic layers were washed with saturated sodium bisulfate, brine, dried over anhydrous sodium sulfate, filtered and concentrated. The crude product was used directly in the next step without further purification.

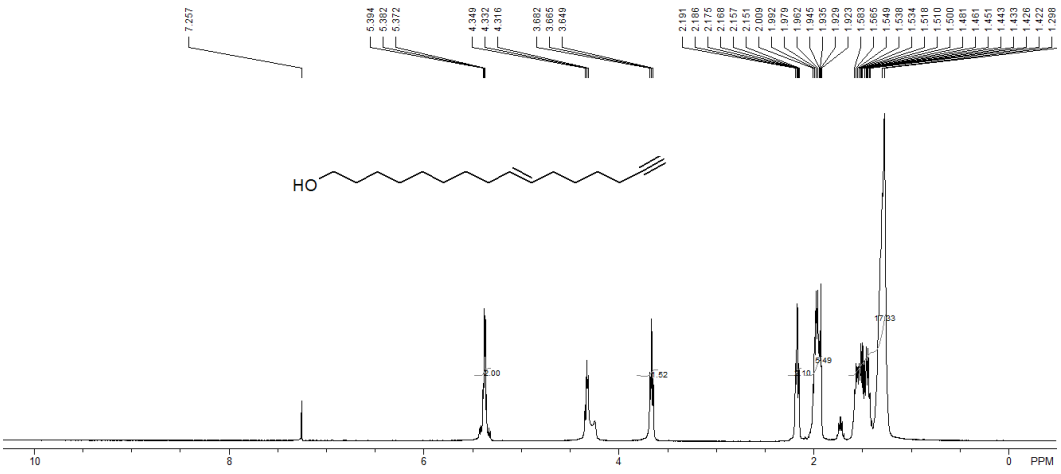
5-(9-(*tert*-Butyldimethylsilyloxy)nonylsulfonyl)-1-phenyl-1*H*-tetrazole (6). To a solution of **5** in dichloromethane (80 mL) was added imidazole (1.48 g, 21.84 mmol), *tert*-butyldimethylsilyl chloride (1.64 g, 10.9 mmol) and 4-dimethylaminopyridine (89 mg, 0.72 mmol) at 23 °C. After stirring overnight, the reaction was quenched with saturated sodium chloride and extracted with dichloromethane (50 mL ×3). The combined organic layers were dried over anhydrous sodium sulfate, concentrated and purified by silica gel column chromatography (20% ethyl acetate/hexanes) to provide **6** as an oil (2.3 g, 67% yield over three steps). ^{1}H NMR (400 MHz, $CDCl_3$): δ 7.69-7.58 (m, 5H), 3.71 (t, J = 8.0 Hz, 2H), 3.57 (t, J = 6.8 Hz, 2H), 1.95-1.91 (m, 2H), 1.49-1.46 (m, 4H), 1.34-1.27 (m, 8H), 0.87 (s, 9H), 0.02 (s, 6H).

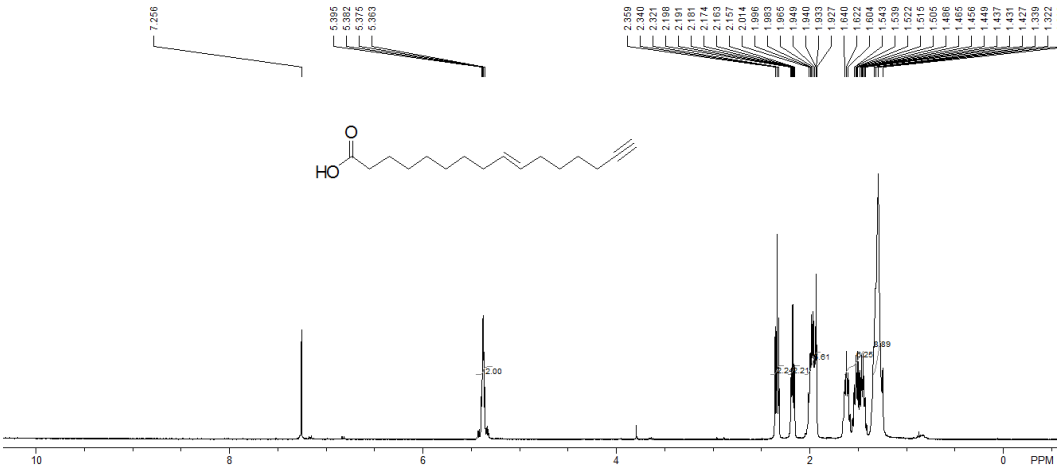
***tert*-Butyl(hexadec-9-en-15-ynyloxy)dimethylsilane (8).** To a soluion of **6** (953 mg, 2.04 mmol) in *N*,*N*-dimethylformamide (7.5 mL) and *N*,*N*′-dimethylpropylene urea (2.5 mL) was added lithium bis(trimethylsilyl)amide (1.0 M in tetrahydrofuran, 2.0 mL) slowly at –78°C. After stirring for 30 min, aldehyde **7** (150 mg, 1.36 mmol) in *N*,*N*-dimethylformamide (3 mL) and *N*,*N*′-dimethylpropylene urea (1 mL) was added dropwise. After stirring at 23 °C overnight, the reaction was quenched with saturated ammonium chloride solution and extracted with ethyl acetate (10 mL ×3).The organic layers were washed with brine (10 mL ×2), dried over anhydrous sodium sulfate, filtered and concentrated. The crude product (*Z*/*E* = 1:1) was used directly for the next step without purification.

(*E*)-Hexadec-9-en-15-yn-1-ol (9). To a solution of crude **8** in tetrahydrofuran (5 mL) was added tetra-*n*-butylammonium fluoride (1.0 M in tetrahydrofuran, 1.7 mL) at 23 °C. After stirring for 4 h, the reaction was quenched with water and extracted with ethyl acetate (10 mL ×3). The combined organic layers were dried over anhydrous sodium sulfate, filtered, concentrated and purified by prepared HPLC to provide **9** as a colorless oil (64 mg, 20% yield). ^{1}H NMR (400 MHz, $CDCl_3$): δ 5.39-5.36 (m, 2H), 3.66 (t, J = 6.8 Hz, 2H), 2.19-1.92 (m, 7H), 1.47-1.23 (m, 17 H).

(*E*)-Hexadec-9-en-15-ynoic acid (10). To a solution of **9** (30 mg, 0.127 mmol) in acetone (1 mL) was added chromium(VI) oxide (32 mg, 0.315 mmol) in 20% aqueous sulfuric acid (0.15 mL) dropwise at 0 °C. After stirring at 0°C for 30 min, the reaction mixture was diluted with water (3 mL) and extracted by ethyl acetate (5 mL ×3). The combined organic layers were dried over anhydrous sodium sulfate, filtered, concentrated, and purified by silica gel column chromatography(25% ethyl acetate/hexanes) to provide **10** as a colorless oil (13 mg, 43% yield). ^{1}H NMR (400 MHz, $CDCl_3$): δ 5.39-5.36 (m, 2 H), 2.34 (t, J = 7.6 Hz, 2H), 2.17 (dt, J = 6.8, 2.8 Hz, 2H), 2.01-1.92 (m, 5H), 1.64-1.24 (m, 14 H).


7.689
7.686
7.671
7.666
7.617
7.602
7.582
7.563
7.243
3.732
3.712
3.692
3.591
3.574
3.558
1.952
1.932
1.913
1.540
1.495
1.478
1.464
1.339
1.320
1.279
1.093
0.895
0.872
0.843
0.080
0.026
TBSO
O
S
O
N
N
N
Ph
4.44
1.98
2.00
1.96
4.09
7.99
8.75
6.70
10
8
6
4
2
0
PPM


7.257
5.394
5.382
5.372
4.349
4.332
4.316
3.682
3.665
3.649
2.191
2.186
2.176
2.168
2.157
2.151
2.009
1.992
1.979
1.962
1.945
1.935
1.929
1.923
1.583
1.565
1.549
1.538
1.534
1.518
1.510
1.500
1.481
1.461
1.455
1.443
1.433
1.426
1.422
1.298
HO
2.00
1.52
5.10
5.49
17.33
10
8
6
4
2
0
PPM


7.256
5.395
5.382
5.375
5.363
2.359
2.340
2.321
2.198
2.191
2.181
2.174
2.163
2.157
2.014
1.996
1.983
1.965
1.949
1.940
1.933
1.927
1.640
1.622
1.604
1.543
1.539
1.522
1.515
1.496
1.486
1.465
1.456
1.449
1.437
1.431
1.427
1.339
1.322
O
HO
2.00
2.61
2.25
2.21
8.89
10
8
6
4
2
0
PPM

sgRNAs

Single guide RNAs (sgRNAs) targeting *LKB1* (exon 1) and Sufu (exons 2, 3 and 8) (listed in Table 1) cloned in pCas-Guide plasmid were purchased from Blue Heron Biotech. In order to screen for the prevalence of CRISPR-mediated exon skipping events, GeckoV2 library (1 vector lentiCRISPRv2; Addgene) was pooled together and electroporated into electrocompetent cells (Lucigen 10F elite cells) with an efficiency of > 2 x 10^{10} and individual sgRNAs were randomly selected and sequenced. sgRNAs were also designed to target exons 2 and 8 of Sufu and exons 39, 51 and 57 of DMD using DNA2.0 CRISPR gRNA Design tool (www.dna20.com) and cloned into lentiCRISPRv2 plasmid as described in Ran et al. [198] CRISPR RNA (crRNA) targeting exon 5 of LKB1 gene was purchased from Dharmacon.

Western blot analysis

Cell lysates were generated with PBS/1% NP40 buffer supplemented with protease inhibitor cocktail (Sigma).

Luciferase reporter assay

Reporter assay was executed as described using a Dual Luciferase kit (Promega). SuperTopFlash and control SV40 driven Renilla luciferase reporters used in the reporter assay were previously described [71] [199].

Click chemistry assay

HEK293 cells were transiently transfected with either the IgG-Fc, the Wnt3a-Fc or the ShhN-Fc DNA expression constructs and treated with various ω-alkynyl fatty acid probes (see “Reagents”, 100 mM final concentration) for 6 hours. Wnt-Fc and ShhN-Fc proteins modified with the

synthetic alkynylated fatty acid probes were then purified from the lysate using Protein A sepharose and subjected to copper catalized alkyne-azide cycloaddition reaction in the presence of biotin-azide. The biotinylated Wnt3a-Fc and ShhN-Fc proteins were run on a SDS-PAGE gel and detected using Streptavidin-HRP. The expression of total IgG-Fc, Wnt3a-Fc and ShhN-Fc proteins was detected using Fc-HRP.

Pulse-chase assay to monitor secretion of acylated Wnt proteins

HEK293 cells transiently transfected with Wnt3a cDNA were labeled with either *cis* or *trans* alkynylated palmitoleic acid for 6 hours. The cells were thoroughly washed with PBS to remove any remaining fatty acid isomers after the end of the labeling period and replenished with fresh culture medium. The Wnt proteins labeled with alkynylated probes were collected from the medium and the lysate after 24 hours and immobilized on Con A-Sepharose and Protein A-Sepharose using Wnt3a antibody respectively. Click chemistry reaction was performed on the sepharose-bound Wnt3a proteins and the palmitoleated Wnt proteins were detected using Streptavidin-HRP. The total Wnt3a proteins isolated from either of the sources were detected using Wnt3a antibody. The value for total acylated Wnts in lysate was used for normalizing the Wnt secretion value derived from the medium.

Exosome precipitation assay

HEK293 cells stably expressing either the control or the Wnt3a cDNAs were treated with DMSO or LGK974 for 48 hours. Total secreted Wnt3a proteins from the culture medium were immobilized on Con A-Sepharose beads, whereas the Wnt3a proteins secreted in exosomes were precipitated using ExoQuick-TC kit using the manufacturer's instructions (System Biosciences). In order to determine the population of Wnt3a proteins secreted in exosome-independent

manner, Con A-sepharose beads were used to enrich the Wnt3a proteins from the exosome depleted culture medium.

Transfection of sgRNAs

1 x 10^6 Mia pancreatic cells were seeded per 6 well and co-transfected with 0.35 ug sgRNA pCas-Guide plasmid and 0.15 ug puromycin expression plasmid using effectene transfection reagent (Qiagen). 24 hours after transfection, cells were trypisinized and plated in 150 mm culture dishes in various dilutions for clonal selection.

Clonal isolation of CRISPR-edited cells

Cells in 150 mm culture plates were treated with 0.5 ug/ml of puromycin in order to select for the cells with expression of pCas-Guide and puromycin plasmid. Puromycin selection was maintained for 10 days after which single colonies were isolated and grown in a 96 well plate. Cells from single colonies were passaged multiple times until sufficient cells were available for analyzing genomic DNA, RNA and protein.

Genomic DNA extraction and genomic sequencing

Genomic DNA was extracted from the CRISPR-edited cells using Genomic DNA Mini kit (Bioland Scientific) according to manufacturer's instructions and used as template for PCR amplification. PCR primers encompassing the CRIPSR-targeted region were designed (and is listed in the Table 2). PCR was performed with GoTaq Green Master Mix (Promega M7122) with the PCR following condition: 98^o C for 2 mins (Initial denaturation), 25 cycles of 98^o C for 30 secs, 56^o C for 30 secs, 72^o C for 30 secs (denaturation, annealing, extension) and final 70^o C for 5 minutes (final extension). Gel electrophoresis in a 1% agarose gel was performed and the

PCR products were purified from gel using QIA Quick PCR Purification Kit (Qiagen) and cloned into pCR-TOPO plasmid using TOPO TA cloning kit for Subcloning (ThermoFisher Scientific). pCR-TOPO plasmids containing genomic DNA sequences were transformed into TOP10 competent cells and individual colonies were selected and sequenced.

RNA extraction and analysis

RNA extraction was performed using RNeasy Mini Kit (Qiagen) according to manufacturer's instructions. cDNA synthesis was performed on 1 ug of RNA using ProtoScript First Strand cDNA Synthesis Kit (Promega). Primers recognizing distal exons within the targeted gene were designed (and is listed in Table 3) and used to amplify the cDNA sequences isolated from the CRISPR-edited cells. PCR products were electrophoresed in 1% agarose gel and the gel bands were isolated using QIA Quick PCR Purification Kit (Qiagen), cloned into pCR-TOPO plasmids using TOPO TA cloning kit for subcloning (ThermoFisher Scientific) and sequenced at UTSW Sequencing Core.

Mini-screen for exon skipping

sgRNAs randomly selected from the GeckoV2 library together with packaging plasmids VSVg (AddGene 8454) and psPAX2 (AddGene 12260) were co-transfected into HEK293FT cells in the ratio 2: 1: 1.5 using effectene transfection reagent. 12 hours post-transfection, the existing medium was replaced with fresh culture medium and 48 hours after the initial transfection, viral supernatant was collected and centrifuged for 5 minutes at 1000 x g to get rid of any cell debris. For transduction, 1 x 10^6 Mia cell seeded in 6 well plate was treated with viral supernatant for 48 hours. After 48 hours, viral medium was replaced with fresh medium containing 0.5 ug/ml of

puromycin to select for the transduced cells. 7-10 days after puromycin selection, the cells were split into multiple plates for generation of genomic DNA, cDNA and protein.

Humanized *Shistosoma* Porcn-myc sequence:

GCGGCCGCATGGACGAGCAGGAGGAGTCTTGGTACATGGACGAGGACGAGGAGGA
GCTGTTCGAGGAAGTGGACGAGGAGAACGGCGATAGCGATCTGAGCGGCAACTACC
TGGACGACCTGGAAACCGTGGAGGAAGACGACGGCGATTTCGAGGACGTGTCCTTC
GAGAACGCCCTGAGCATGAAGAGCATCAGCAGCGAGATCCTGTCTTGCAGCATCCC
CGTGCTGAAGCAGGCCTATTGGCTGCTGATCCTGATCCTGGTGTCTTGCCTCATCTGG
AGGGTGGTGCAGTTCCTGCTGAATCGCTGCTTCCCCAGCAACGGCAAGATCTCTCGG
ACCTTCGTGGCCATCATCCTGCACATCTTCAGCATTTGCCTGGGCCTGCTGGTGCTGT
ACTACGCCTCCTACGACCTCTGGTGGATCGTGATCGGCCTGATCATCACCCTGACCC
TGCTGTTCACCATCAACATCAACAGCGACATCAGCAACAGCCACGGCGAGTACACC
AATTGGGAGCTGTTCGTGACCATCGGCATTTGCAGCAGCGTGCAGCTGTATTGCGAG
TTCTACAAGAACCCAGTCAAGTGGCACCAGATCAGAGGCAGCATCATGATCATCAT
CATGAAGTGCATCAGCTTCAGCATGGAGAACAAGAACTTCTACAAGACCGTGAACA
AGGAGCCTTGTTGCATCCTCCTCTGGACCCCTCTGTATAGAGGCCTGCTGTGGCTGA
GCTATTGCCTCTGTCCAGCCAGCCTGCTGTTCGGACCTTGGTTCAACCCCCTGCGCTA
CGAGGAGATGGTCCGGAACTACGCAGTGAACGACCGGCCCTTCAGCTTCAAGAACA
CCATGATCAGCCTGTTCCAGACCGCCAAGCTGTTCGGCCTGAGCTTCCTGTACCTGA
CCTACAGCACTTGCTTCACCGACACCCTGGCCAGCATCCTGAAACTGCAGCCTTGGC
TGCACGCCTACTTTGCCAGCCAGAGCTTCCGGTTCAGCCACTACTTCATCTGCATCA
GCAGCGAGAGCTCTATGACCGCCCTGGGCTATTGCGACAAGTACAGCATCCAGAGC
AACAAGCAGGAGAACACCAAGGAGAAGAACCGGAGCAAGGACGTGTACAAGCCAG
TGGTGGTGACCAGGCCTCTGTTCATCGAGTTCCCCAGGAGCCTGGTCGAGGTCGTGA
TCTATTGGAACCTGCCCATGCACACTTGGCTGAAGCAGTACGTGTACAAGCCTCTGA
GACCTTACGGCCACGTGTACGCAGTGCTGGGAACATACACAGCCAGCAGCCTGCTG
CACGGACTGAACTTCCAGCTGAGCGCAGTGCTGTTCTCCATCGGCATCTACGCCTAC
ACCGAGTACGTGTTCCGGGAGAAGCTGGCTAAGATCCTGGACGCTTGCATCGCCAG
CAGGCCTTGCCCCGAGAGTTGTTGCCACTCCAACAAGAACAGCAGCTGGGTCTACTG
GTGCAATGTGGCCTTCTCTTGCCTGGCCATCTTCCACCTGGCCTACCTGGCCGTGATG
TTCGATACCAGCGAGCAGCAGTTCCAGGGCTACAACATGTGGCACACCATGAACAA
GTGGTACAACCTGGGCTTCCTGTCTCACATCGTGGCCTTTGCCAACTTCCTGTTCTAC
CTGTACCTGGCGGCCGCTCGAGGTCACCCATTCGAACAAAAACTCATCTCAGAAGA
GGATCTGAATATGCATACCGGTCATCATCACCATCACCATTGA

BIBLIOGRAPHY

1. Clevers, H., K.M. Loh, and R. Nusse, *Stem cell signaling. An integral program for tissue renewal and regeneration: Wnt signaling and stem cell control.* Science, 2014. **346**(6205): p. 1248012.
2. Logan, C.Y. and R. Nusse, *The Wnt signaling pathway in development and disease.* Annu Rev Cell Dev Biol, 2004. **20**: p. 781-810.
3. Nusse, R. and H.E. Varmus, *Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome.* Cell, 1982. **31**(1): p. 99-109.
4. Sharma, R.P. and V.L. Chopra, *Effect of the wingless (wg1) mutation on wing and haltere development in Drosophila melanogaster.* Developmental Biology, 1976. **48**(2): p. 461-465.
5. Rijsewijk, F., et al., *The Drosophila homology of the mouse mammary oncogene int-1 is identical to the segment polarity gene wingless.* Cell, 1987. **50**(4): p. 649-657.
6. Reya, T. and H. Clevers, *Wnt signalling in stem cells and cancer.* Nature, 2005. **434**(7035): p. 843-50.
7. Willert, K. and R. Nusse, *Wnt Proteins.* Cold Spring Harbor Perspectives in Biology, 2012. **4**(9).
8. Tang, X.F., et al., *Roles of N-glycosylation and lipidation in Wg secretion and signaling.* Developmental Biology, 2012. **364**(1): p. 32-41.
9. Herr, P. and K. Basler, *Porcupine-mediated lipidation is required for Wnt recognition by Wls.* Developmental Biology, 2012. **361**(2): p. 392-402.
10. Hofmann, K., *A superfamily of membrane-bound O-acyltransferases with implications for Wnt signaling.* Trends in Biochemical Sciences, 2000. **25**(3): p. 111-112.
11. Yang, J., et al., *Identification of the acyltransferase that octanoylates ghrelin, an appetite-stimulating peptide hormone.* Cell, 2008. **132**(3): p. 387-396.
12. Chamoun, Z., et al., *Skinny Hedgehog, an acyltransferase required for palmitoylation and activity of the Hedgehog signal.* Science, 2001. **293**(5537): p. 2080-2084.
13. Willert, K., et al., *Wnt proteins are lipid-modified and can act as stem cell growth factors.* Nature, 2003. **423**(6938): p. 448-52.
14. Takada, R., et al., *Monounsaturated fatty acid modification of Wnt protein: its role in Wnt secretion.* Dev Cell, 2006. **11**(6): p. 791-801.
15. Janda, C.Y., et al., *Structural basis of Wnt recognition by Frizzled.* Science, 2012. **337**(6090): p. 59-64.
16. Banziger, C., et al., *Wntless, a conserved membrane protein dedicated to the secretion of Wnt proteins from signaling cells.* Cell, 2006. **125**(3): p. 509-22.
17. Bartscherer, K., et al., *Secretion of Wnt ligands requires Evi, a conserved transmembrane protein.* Cell, 2006. **125**(3): p. 523-33.
18. Goodman, R.M., et al., *Sprinter: a novel transmembrane protein required for Wg secretion and signaling.* Development, 2006. **133**(24): p. 4901-11.
19. Herr, P. and K. Basler, *Porcupine-mediated lipidation is required for Wnt recognition by Wls.* Dev Biol, 2012. **361**(2): p. 392-402.

20. Coombs, G.S., et al., *WLS-dependent secretion of WNT3A requires Ser209 acylation and vacuolar acidification.* J Cell Sci, 2010. **123**(Pt 19): p. 3357-67.
21. Port, F., G. Hausmann, and K. Basler, *A genome-wide RNA interference screen uncovers two p24 proteins as regulators of Wingless secretion.* EMBO Rep, 2011. **12**(11): p. 1144-52.
22. Buechling, T., et al., *p24 proteins are required for secretion of Wnt ligands.* EMBO Rep, 2011. **12**(12): p. 1265-72.
23. Mulligan, K.A., et al., *Secreted Wingless-interacting molecule (Swim) promotes long-range signaling by maintaining Wingless solubility.* Proc Natl Acad Sci U S A, 2012. **109**(2): p. 370-7.
24. Mihara, E., et al., *Active and water-soluble form of lipidated Wnt protein is maintained by a serum glycoprotein afamin/alpha-albumin.* Elife, 2016. **5**.
25. Gross, J.C., et al., *Active Wnt proteins are secreted on exosomes.* Nat Cell Biol, 2012. **14**(10): p. 1036-45.
26. Koles, K. and V. Budnik, *Exosomes go with the Wnt.* Cell Logist, 2012. **2**(3): p. 169-173.
27. Zhang, L. and J.L. Wrana, *The emerging role of exosomes in Wnt secretion and transport.* Curr Opin Genet Dev, 2014. **27**: p. 14-9.
28. Belenkaya, T.Y., et al., *The retromer complex influences Wnt secretion by recycling wntless from endosomes to the trans-Golgi network.* Dev Cell, 2008. **14**(1): p. 120-31.
29. Port, F., et al., *Wingless secretion promotes and requires retromer-dependent cycling of Wntless.* Nat Cell Biol, 2008. **10**(2): p. 178-85.
30. Franch-Marro, X., et al., *Wingless secretion requires endosome-to-Golgi retrieval of Wntless/Evi/Sprinter by the retromer complex.* Nat Cell Biol, 2008. **10**(2): p. 170-7.
31. Yu, J., et al., *WLS retrograde transport to the endoplasmic reticulum during Wnt secretion.* Dev Cell, 2014. **29**(3): p. 277-91.
32. Gasnereau, I., et al., *Identification of an endocytosis motif in an intracellular loop of Wntless protein, essential for its recycling and the control of Wnt protein signaling.* J Biol Chem, 2011. **286**(50): p. 43324-33.
33. Amit, S., et al., *Axin-mediated CKI phosphorylation of beta-catenin at Ser 45: a molecular switch for the Wnt pathway.* Genes Dev, 2002. **16**(9): p. 1066-76.
34. Behrens, J., et al., *Functional interaction of an axin homolog, conductin, with beta-catenin, APC, and GSK3beta.* Science, 1998. **280**(5363): p. 596-9.
35. Liu, C., et al., *Control of beta-catenin phosphorylation/degradation by a dual-kinase mechanism.* Cell, 2002. **108**(6): p. 837-47.
36. Cavallo, R.A., et al., *Drosophila Tcf and Groucho interact to repress Wingless signalling activity.* Nature, 1998. **395**(6702): p. 604-8.
37. Roose, J., et al., *The Xenopus Wnt effector XTcf-3 interacts with Groucho-related transcriptional repressors.* Nature, 1998. **395**(6702): p. 608-12.
38. Chen, C.M. and G. Struhl, *Wingless transduction by the Frizzled and Frizzled2 proteins of Drosophila.* Development, 1999. **126**(23): p. 5441-52.
39. Bhanot, P., et al., *A new member of the frizzled family from Drosophila functions as a Wingless receptor.* Nature, 1996. **382**(6588): p. 225-30.

40. Wehrli, M., et al., *arrow encodes an LDL-receptor-related protein essential for Wingless signalling.* Nature, 2000. **407**(6803): p. 527-30.
41. Cliffe, A., F. Hamada, and M. Bienz, *A role of Dishevelled in relocating Axin to the plasma membrane during wingless signaling.* Curr Biol, 2003. **13**(11): p. 960-6.
42. Mao, J., et al., *Low-density lipoprotein receptor-related protein-5 binds to Axin and regulates the canonical Wnt signaling pathway.* Mol Cell, 2001. **7**(4): p. 801-9.
43. Schwarz-Romond, T., et al., *The DIX domain of Dishevelled confers Wnt signaling by dynamic polymerization.* Nat Struct Mol Biol, 2007. **14**(6): p. 484-92.
44. Zeng, X., et al., *Initiation of Wnt signaling: control of Wnt coreceptor Lrp6 phosphorylation/activation via frizzled, dishevelled and axin functions.* Development, 2008. **135**(2): p. 367-75.
45. Riese, J., et al., *LEF-1, a nuclear factor coordinating signaling inputs from wingless and decapentaplegic.* Cell, 1997. **88**(6): p. 777-87.
46. Behrens, J., et al., *Functional interaction of beta-catenin with the transcription factor LEF-1.* Nature, 1996. **382**(6592): p. 638-42.
47. van de Wetering, M., et al., *Armadillo coactivates transcription driven by the product of the Drosophila segment polarity gene dTCF.* Cell, 1997. **88**(6): p. 789-99.
48. Baker, N.E., *Molecular cloning of sequences from wingless, a segment polarity gene in Drosophila: the spatial distribution of a transcript in embryos.* Embo j, 1987. **6**(6): p. 1765-73.
49. van den Heuvel, M., et al., *Mutations in the segment polarity genes wingless and porcupine impair secretion of the wingless protein.* Embo j, 1993. **12**(13): p. 5293-302.
50. McMahon, A.P. and R.T. Moon, *Ectopic expression of the proto-oncogene int-1 in Xenopus embryos leads to duplication of the embryonic axis.* Cell, 1989. **58**(6): p. 1075-84.
51. van Amerongen, R. and R. Nusse, *Towards an integrated view of Wnt signaling in development.* Development, 2009. **136**(19): p. 3205-14.
52. Millar, S.E., et al., *WNT signaling in the control of hair growth and structure.* Dev Biol, 1999. **207**(1): p. 133-49.
53. Alonso, L. and E. Fuchs, *Stem cells in the skin: waste not, Wnt not.* Genes Dev, 2003. **17**(10): p. 1189-200.
54. Pinto, D. and H. Clevers, *Wnt control of stem cells and differentiation in the intestinal epithelium.* Exp Cell Res, 2005. **306**(2): p. 357-63.
55. Nemeth, M.J., et al., *beta-Catenin expression in the bone marrow microenvironment is required for long-term maintenance of primitive hematopoietic cells.* Stem Cells, 2009. **27**(5): p. 1109-19.
56. Malhotra, S. and P.W. Kincade, *Wnt-related molecules and signaling pathway equilibrium in hematopoiesis.* Cell Stem Cell, 2009. **4**(1): p. 27-36.
57. Rubinfeld, B., et al., *Association of the APC gene product with beta-catenin.* Science, 1993. **262**(5140): p. 1731-4.
58. Su, L.K., B. Vogelstein, and K.W. Kinzler, *Association of the APC tumor suppressor protein with catenins.* Science, 1993. **262**(5140): p. 1734-7.

59. Moser, A.R., H.C. Pitot, and W.F. Dove, *A dominant mutation that predisposes to multiple intestinal neoplasia in the mouse.* Science, 1990. **247**(4940): p. 322-4.
60. Vogelstein, B., et al., *Cancer genome landscapes.* Science, 2013. **339**(6127): p. 1546-58.
61. Klaus, A. and W. Birchmeier, *Wnt signalling and its impact on development and cancer.* Nat Rev Cancer, 2008. **8**(5): p. 387-98.
62. Cancer Genome Atlas, N., *Comprehensive molecular characterization of human colon and rectal cancer.* Nature, 2012. **487**(7407): p. 330-7.
63. Polakis, P., *The oncogenic activation of beta-catenin.* Curr Opin Genet Dev, 1999. **9**(1): p. 15-21.
64. Major, M.B., et al., *Wilms tumor suppressor WTX negatively regulates WNT/beta-catenin signaling.* Science, 2007. **316**(5827): p. 1043-6.
65. Moon, R.T., et al., *WNT and beta-catenin signalling: diseases and therapies.* Nat Rev Genet, 2004. **5**(9): p. 691-701.
66. Wu, J., et al., *Whole-exome sequencing of neoplastic cysts of the pancreas reveals recurrent mutations in components of ubiquitin-dependent pathways.* Proc Natl Acad Sci U S A, 2011. **108**(52): p. 21188-93.
67. Hao, H.X., et al., *ZNRF3 promotes Wnt receptor turnover in an R-spondin-sensitive manner.* Nature, 2012. **485**(7397): p. 195-200.
68. Lazo, J.S. and E.R. Sharlow, *Drugging Undruggable Molecular Cancer Targets.* Annu Rev Pharmacol Toxicol, 2016. **56**: p. 23-40.
69. Bergmann, C. and J.H. Distler, *Canonical Wnt signaling in systemic sclerosis.* Lab Invest, 2016. **96**(2): p. 151-5.
70. Holstein, T.W., *The evolution of the Wnt pathway.* Cold Spring Harb Perspect Biol, 2012. **4**(7): p. a007922.
71. Chen, B., et al., *Small molecule-mediated disruption of Wnt-dependent signaling in tissue regeneration and cancer.* Nat Chem Biol, 2009. **5**(2): p. 100-7.
72. Jiang, X., et al., *Inactivating mutations of RNF43 confer Wnt dependency in pancreatic ductal adenocarcinoma.* Proc Natl Acad Sci U S A, 2013. **110**(31): p. 12649-54.
73. Bovolenta, P., et al., *Beyond Wnt inhibition: new functions of secreted Frizzled-related proteins in development and disease.* J Cell Sci, 2008. **121**(Pt 6): p. 737-46.
74. Malinauskas, T., et al., *Modular mechanism of Wnt signaling inhibition by Wnt inhibitory factor 1.* Nat Struct Mol Biol, 2011. **18**(8): p. 886-93.
75. Gurney, A., et al., *Wnt pathway inhibition via the targeting of Frizzled receptors results in decreased growth and tumorigenicity of human tumors.* Proc Natl Acad Sci U S A, 2012. **109**(29): p. 11717-22.
76. Huang, S.M., et al., *Tankyrase inhibition stabilizes axin and antagonizes Wnt signalling.* Nature, 2009. **461**(7264): p. 614-20.
77. Fiedler, M., et al., *Dishevelled interacts with the DIX domain polymerization interface of Axin to interfere with its function in down-regulating beta-catenin.* Proc Natl Acad Sci U S A, 2011. **108**(5): p. 1937-42.

78. Wong, H.C., et al., *Direct binding of the PDZ domain of Dishevelled to a conserved internal sequence in the C-terminal region of Frizzled.* Mol Cell, 2003. **12**(5): p. 1251-60.
79. Shan, J., et al., *Identification of a specific inhibitor of the dishevelled PDZ domain.* Biochemistry, 2005. **44**(47): p. 15495-503.
80. Fujii, N., et al., *An antagonist of dishevelled protein-protein interaction suppresses beta-catenin-dependent tumor cell growth.* Cancer Res, 2007. **67**(2): p. 573-9.
81. Grandy, D., et al., *Discovery and characterization of a small molecule inhibitor of the PDZ domain of dishevelled.* J Biol Chem, 2009. **284**(24): p. 16256-63.
82. Rubinfeld, B., et al., *Binding of GSK3beta to the APC-beta-catenin complex and regulation of complex assembly.* Science, 1996. **272**(5264): p. 1023-6.
83. Thorne, C.A., et al., *Small-molecule inhibition of Wnt signaling through activation of casein kinase 1alpha.* Nat Chem Biol, 2010. **6**(11): p. 829-36.
84. Saraswati, S., et al., *Pyrvinium, a potent small molecule Wnt inhibitor, promotes wound repair and post-MI cardiac remodeling.* PLoS One, 2010. **5**(11): p. e15521.
85. Kahn, M., *Can we safely target the WNT pathway?* Nat Rev Drug Discov, 2014. **13**(7): p. 513-32.
86. Lepourcelet, M., et al., *Small-molecule antagonists of the oncogenic Tcf/beta-catenin protein complex.* Cancer Cell, 2004. **5**(1): p. 91-102.
87. Hahne, G. and T.N. Grossmann, *Direct targeting of beta-catenin: Inhibition of protein-protein interactions for the inactivation of Wnt signaling.* Bioorg Med Chem, 2013. **21**(14): p. 4020-6.
88. Tian, W., et al., *Structure-based discovery of a novel inhibitor targeting the beta-catenin/Tcf4 interaction.* Biochemistry, 2012. **51**(2): p. 724-31.
89. Gonsalves, F.C., et al., *An RNAi-based chemical genetic screen identifies three small-molecule inhibitors of the Wnt/wingless signaling pathway.* Proc Natl Acad Sci U S A, 2011. **108**(15): p. 5954-63.
90. Emami, K.H., et al., *A small molecule inhibitor of beta-catenin/CREB-binding protein transcription [corrected].* Proc Natl Acad Sci U S A, 2004. **101**(34): p. 12682-7.
91. Haynes, C.A., et al., *Quantitation of fatty acyl-coenzyme As in mammalian cells by liquid chromatography-electrospray ionization tandem mass spectrometry.* J Lipid Res, 2008. **49**(5): p. 1113-25.
92. Paton, C.M. and J.M. Ntambi, *Biochemical and physiological function of stearoyl-CoA desaturase.* Am J Physiol Endocrinol Metab, 2009. **297**(1): p. E28-37.
93. Hodson, L. and B.A. Fielding, *Stearoyl-CoA desaturase: rogue or innocent bystander?* Prog Lipid Res, 2013. **52**(1): p. 15-42.
94. Brown, J.M. and L.L. Rudel, *Stearoyl-coenzyme A desaturase 1 inhibition and the metabolic syndrome: considerations for future drug discovery.* Curr Opin Lipidol, 2010. **21**(3): p. 192-7.
95. Rios-Esteves, J. and M.D. Resh, *Stearoyl CoA desaturase is required to produce active, lipid-modified Wnt proteins.* Cell Rep, 2013. **4**(6): p. 1072-81.

96. Thorn, K. and P. Bergsten, *Fatty acid-induced oxidation and triglyceride formation is higher in insulin-producing MIN6 cells exposed to oleate compared to palmitate.* J Cell Biochem, 2010. **111**(2): p. 497-507.
97. Mann, R.K. and P.A. Beachy, *Novel lipid modifications of secreted protein signals.* Annu Rev Biochem, 2004. **73**.
98. Pepinsky, R.B., et al., *Identification of a palmitic acid-modified form of human Sonic hedgehog.* J Biol Chem, 1998. **273**(22): p. 14037-45.
99. Burke, R., et al., *Dispatched, a novel sterol-sensing domain protein dedicated to the release of cholesterol-modified hedgehog from signaling cells.* Cell, 1999. **99**(7): p. 803-15.
100. Ma, Y., et al., *Hedgehog-mediated patterning of the mammalian embryo requires transporter-like function of dispatched.* Cell, 2002. **111**(1): p. 63-75.
101. Zeng, X., et al., *A freely diffusible form of Sonic hedgehog mediates long-range signalling.* Nature, 2001. **411**(6838): p. 716-20.
102. Alexandre, C., A. Baena-Lopez, and J.P. Vincent, *Patterning and growth control by membrane-tethered Wingless.* Nature, 2014. **505**(7482): p. 180-5.
103. Bartscherer, K. and M. Boutros, *Regulation of Wnt protein secretion and its role in gradient formation.* EMBO Rep, 2008. **9**(10): p. 977-82.
104. Clevers, H. and R. Nusse, *Wnt/beta-catenin signaling and disease.* Cell, 2012. **149**(6): p. 1192-205.
105. Janda, C.Y. and K.C. Garcia, *Wnt acylation and its functional implication in Wnt signalling regulation.* Biochem Soc Trans, 2015. **43**(2): p. 211-6.
106. Bruhn, H., *A short guided tour through functional and structural features of saposin-like proteins.* Biochemical Journal, 2005. **389**: p. 249-257.
107. Bazan, J.F., C.Y. Janda, and K.C. Garcia, *Structural architecture and functional evolution of Wnts.* Dev Cell, 2012. **23**(2): p. 227-32.
108. Ben-David, U., et al., *Selective elimination of human pluripotent stem cells by an oleate synthesis inhibitor discovered in a high-throughput screen.* Cell Stem Cell, 2013. **12**(2): p. 167-79.
109. Noto, A., et al., *Stearoyl-CoA desaturase-1 is a key factor for lung cancer-initiating cells.* Cell Death Dis, 2013. **4**: p. e947.
110. Nile, A.H. and R.N. Hannoush, *Fatty acylation of Wnt proteins.* Nat Chem Biol, 2016. **12**(2): p. 60-9.
111. Liu, J., et al., *Targeting Wnt-driven cancer through the inhibition of Porcupine by LGK974.* Proc Natl Acad Sci U S A, 2013. **110**(50): p. 20224-9.
112. Proffitt, K.D., et al., *Pharmacological inhibition of the Wnt acyltransferase PORCN prevents growth of WNT-driven mammary cancer.* Cancer Res, 2012.
113. Gao, X. and R.N. Hannoush, *Single-cell imaging of Wnt palmitoylation by the acyltransferase porcupine.* Nat Chem Biol, 2014. **10**(1): p. 61-8.
114. Biechele, S., B.J. Cox, and J. Rossant, *Porcupine homolog is required for canonical Wnt signaling and gastrulation in mouse embryos.* Dev Biol, 2011. **355**(2): p. 275-85.

115. Kadowaki, T., et al., *The segment polarity gene porcupine encodes a putative multitransmembrane protein involved in Wingless processing.* Genes Dev, 1996. **10**(24): p. 3116-28.
116. Hardy, R.Y. and M.D. Resh, *Identification of N-terminal residues of Sonic Hedgehog important for palmitoylation by Hedgehog acyltransferase.* J Biol Chem, 2012. **287**(51): p. 42881-9.
117. Rios-Esteves, J., B. Haugen, and M.D. Resh, *Identification of key residues and regions important for porcupine-mediated Wnt acylation.* J Biol Chem, 2014. **289**(24): p. 17009-19.
118. Buglino, J.A. and M.D. Resh, *Hhat is a palmitoylacyltransferase with specificity for N-palmitoylation of Sonic Hedgehog.* J Biol Chem, 2008. **283**(32): p. 22076-88.
119. Wakil, S.J., J.K. Stoops, and V.C. Joshi, *Fatty acid synthesis and its regulation.* Annu Rev Biochem, 1983. **52**: p. 537-79.
120. Barrott, J.J., et al., *Deletion of mouse Porcn blocks Wnt ligand secretion and reveals an ectodermal etiology of human focal dermal hypoplasia/Goltz syndrome.* Proc Natl Acad Sci U S A, 2011. **108**(31): p. 12752-7.
121. Miranda, M., et al., *Identification of the WNT1 residues required for palmitoylation by Porcupine.* FEBS Lett, 2014. **588**(24): p. 4815-24.
122. Wang, B., J.F. Fallon, and P.A. Beachy, *Hedgehog-regulated processing of Gli3 produces an anterior/posterior repressor gradient in the developing vertebrate limb.* Cell, 2000. **100**(4): p. 423-34.
123. Fu, J., et al., *Reciprocal regulation of Wnt and Gpr177/mouse Wntless is required for embryonic axis formation.* Proc Natl Acad Sci U S A, 2009. **106**(44): p. 18598-603.
124. Mulligan, K.A., et al., *Secreted Wingless-interacting molecule (Swim) promotes long-range signaling by maintaining Wingless solubility.* Proc Natl Acad Sci U S A, 2011.
125. Zhang, X., et al., *Notum is required for neural and head induction via Wnt deacylation, oxidation, and inactivation.* Dev Cell, 2015. **32**(6): p. 719-30.
126. Kakugawa, S., et al., *Notum deacylates Wnt proteins to suppress signalling activity.* Nature, 2015. **519**(7542): p. 187-92.
127. Abrami, L., et al., *Palmitoylation and ubiquitination regulate exit of the Wnt signaling protein LRP6 from the endoplasmic reticulum.* Proc Natl Acad Sci U S A, 2008. **105**(14): p. 5384-9.
128. Hodson, L., C.M. Skeaff, and B.A. Fielding, *Fatty acid composition of adipose tissue and blood in humans and its use as a biomarker of dietary intake.* Prog Lipid Res, 2008. **47**(5): p. 348-80.
129. Kirchner, H., et al., *GOAT links dietary lipids with the endocrine control of energy balance.* Nat Med, 2009. **15**(7): p. 741-5.
130. Innis, S.M. and D.J. King, *trans Fatty acids in human milk are inversely associated with concentrations of essential all-cis n-6 and n-3 fatty acids and determine trans, but not n-6 and n-3, fatty acids in plasma lipids of breast-fed infants.* Am J Clin Nutr, 1999. **70**(3): p. 383-90.

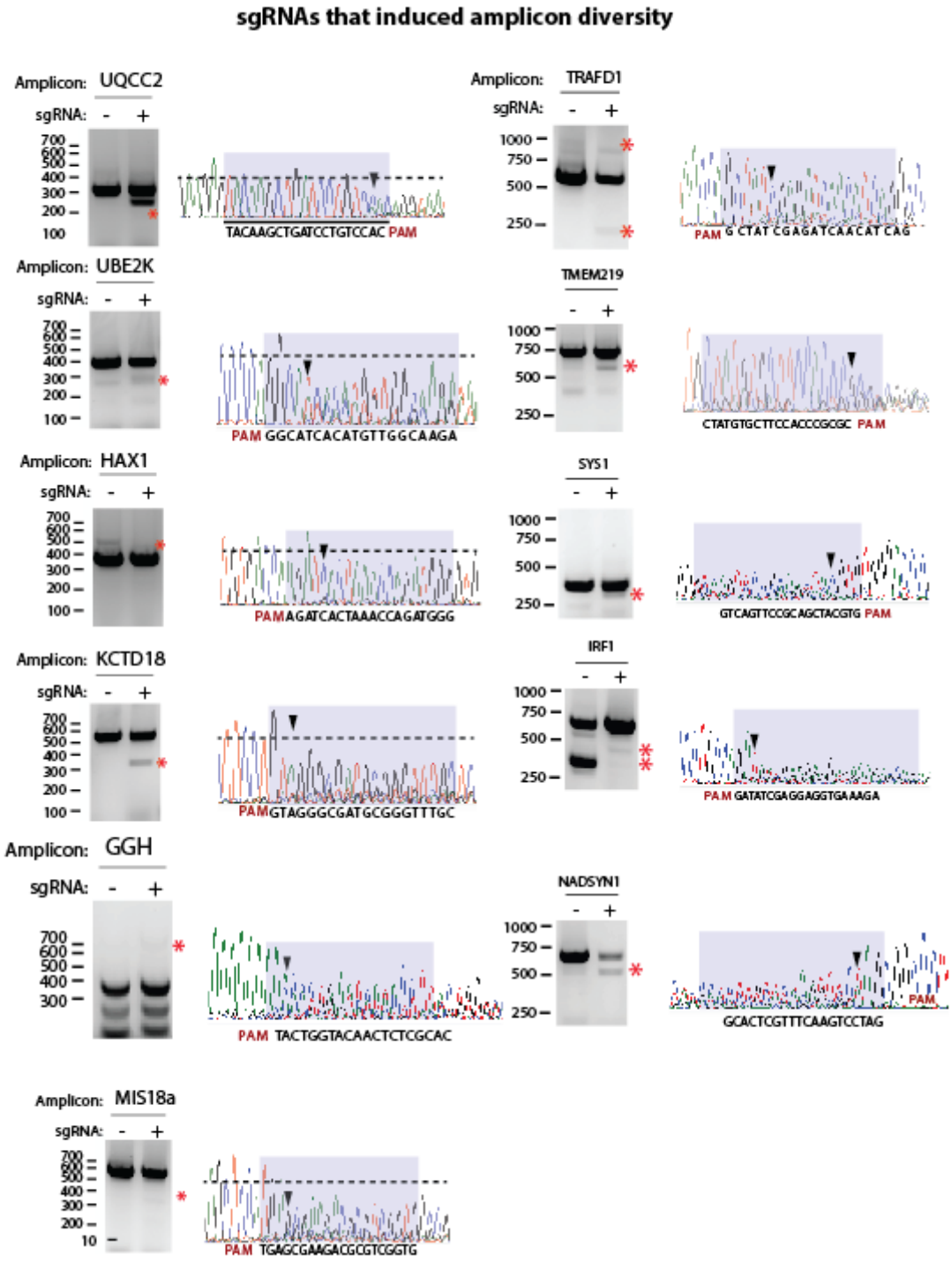
131. Dodge, M.E., et al., *Diverse chemical scaffolds support direct inhibition of the membrane-bound O-acyltransferase porcupine.* J Biol Chem, 2012. **287**(27): p. 23246-54.
132. Liu, A., A.L. Joyner, and D.H. Turnbull, *Alteration of limb and brain patterning in early mouse embryos by ultrasound-guided injection of Shh-expressing cells.* Mech Dev, 1998. **75**(1-2): p. 107-15.
133. Duraiswamy, A.J., et al., *Discovery and Optimization of a Porcupine Inhibitor.* J Med Chem, 2015.
134. Hemminki, A., et al., *A serine/threonine kinase gene defective in Peutz-Jeghers syndrome.* Nature, 1998. **391**(6663): p. 184-7.
135. Giardiello, F.M., et al., *Increased risk of cancer in the Peutz-Jeghers syndrome.* N Engl J Med, 1987. **316**(24): p. 1511-4.
136. Hearle, N., et al., *Frequency and spectrum of cancers in the Peutz-Jeghers syndrome.* Clin Cancer Res, 2006. **12**(10): p. 3209-15.
137. Ji, H., et al., *LKB1 modulates lung cancer differentiation and metastasis.* Nature, 2007. **448**(7155): p. 807-10.
138. Shah, U., N.E. Sharpless, and D.N. Hayes, *LKB1 and lung cancer: more than the usual suspects.* Cancer Res, 2008. **68**(10): p. 3562-5.
139. Sanchez-Cespedes, M., et al., *Inactivation of LKB1/STK11 is a common event in adenocarcinomas of the lung.* Cancer Res, 2002. **62**(13): p. 3659-62.
140. Zhang, X., et al., *Expression and transcriptional profiling of the LKB1 tumor suppressor in cervical cancer cells.* Gynecol Oncol, 2014. **134**(2): p. 372-8.
141. Wingo, S.N., et al., *Somatic LKB1 mutations promote cervical cancer progression.* PLoS One, 2009. **4**(4): p. e5137.
142. Alessi, D.R., K. Sakamoto, and J.R. Bayascas, *LKB1-dependent signaling pathways.* Annu Rev Biochem, 2006. **75**: p. 137-63.
143. Shaw, R.J., et al., *The LKB1 tumor suppressor negatively regulates mTOR signaling.* Cancer Cell, 2004. **6**(1): p. 91-9.
144. Shackelford, D.B. and R.J. Shaw, *The LKB1-AMPK pathway: metabolism and growth control in tumour suppression.* Nat Rev Cancer, 2009. **9**(8): p. 563-75.
145. Jacob, L.S., et al., *Genome-wide RNAi screen reveals disease-associated genes that are common to Hedgehog and Wnt signaling.* Sci Signal, 2011. **4**(157): p. ra4.
146. Goodwin, J.M., et al., *An AMPK-independent signaling pathway downstream of the LKB1 tumor suppressor controls Snail1 and metastatic potential.* Mol Cell, 2014. **55**(3): p. 436-50.
147. Qian, D., et al., *Wnt5a functions in planar cell polarity regulation in mice.* Dev Biol, 2007. **306**(1): p. 121-33.
148. Gao, B., et al., *Wnt signaling gradients establish planar cell polarity by inducing Vangl2 phosphorylation through Ror2.* Dev Cell, 2011. **20**(2): p. 163-76.
149. Oishi, I., et al., *The receptor tyrosine kinase Ror2 is involved in non-canonical Wnt5a/JNK signalling pathway.* Genes Cells, 2003. **8**(7): p. 645-54.

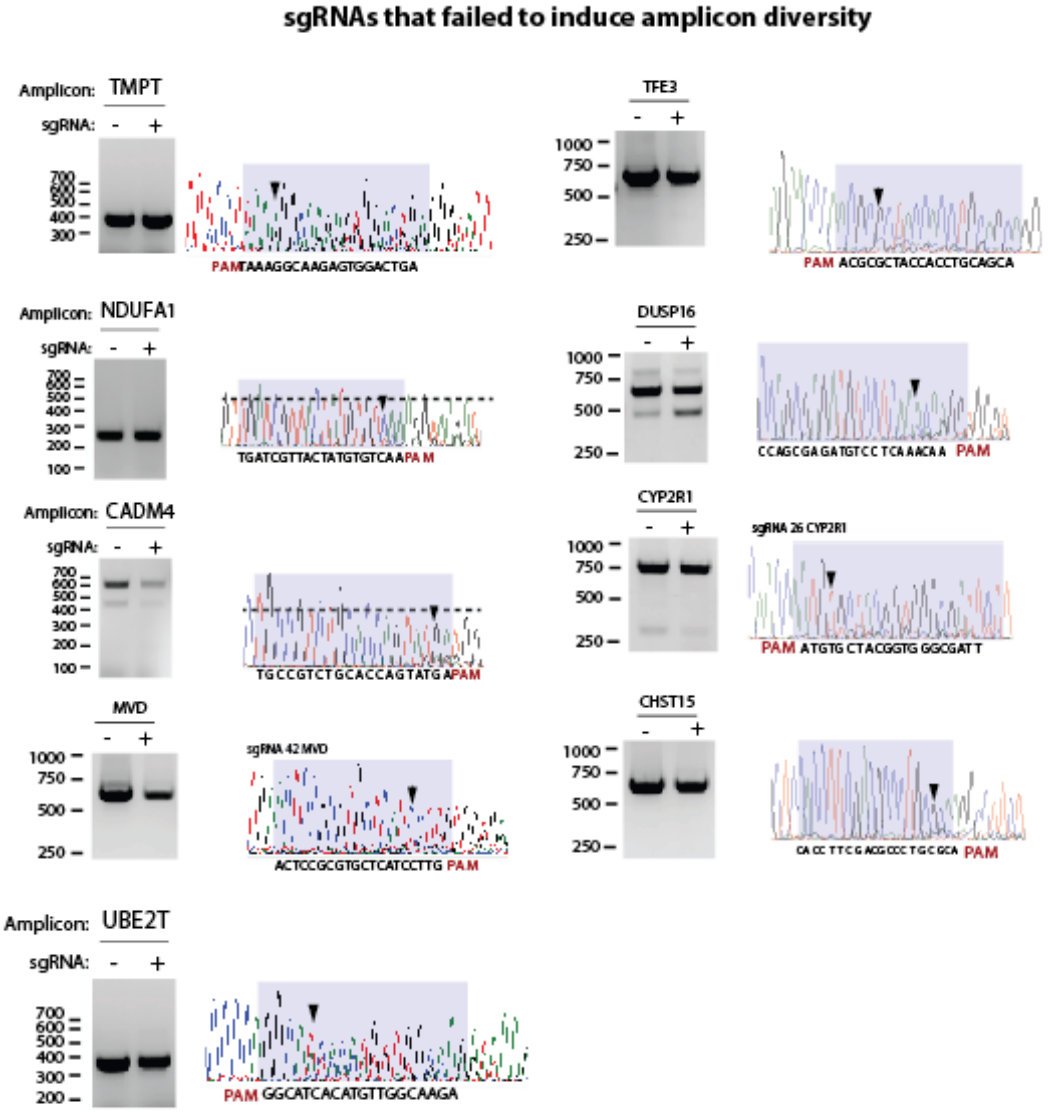
150. Ho, H.Y., et al., *Wnt5a-Ror-Dishevelled signaling constitutes a core developmental pathway that controls tissue morphogenesis.* Proc Natl Acad Sci U S A, 2012. **109**(11): p. 4044-51.
151. Faubert, B., et al., *Loss of the tumor suppressor LKB1 promotes metabolic reprogramming of cancer cells via HIF-1alpha.* Proc Natl Acad Sci U S A, 2014. **111**(7): p. 2554-9.
152. Mick, D.U., et al., *Proteomics of Primary Cilia by Proximity Labeling.* Dev Cell, 2015. **35**(4): p. 497-512.
153. Contreras, C.M., et al., *Lkb1 inactivation is sufficient to drive endometrial cancers that are aggressive yet highly responsive to mTOR inhibitor monotherapy.* Dis Model Mech, 2010. **3**(3-4): p. 181-93.
154. Pena, C.G., et al., *LKB1 loss promotes endometrial cancer progression via CCL2-dependent macrophage recruitment.* J Clin Invest, 2015. **125**(11): p. 4063-76.
155. van Lier, M.G., et al., *High cancer risk in Peutz-Jeghers syndrome: a systematic review and surveillance recommendations.* Am J Gastroenterol, 2010. **105**(6): p. 1258-64; author reply 1265.
156. Giardiello, F.M., et al., *Very high risk of cancer in familial Peutz-Jeghers syndrome.* Gastroenterology, 2000. **119**(6): p. 1447-53.
157. Klumpen, H.J., et al., *mTOR inhibitor treatment of pancreatic cancer in a patient With Peutz-Jeghers syndrome.* J Clin Oncol, 2011. **29**(6): p. e150-3.
158. Morton, J.P., et al., *LKB1 haploinsufficiency cooperates with Kras to promote pancreatic cancer through suppression of p21-dependent growth arrest.* Gastroenterology, 2010. **139**(2): p. 586-97, 597 e1-6.
159. Segal, D.J. and J.F. Meckler, *Genome engineering at the dawn of the golden age.* Annu Rev Genomics Hum Genet, 2013. **14**: p. 135-58.
160. Lykke-Andersen, S. and T.H. Jensen, *Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes.* Nat Rev Mol Cell Biol, 2015. **16**(11): p. 665-77.
161. Mendell, J.T. and H.C. Dietz, *When the message goes awry: disease-producing mutations that influence mRNA content and performance.* Cell, 2001. **107**(4): p. 411-4.
162. You, K.T., et al., *Selective translational repression of truncated proteins from frameshift mutation-derived mRNAs in tumors.* PLoS Biol, 2007. **5**(5): p. e109.
163. Shi, J., et al., *Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains.* Nat Biotechnol, 2015. **33**(6): p. 661-7.
164. Zhou, M., et al., *Non-optimal codon usage affects expression, structure and function of clock protein FRQ.* Nature, 2013. **495**(7439): p. 111-5.
165. Yu, C.H., et al., *Codon Usage Influences the Local Rate of Translation Elongation to Regulate Co-translational Protein Folding.* Mol Cell, 2015. **59**(5): p. 744-54.
166. Popp, M.W. and L.E. Maquat, *Leveraging Rules of Nonsense-Mediated mRNA Decay for Genome Engineering and Personalized Medicine.* Cell, 2016. **165**(6): p. 1319-22.
167. Garneau, N.L., J. Wilusz, and C.J. Wilusz, *The highways and byways of mRNA decay.* Nat Rev Mol Cell Biol, 2007. **8**(2): p. 113-26.

168. Burset, M., I.A. Seledtsov, and V.V. Solovyev, *Analysis of canonical and non-canonical splice sites in mammalian genomes.* Nucleic Acids Res, 2000. **28**(21): p. 4364-75.
169. Srebrow, A. and A.R. Kornblihtt, *The connection between splicing and cancer.* J Cell Sci, 2006. **119**(Pt 13): p. 2635-41.
170. Masuda, K., et al., *Characterization of the STK11 splicing variant as a normal splicing isomer in a patient with Peutz-Jeghers syndrome harboring genomic deletion of the STK11 gene.* Hum Genome Var, 2016. **3**: p. 16002.
171. Ishigaki, Y., et al., *Evidence for a pioneer round of mRNA translation: mRNAs subject to nonsense-mediated decay in mammalian cells are bound by CBP80 and CBP20.* Cell, 2001. **106**(5): p. 607-17.
172. Maquat, L.E., W.Y. Tarn, and O. Isken, *The pioneer round of translation: features and functions.* Cell, 2010. **142**(3): p. 368-74.
173. Frith, M.C., et al., *Pseudo-messenger RNA: phantoms of the transcriptome.* PLoS Genet, 2006. **2**(4): p. e23.
174. Andrews, S.J. and J.A. Rothnagel, *Emerging evidence for functional peptides encoded by short open reading frames.* Nat Rev Genet, 2014. **15**(3): p. 193-204.
175. Ingolia, N.T., *Ribosome profiling: new views of translation, from single codons to genome scale.* Nat Rev Genet, 2014. **15**(3): p. 205-13.
176. Martinez-Salas, E., D. Pineiro, and N. Fernandez, *Alternative Mechanisms to Initiate Translation in Eukaryotic mRNAs.* Comp Funct Genomics, 2012. **2012**: p. 391546.
177. Weingarten-Gabbay, S., et al., *Comparative genetics. Systematic discovery of cap-independent translation sequences in human and viral genomes.* Science, 2016. **351**(6270).
178. Pasca di Magliano, M. and M. Hebrok, *Hedgehog signalling in cancer formation and maintenance.* Nat Rev Cancer, 2003. **3**(12): p. 903-11.
179. Taylor, M.D., et al., *Mutations in SUFU predispose to medulloblastoma.* Nat Genet, 2002. **31**(3): p. 306-10.
180. Matera, A.G. and Z. Wang, *A day in the life of the spliceosome.* Nat Rev Mol Cell Biol, 2014. **15**(2): p. 108-21.
181. Fu, X.D. and M. Ares, Jr., *Context-dependent control of alternative splicing by RNA-binding proteins.* Nat Rev Genet, 2014. **15**(10): p. 689-701.
182. Cartegni, L., S.L. Chew, and A.R. Krainer, *Listening to silence and understanding nonsense: exonic mutations that affect splicing.* Nat Rev Genet, 2002. **3**(4): p. 285-98.
183. Liu, H.X., M. Zhang, and A.R. Krainer, *Identification of functional exonic splicing enhancer motifs recognized by individual SR proteins.* Genes Dev, 1998. **12**(13): p. 1998-2012.
184. Liu, G., et al., *A novel mechanism for Wnt activation of canonical signaling through the LRP6 receptor.* Mol Cell Biol, 2003. **23**(16): p. 5825-35.
185. Rallabhandi, P., et al., *Sumoylation of topoisomerase I is involved in its partitioning between nucleoli and nucleoplasm and its clearing from nucleoli in response to camptothecin.* J Biol Chem, 2002. **277**(42): p. 40020-6.
186. Lu, Q.L., et al., *The status of exon skipping as a therapeutic approach to duchenne muscular dystrophy.* Mol Ther, 2011. **19**(1): p. 9-15.

187. Touznik, A., J.J. Lee, and T. Yokota, *New developments in exon skipping and splice modulation therapies for neuromuscular diseases.* Expert Opin Biol Ther, 2014. **14**(6): p. 809-19.
188. Cirak, S., et al., *Exon skipping and dystrophin restoration in patients with Duchenne muscular dystrophy after systemic phosphorodiamidate morpholino oligomer treatment: an open-label, phase 2, dose-escalation study.* Lancet, 2011. **378**(9791): p. 595-605.
189. Goyenvalle, A., et al., *Functional correction in mouse models of muscular dystrophy using exon-skipping tricyclo-DNA oligomers.* Nat Med, 2015. **21**(3): p. 270-5.
190. Tabebordbar, M., et al., *In vivo gene editing in dystrophic mouse muscle and muscle stem cells.* Science, 2016. **351**(6271): p. 407-11.
191. Long, C., et al., *Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy.* Science, 2016. **351**(6271): p. 400-3.
192. Nelson, C.E., et al., *In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy.* Science, 2016. **351**(6271): p. 403-7.
193. Kim, H. and J.S. Kim, *A guide to genome engineering with programmable nucleases.* Nat Rev Genet, 2014. **15**(5): p. 321-34.
194. Julien, P., et al., *The complete local genotype-phenotype landscape for the alternative splicing of a human exon.* Nat Commun, 2016. **7**: p. 11558.
195. Pan, Q., et al., *Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing.* Nat Genet, 2008. **40**(12): p. 1413-5.
196. Wang, E.T., et al., *Alternative isoform regulation in human tissue transcriptomes.* Nature, 2008. **456**(7221): p. 470-6.
197. Mills, R.E., et al., *Natural genetic variation caused by small insertions and deletions in the human genome.* Genome Res, 2011. **21**(6): p. 830-9.
198. Ran, F.A., et al., *Genome engineering using the CRISPR-Cas9 system.* Nat Protoc, 2013. **8**(11): p. 2281-308.
199. Kulak, O. and L. Lum, *A multiplexed luciferase-based screening platform for interrogating cancer-associated signal transduction in cultured cells.* J Vis Exp, 2013(77): p. e50369.

APPENDIX A
Location of INDELS generated in the CRISPR screen using Gecko V2 library




sgRNAs that failed to induce amplicon diversity
Amplicon: TMPT
sgRNA: - +
700
600
500
400
300
PAMTAAAGGCAAGAGTGGACTGA
TFE3
- +
1000
750
500
250
PAM ACGCGCTACCACCTGCAGCA
Amplicon: NDUFA1
sgRNA: - +
700
600
500
400
300
200
100
TGATCGTTACTATGTGTCAAPAM
DUSP16
- +
1000
750
500
250
CCAGCGAGATGTCCTCAAACAA PAM
Amplicon: CADM4
sgRNA: - +
700
600
500
400
300
200
100
TGCCGTCTGCACCAGTATGAPAM
CYP2R1
- +
1000
750
500
250
sgRNA 26 CYP2R1
PAM ATGTG CTACGGTG GGCGATT
MVD
- +
1000
750
500
250
sgRNA 42 MVD
ACTCCGCGTGCTCATCCTTG PAM
CHST15
- +
1000
750
500
250
CA CC TT C G ACG CCC T GC GCA PAM
Amplicon: UBE2T
sgRNA: - +
700
600
500
400
300
200
PAM GGCATCACATGTTGGCAAGA

APPENDIX B
Average distance of sgRNA targeted sequences from exon junction sites

sgRNAs	Distance from exon junction 5'	Distance from exon junction 3'	Exon skipping?
MIS18A	163	161	Yes
GGH	93	2	Yes
UBE2K	21	59	Yes
UQCC2	48	2	Yes
KCTD18	85	107	Yes
TRAFD1	112	208	Yes
TMEM219	85	33	Yes
SYS1	7	135	Yes
IRF1	80	77	Yes
NADSYN1	45	53	Yes
Average distance from exon junction	73.9	83.7	
NDUFA1	68	2	No
DUSP16	142	2	No
CYP2R1	582	31	No
CHST15	135	186	No
TFE3	24	201	No
MVD	178	2	No
UBE2T	60	29	No
TMPT	30	43	No
CADM4	65	62	No
HAX1	139	0	No
Average distance from exon junction	142.3	55.8	

APPENDIX C
sgRNAs and primer sequences

sgRNA sequences

sgRNA name	Strand	sgRNA sequence
Lkb1 Exon 1	Anti-sense	GAGCTTGGCCCGCTTGCGGCG
Sufu Exon 2	Sense	CTGCTAACATCCCCGAGCAC
Sufu Exon 3a	Sense	ATGGCCCGCAGAGTTAATGC
Sufu Exon 3b	Sense	CACGATACGTGTTCCAGTC
Sufu Exon 8	Anti-Sense	CCAATCAACCCTCAGCGGCA

PCR primers for amplifying genomic DNA sequences

Primer name	Sequence
Stk11 Genomic 5' UTR F	GTTTCTGTTGGAAGAAGGGTTT
Stk11 Genomic Intron 1 R	AGCTCAGGGTGTTAAGAGGAAG
Sufu Exon 2	

PCR primers for amplifying cDNA sequences

Primer name	Sequence
Lkb1 5’ UTR F	AGGGCTGGCGGCGGGACTCCAG
Lkb1 Exon 4 R	TTTGAGGGTGCCACCGGTGGTG

sgRNA sequences from the screen

sgRNA name	sgRNA sequence
TTC13	GACCAGTACTCTGCCGAGTG
SMAD1	TTAGCTCAGTTCCGTAACTT
GHITM	TGGCTCCTCTGACAATATTA
MIS18A	CACCGACGCGTCTTCGCTCA
UBE2T	TCTTGCCAACATGTGATGCC
PID1	GGCTCATGGAAGATGACCGC
TPMT	TCAGTCCACTCTTGCCTTTA
G6PC2	GTTTCTTGGACCCACCAGTA
PPP2R3A	TCTGTGGCTGTCCTCTCTAT
TNFRSF1A	AGAACCAGTACCGGCATTAT
STK11	CAGCCGCCCGAGATTGCCAA
TDRD12	GTCTTCCGCAGACCAGTACC
ZNF480	TCACTTACATCTGTCTGAAC
CENPL	ACCCTGTAGCCCTCAAAGTC
PDLIM2	AGTGCTGGCGACTCGCTTCC
GTF3C3	TACTCTAGCCATGATATATG
PCDP1	CCATACGAATGTGTCTTCAC
UPB1	CATGCTCGCTGTCTCGTTCC
C1orf61	ATGTGGCTCTATCGTCCACA
IMP4	TCCCACTGAGTTACGCCGAG
CADM4	TGCCGTCTGCACCAGTATGA
UBE2K	GCAATGACAATAATACCGTG
OXA1L	CGGTTTCCCAAGCCATTGCG
TMEM66	AGACTACAAACGCGATCCCA
SEC13	TGCCAGGATGTTGCCGTACA
MRPL43	AACAAGCCGACCACGTTCCG
QTRTD1	GTGCTCGAGTGTATTGAAAG
KDM1A	CCGGCCCTACTGTCGTGCCT
ARL8A	GTTGTGGAGCTCGTTCTTAG
RABGAP1	GGAGGCCGAGTGTAAGATAC
SCFD2	CCTGCCAATCAAGATGAGTA
GGH	GTGCGAGAGTTGTACCAGTA
NFXL1	GGAGTCCATCTCCGCGGCGC
RAD17	TTCAAGGGGATGTTTAATAC
TMEM47	CTGACCCCCTTGAAGCTGGT
NDUFA1	TGATCGTTACTATGTGTCAA
GLYATL1	CCGTGTTCACTTCGCGCTCC
UQCC2	TACAAGCTGATCCTGTCCAC
RHBDL3	AGCCGATCTGCCCATCCGCG
KCTD18	GCAAACCCGCATCGCCCTAC
NCKAP1L	CCCCCTGACCGCGATCATTC
CORO7	TGCTCGAATGTCACTGATCC

DSCAM	GGAGCCCTATACAGTCCGTG
ANKRD40	TGCCAGTCCAGTTAACATCA
PLA2G4A	ACACCACTACCGTAAACTTG
SBSN	ACAACCACGCCGTTAGCCTC
HAX1	CCCATCTGGTTTAGTGATCT
MSANTD2	GTCATAATGATCCGTCCTTC
SYS1	GTCAGTTCCGCAGCTACGTG
MANSC4	CCCAGGTCTTCTAATCAATC
E2F1	CGTCATCATCTCCGCCGCGC
SLC9A4	CTCAGTATGCATCTAACGCT
NPY	TCCGCTGGTGCGTCCTCGCC
TMIE	CGCACTCGGGGTGTGCCTCG
ALS2CR12	AGTTCTTGCGTGGTAGTTGA
ALS2CR12	AGTTCTTGCGTGGTAGTTGA
AHCYL2	TCCTAGTTGCGACTCCGCCT
PDE4DIP	CCCGCTTGTAGATGTCCTCC
YWHAZ	GCCGCTGGTGATGACAAGAA
FGF14	TACTTGCAAATGCACCCCGA
NABP1	GATATTATTCGGTTGACCAG
TNFRSF21	CGCCATGTTGACCGTGCCAC
CHST15	CACCTTCGACGCCCTGCGCA
USP50	TCTCTGCAGCATCTTGCCGC
NAPSA	TCGTACCTCTCTCGAACTAC
CCNJ	TGTCATAGCGGTCCATAAAC
PARK2	GTGTCAGAATCGACCTCCAC
CUX1	CTGTTCGATAAGCCGCTTTC
SC5D	CAGAGGCCTTCATCATAGAC
DUSP16	CCAGCGAGATGTCCTCAACA
ATPAF2	GCTCTTTACCGTCCCCAGCG
NADSYN1	GCACTCGTTTCAAGTCCTAG
TMEM2	CGGATTGTTGCCATAGCTGT
CYP2R1	AATCGCCCACCGTAGCACAT
FGF18	CATCGTCCCGAGCCCGCGTC
IRF1	TCTTTCACCTCCTCGATATC
GLOD4	GCGACAAAATGATCATCCTC
LRCH4	CTTCCCAGGGTGCCGATGTC
CCDC33	TTCACCGTGTCCCCCCAGAT
TICRR	GTGCCCAAAGTGCTGCTCGT
TMEM219	CTATGTGCTTCCACCCGCGC
ATP1B1	ATAAAGCTCAACCGAGTTCT
TMEM65	ACCCATATTGAAATGTCTAT
GAS2L3	ATACCGCAAACTTCCTTCAC
PKN1	TCATCCGCATGCAACTCCGC
USP36	ATCGCCCGACACTTCCGCTT
TPR	CAAGAACTTGAATACTTAAC
TFE3	TGCTGCAGGTGGTAGCGCGT
RAVER2	GCTGCTAATGTACTTCGCCC
MVD	ACTCCGCGTGCTCATCCTTG
HSP90B1	CTTGGTACCATAGCCAAATC

CRISPLD1	CAGGGCCGCCCATGTTTGTA
GPNMB	ATGAATTTCGATCGTTCTTC
SPTA1	TCGTTGTAAGTTTGCAGCAT
TRAFD1	CTGATGTTGATCTCGATAGC
HOOK2	GGGCCAGGCTAGTCAGCGCC

PCR primers for amplifying genomic DNA from the screen

Primer name (5')	Sequence	Primer name (3')	Sequence
SMAD1_gDNA F	GTATTTGTTCCTGCATTTGTTC	SMAD1_gDNA R	GTGACACTGTGATAACACTGT
MIS18A_gDNA F	CGATTTGTAGGTAATGGCAG	MIS18A_gDNA R	CTGCCACAGCTCACACTCT
UBE2T_gDNA F	TGACAGAGTGAAACTGTGTCT	UBE2T_gDNA R	TTAAATCTTGGCTATAAATCAGC
TPMT_gDNA F	AATTACTAATCAGGAAAGCCCT	TPMT_gDNA R	CCAGAAAGACTTCATACCTGT
PPP2R3A_gDNA F	TGCTTTAACTGTCAGGTTGGA	PPP2R3A_gDNA R	GTATGTGAGTATAAATGGTTGTA
ZNF480_gDNA F	AGTGCATCACCTTACCCTTT	ZNF480_gDNA R	GCCTTACTAAACATCAAGTAAT
CENPL_gDNA F	GCTGTGTATTTGGAGACAGT	CENPL_gDNA R	TAACTAGATTGTTCAAACTAACC
GTF3C3_gDNA F	CCTTGCTTAGAGAAAACTGCT	GTF3C3_gDNA R	TTTATCTCAGCCAATATCACAC
PCDP1_gDNA F	CTAATCTCCACATTGTCAGAG	PCDP1_gDNA R	TTGTCAGAAGCAACTCTTGC
IMP4_gDNA F	AGTCCTCAAGTCATGAGACTA	IMP4_gDNA R	CAGAGTCAGTGAGGCACTC
CADM4_gDNA F	CAGATTTCCTCCCTGTCCT	CADM4_gDNA R	GAGTCTGAGGTGTCCAACTA
UBE2K_gDNA F	GGCAAATTGTAAAGTAAGGAAG	UBE2K_gDNA R	TAGGAAAACAGCATCTATATCC
OXA1L_gDNA F	GAACCCAGGTTCGAGCTTC	OXA1L_gDNA R	AGAGATGCTATCATTTTGGAGA
MRPL43_gDNA F	TATCCTCCTTCTCGCTTGGT	MRPL43_gDNA R	GAGTGGTTCACAAGGTCACT
KDM1A_gDNA F	AGATGTTATCTGGGAAGAAGG	KDM1A_gDNA R	CAAAGACTTTTCCTTTCTCAC
GGH_gDNA F	GACTTGGAGGAGAGACTTC	GGH_gDNA R	TCCCTGGTTCAAGCGATTCT
RAD17_gDNA F	CTAGCATTCCAGTGATTTTGG	RAD17_gDNA R	TCAGCCCTTGAGTAGTTAGA
TMEM47_gDNA F	CTTCTCCCTTCGGGAGGT	TMEM47_gDNA R	TACTCCAGAACAGGAGGAAT
NDUFA1_gDNA F	AGGAGATTCTGTTCTTTGTACT	NDUFA1_gDNA R	GATATAGGCAACAATCCACAG
UQCC2_gDNA F	GAGACTGTACTGGTTCTTGC	UQCC2_gDNA R	CTCTGTTCCCTGATGCCCA
KCTD18_gDNA F	GGTGTTTGAATAACCAAAAGTC	KCTD18_gDNA R	GGTTAATGTGAAAACACAGGC
CORO7_gDNA F	AGAGGCAGAGATTGCAGTG	CORO7_gDNA R	GTCTAAGGTCTGGCATGGG
ANKRD40_gDNA F	GTTACCTGGTTCTAGCTGAA	ANKRD40_gDNA R	AGGAGAATCACTTGAACCTG
SBSN_gDNA F	TCTTCTGGACAGTGAGGAG	SBSN_gDNA R	GAGCTGTCCAAACAGACAC
HAX1_gDNA F	GGAAGTTTACCAGCCTTTTG	HAX1_gDNA R	TGAGATGAACCCCTTTGTCT
MSANTD2_gDNA F	GAGACATCATGCAGAATATTGT	MSANTD2_gDNA R	TCTCAATACATTCTTCATAGCC

PCR primers for amplifying cDNA sequences from the screen

Primer name (5')	5' Primers	Primer name (3')	3' Primers
TTC13 Exon 1 F	CACCGAGCACTACTCGCCGCT	TTC13 Exon 7 R	CACTGCTTCATTAATTCGTCC
SMAD1 Exon 1 F	GGCAGGCTGCAAGTCTCCCAC	SMAD1 Exon 4 R	CCCACACGATTGTTGAGCTCATAG
GHITM Exon 4 F	GGATAGAATTCATTCCACCTA	GHITM Exon 8 R	CAGTTGCCAGCATAGTTGCAACTCG

MIS18A Exon 1 F	ATGGCAGGCGTTCGGTCACTG	MIS18A Exon 5 R	AATGCTTTCAAGACATCTTCC
UBE2T Exon 1 F	ATGCAGAGAGCTTCACGTCTG	UBE2T Exon 5 R	TCTTGCATGCTTCTCTGTCCAC
PID1 Exon 1 F	TGCCAAGGATTGCTGGCAACC	PID1 Exon 3 R	CTGGAAGGTATCCATGTGCAC
TPMT Exon 1 F	ATGGATGGTACAAGAACTTCA	TPMT Exon 4 R	ACAGTACAATGAAATGTTCC
G6PC2 Exon 1 F	ATGGATTTCCTTCACAGGAAT	G6PC2 Exon 5 R	AAGGCCTCTGCCACCAGCATG
PPP2R3A Exon 1 F	CAGTCTGTAGAAGTCCTGTTG	PPP2R3A Exon 6 R	TTCCACTCCAAGATCTGTTGA
TNFRSF1A Exon 1 F	ATGGGCCTCTCCACCGTGCCT	TNFRSF1A Exon 6 R	ATCTGGGGTAGGCACAACTTC
STK11 Exon 3 F	GGTGATGGAGTACTGCGTGTG	STK11 Exon 8 R	CAGCCGGAGGATGTTTCTTCC
TDRD12 Exon 1 F	ATGCTCCAGCTCCTGGTGCTG	TDRD12 Exon 5 R	TACTGAATAGCTGCATTGTCC
ZNF480 Exon 1 F	GGAGTCAGGGATGGCTCTTCC	ZNF480 Exon 3 R	AAGGTTTTTCTCTAAGGTGTAC
CENPL Exon 2 F	GGGACCCGGAAGCATTTCTTG	CENPL Exon 4 R	TGTCAAATATGCTAACACTCC
PDLIM2 Exon 1 F	GCGCGGAGTCCACTGACCGGC	PDLIM2 Exon 6 R	GCCTGAGTAGGACAGGCGGTC
GTF3C3 Exon F	GGAGATGGTTCTCAATCGTGA	GTF3C3 Exon 12 R	TGTTAATAGCAGAAGTAACATC
PCDP1 Exon 5 F	CATGCCTATCCAGTCATGAACT	PCDP1 Exon 12 R	TGAAATTTCAGTGCCCTGGTCC
UPB1 Exon 2 F	CGCAAATGCCCCTGTGGCAGA	UPB1 Exon 7 R	GCGTTTCTGGCCTCGATGGGC
C1orf61 Exon 1 F	ATGTTCCTGACTGAGGATCTC	C1orf61 Exon 5 R	TCAGAGGGAATCCAGCCCAAC
IMP4 Exon 1 F	CTGCGCCGCGAGGCCCGCCTG	IMP4 Exon 5 R	CGACCTCGGTTCATTCGCTGG
CADM4 Exon 1 F	ATGGGCCGGGCCCGGCGCTTC	CADM4 Exon 5 R	TGCCACGCTCCAGACCTTGCC
UBE2K Exon 4 F	TAGTTCCGTCACAGGGGCTAT	UBE2K Exon 7 R	ACTCAGAAGCAATTCTGTTGC
OXA1L Exon 1 F	GACTAATGTGCGGACGCCGGG	OXA1L Exon 4 R	CTGGCAAGTGATTGTGGATCC
TMEM66 Exon 2 F	GTGTCAGAACAAAGGCTGGGA	TMEM66 Exon 4 R	TCCCAAGCCTGTCCAGAACCC
SEC13 Exon 2 F	CACTGTGGATACCTCCCATGA	SEC13 Exon 6 R	CCTGATGCAAACCTCTTGATG
MRPL43 Exon 1 F	AGCTTCAGCGTCAGCCGCGA	MRPL43 Exon 5 R	GGACAGCATAGGGGCTGGAGC
QTRTD1 Exon 4 F	GAAAGTCTGTTGACCGATCAC	QTRTD1 Exon 8 R	CAGGAACATCCTCTCACCAGC
KDM1A Exon 1 F	GCAAGAAAGAGCCTCCGCGGG	KDM1A Exon 4 R	CTAATTGTTGGAGAGTAGCCTC
ARL8A Exon 1 F	CAAGACCACCTTCGTCAACGT	ARL8A Exon 7 R	GTTGACGAAGGTGGTCTTGC
RABGAP1 Exon 20 F	GGCTGGAAGAAGAGTCTGCTC	RABGAP1 Exon 25 R	TCAGCAAGTCTCTTTCCCTTGA
SCFD2 Exon 1 F	GACTGCTGCAGGCAGGGCATC	SCFD2 Exon 5 R	CTTAATCATGGGCAGCAGCTG
GGH Exon 1 F	CTGCTACTCTGCGGGGCGGCG	GGH Exon 6 R	TCAGAGGTTCTACTGCTAATG
NFXL1 Exon 1 F	ATGGAAGCTTCCTGGCGCCAG	NFXL1 Exon 4 R	GAAGATACAAGAAACTGGCTG
RAD17 Exon 2 F	GAACCATGGGTGGATAAATATA	RAD17 Exon 7 R	AGAGGACATCGACCAATCCTC
TMEM47 Exon 1 F	ATGGCTTCGGCGGGCAGCGGC	TMEM47 Exon 3 R	CTAGTAGTAGTCTTCATAGTT
NDUFA1 Exon 1 F	ATGTGGTTCGAGATTCTCCCC	NDUFA1 Exon 3 R	TTAATCAATGTTCTCCAAACC
GLYATL1 Exon 1 F	ATGTTCAAACTGTGTTCAAATA	GLYATL1 Exon 6 R	CTTCCAAGCTTGCTTTTACTG
UQCC2 Exon 1 F	GACTTGGGCGCTTACCTGCGA	UQCC2 Exon 4 R	TCAGGCCTTATGATCCTCCTC
RHBDL3 Exon 1 F	GGCCGAGGAGCGGCTGCCCGC	RHBDL3 Exon 5 R	TAACGTGGATGAGTTACCTGC
KCTD18 Exon 1 F	GGCATCTATGTTCAGTGGTCG	KCTD18 Exon 4 R	TCCCAAGCATCAAAGGCATCC
NCKAP1L Exon 1 F	ATGTCTTTGACATCTGCTTAC	NCKAP1L Exon 5 R	ACATGCCAATGAGTATCCGCC
CORO7 Exon 1 F	ATGAACCGCTTCAGGGTGTCC	CORO7 Exon 6 R	ACGGCGCTCTGCACCAGGTCC
DSCAM Exon 1 F	TGTTCCAGAGCTTCGCGAATG	DSCAM Exon 4 R	ATGTGATGAGAAATCTAGATC
ANKRD40 Exon 1 F	GTGCAGAAACTGGTGGAGAGC	ANKRD40 Exon 3 R	GGCCGACACATCACCATTCTG
PLA2G4A Exon 1 F	ATGTCATTTATAGATCCTTAC	PLA2G4A Exon 6 R	TGTTGTCTGAAAGTCTTCTCC

SBSN Exon 1 F	CCACGGCATGGACAAGGTTGC	SBSN Exon 5 R	TTAGGGCATGATGTTGGCGAC
HAX1 Exon 3 F	GGAGAGTGATGCAAGAAGTGA	HAX1 Exon 7 R	CTACCGGGACCGGAACCAACG
MSANTD2 Exon 1 F	GAGCTGGGCTACGAGCGGACC	MSANTD2 Exon 4 R	CTCGGGTCATAATGATCCGTC
SYS1 5'UTR F	ATGGTCCTGTCTGTCAGCGCT	SYS1 Exon3 R	CCAGACACTGCTTTCCTCGCC
MANSC4 Exon1 F	CTGCAAGGAAGCTCTCCCTTA	MANSC4 Exon3 R	ATTTAGATATGTGGGAGATTGT
E2F1 Exon1 F	AGTGGCCGGGACTTTGCAGGC	E2F1 Exon3 R	GATTCAGTGAGGTCTCATAGC
SLC9A4 Exon 1 F	TGGATGCAGTCACTCTCTAGA	SLC9A4 R	TCTTTGTAAAGGCAATTAACA
NPY 5'UTR F	GTGGAGCAGAGGGGCAGGTCC	NPY Exon3 R	TCACCACATTGCAGGGTCTTC
TMIE Exon 1 F	CTCCGCAAGCGGCGCGGTGGC	TMIE Exon 3 R	ACACGACAGTTGAAGACACAG
ALS2CR12 Exon 1 F	CCTGATCTAGCAGCCAAGACT	ALS2CR12 R	AATGATTTATTAGTCTCTTCT
ALS2CR12 Exon 1 F	CCTGATCTAGCAGCCAAGACT	ALS2CR12 R	AATGATTTATTAGTCTCTTCT
AHCYL2 Exon 1 F	ATGTCGGTGCAGGTTGTGTCA	AHCYL2 Exon 3 R	CGATGTCTCATCATCAGAGCT
PDE4DIP Exon 1 F	ATGAAGGGCACAGACAGCGGG	PDE4DIP Exon 3 R	TGTGACTGCACTGGAGGGCCT
YWHAZ Exon1 F	AGCAGATGGCTCGAGAATACA	YWHAZ Exon4 R	TGTATCAAGTTCAGCAATGGC
FGF14 Exon1 F	ACGCATATGCTGCAGTGTCTT	FGF14 Exon4 R	CAGATTCCTGTTGTCTGTACA
NABP1 Exon1 F	ATGAATAGGGTCAACGACCCA	NABP1 Exon4 R	CTTTGTTCTGCTGTCCTCGAT
TNFRSF21 Exon 1 F	ATGGGGACCTCTCCGAGCAGC	TNFRSF21 Exon 4 R	ACAGAGGCAGAAGAGTTGGAT
CHST15 Exon 1 F	CAGCATCACAACTAGGATTGA	TNFRSF21 Exon 4 R	CGTAGAAGAACGTCCAGGCAT
USP50 Exon 1 F	ATGACTTCTCAGCCGTCTCTC	USP50 Exon 4 R	TCTCATATGATCTTCTCCGGG
NAPSA Exon 1 F	ATGTCTCCACCACCGCTGCTG	NAPSA Exon 4 R	TGGGATCAAATCGGTGGTGTA
CCNJ UTR F	GGCTGTGAGGGCCGCGGTTCC	CCNJ Exon 4 R	CCTCGAAGAAGCCACACATGC
PARK2 Exon 1 F	ATGATAGTGTTTGTCAGGTTC	PARK2 Exon 4 R	AAGCTGTTGTAGATTGATCTA
CUX1 Exon 1 F	ATGGCGGCCAATGTGGGATCG	CUX1 Exon 4 R	GCTTCTTTTACTCAGTGCATC
SC5D EXON 1 F	GTCTTCGATCATGCATTAATG	SC5D EXON 5 R	TGGTGGTCTGTATGATGAGC
DUSP16 EXON 1 F	TACAGAGCTCATCCAGCATT	DUSP16 EXON 6 R	AGTTTGGAGATATAGTAGGTC
ATPAF2 EXON 1 F	CAACCATCCCGTCTCCAG	ATPAF2 EXON 7 R	GCTGGGCAGCTACAAACTC
NADSYN1 EXON 1 F	GGACTTCGAGGGCAATTTG	NADSYN1 EXON 8 R	GATCTCCACGCCATCCAG
TMEM2 EXON 2 F	CACCGTGCATTCTATAGTCAT	TMEM2 EXON 4 R	ACACCAACTAAAGCCCAAGC
CYP2R1 EXON 2 F	CATGAAGATGACAAAAATGGG	CYP2R1 EXON 4 R	GGGCCCATAATTAAATCAATC
FGF18 EXON 1 F	TGCACTTGCCTGTGTTTACA	FGF18 EXON 5 R	TGGTGAAGCCCACGTACCA
IRF1 EXON 1 F	GAGATGCAGATTAATTCCAAC	IRF1 EXON 6 R	GGGGAGAGTGCTGCTGAC
GLOD4 EXON 1 F	CGTTTCTATCGGGACGTCC	GLOD4 EXON 6 R	CAAAAGCTGCTGCATGGTC
LRCH4 EXON 1 F	CCGTAGCTACGACCTGTC	LRCH4 EXON 6 R	CAGGGACTGAAACTCGGG
CCDC33 EXON 1 F	CCTGCCTGCAAGGATGGC	CCDC33 EXON 7 R	CCAGGTCCCTGTTAGCCTG
TICRR EXON 1 F	GTGGATACCACCGAATGGT	TICRR EXON 5 R	CCACTCTGGAACAGGAGAA
TMEM219 EXON 1 F	TCACCCACAGGACTGGCC	TMEM219 3' UTR R	CCAACTCAGAGCCAAGGC
ATP1B1 EXON 1 F	TTCTGGGCAGGACCGGTG	ATP1B1 EXON 6 R	ACTTTATCCTTATCTTCATCTC
TMEM65 EXON 1 F	TGCACCGCTTCGAGTCTAT	TMEM65 EXON 7 R	CATCTTCTTCACCTCCTCCA
GAS2L3 EXON 2 F	AGGCATGAAGCCACTTTGTT	GAS2L3 EXON 6 R	GCATGATTTTGGAATGCTGA
PKN1 EXON 2 F	CAAGGAGCTGAAGCTGAAGG	PKN1 EXON 6 R	CCCAGCTTCTGGTTGGATT
USP36 EXON 2 F	GACTCCACAACCTTGGCAAC	USP36 EXON 7 R	GGCATTCTCTCCACTCAGGA
TPR EXON 1 F	AACAAGCTGCCCAAGTCTGT	TPR EXON 6 R	TTCTTGCTCCAAGCGTTTTT

TFE3 EXON 3 F	TTCTTCATCGAGGGTCTTGC	TFE3 EXON 6 R	CAAGCAGATTCCCTGACACA
RAVER2 EXON 3 F	CTTTGCACAATGGATGGATG	RAVER2 EXON 6 R	AATGCAGGAGAGATGGATGG
MVD EXON 3 F	GACCGGATTTGGCTGAATG	MVD EXON 7 R	GGTTGCTGTCCTTCATGGTC
HSP90B1 EXON 4 F	TGATGAAAATGCTCTTTCTGGA	HSP90B1 EXON 7 R	CATGGGCTCCTCAACAGTTT
CRISPLD1 EXON 2 F	AGAATTTGGGAGCACACTGG	CRISPLD1 EXON 8 R	CCAGCAGGACATTCGTACCT
GPNMB EXON 4 F	CTTCCCTGATGGGAAACCTT	GPNMB EXON 7 R	TCCTGGGGTGTTTGAATCAT
SPTA1 EXON 4 F	AGAAGGGTGACCAGTTGCTG	SPTA1 EXON 8 R	CTCCTGGATCTGAGGTGCAT
TRAFD1 EXON 7 F	GACTCTTTGATGGGCCTGAG	TRAFD1 EXON 10 R	CCTGCCACCTTCACTCCTAC
HOOK2 EXON 8 F	GCAGCGCTGTCTGGATCT	HOOK2 EXON 12 R	ATAGCCATTTCTCGGCCTTC