The Role of an Evolved Novel Splicing Regulatory G Tract in Diversification of Protein Functions

by

Muhammad Sohail

A Thesis submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfilment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Department of Physiology & Pathophysiology

University of Manitoba

Winnipeg

TABLE OF CONTENTS

ABSTRACT

Alternative pre-mRNA splicing greatly contributes to the mammalian proteomic diversity. The novel splice variants often emerge through splicing regulation at/nearby the splice sites (SS). A large group of 3′SS in human genes contain REPA (regulatory elements between the Py and 3′AG) G tracts that mostly appear in mammals as splicing silencers. However, the underlying molecular mechanisms and functional consequences remain unknown.

We have identified a novel class of REPA G tracts $(G)_{5\text{-}8}$ in a group of human genes including *PRMT5* (protein arginine methyl transferase 5) that are significantly enriched in functional clusters of cell growth and proliferation. The *PRMT5* G tracts emerged evolutionarily in mammals and repress splicing through recruitment of mainly hnRNP H that interferes with early spliceosome assembly. The splicing repression creates a shorter PRMT5 isoform (PRMT5S) that inhibits cell cycle progression contrary to the role of the full length protein (PRMT5L). Moreover, the expression of a group of genes involved in cell cycle arrest at interphase is preferentially regulated by PRMT5S.

We further showed that PRMT5S is differentially expressed among cell and tissue types suggesting tissue-specific regulation. It exhibits distinct subcellular localization pattern from that of PRMT5L and opposite effects on cell cycle-specific structural dynamics of the Golgi apparatus. Moreover, these splice variants are differentially expressed during cell differentiation and PRMT5S promotes the differentiation of dendritic cells whereas PRMT5L shows the opposite effect. The expression of a large number of genes including those involved in crucial cellular processes such as differentiation and apoptosis is regulated by these splice variants of PRMT5.

This study provides a direct link between the evolutionary emergence of a novel splicing regulatory G tract element and the generation of a functionally distinct protein isoform. The molecular mechanism underlying the splicing regulation by this G tract is likely common to many mammalian genes and the generation of their protein diversity.

ACKNOWLEDGEMENT

I would like to express my deepest gratitude to my advisor Dr. Jiuyong Xie for his meticulous insights that insinuated the foundation of the studies presented in this thesis. The completion of these studies involves highly valuable and considerate advice from him and his readiness to engage in very thoughtful and reassuring discussions. His passion and dedication for science has been a great source of inspiration and motivation for me to excel in my career.

I am extremely grateful to my advisory committee members Dr.s Janice Dodd, Peter Cattini, Sean Mckenna and Etienne Leygue for their generous and encouraging advice throughout my research progress. Their suggestions and remarks have been really a great stimulation for me to enhance my learning. It has been a great experience to have an excellent group of advisory committee members.

I want to thank all the lab members for their support and a delightful working environment. I want to specially thank Wenguang Cao, Niaz Mahmood and Mike Myschyshyn for their help in my research. I also want to say thanks to Dairong Feng, Lei Lei, Nan Liang and Jason Wu, and former lab members Aleh Razanau, Guodong Liu, Geremy Koumbadinga and Vincent Lobo for their support.

I would like to say thanks to Gail McIndless and Judith Olfert for their valued support throughout my PhD program.

I am thankful to my parents, family, friends, teachers and specially my wife for their unreserved support and compassion.

DEDICATION

This thesis is dedicated to my wife, parents and family.

List of Tables

List of Figures

ABBREVIATIONS

BMDC	bone marrow-derived dendritic cell
BP	branch point
CaMK	Calcium/calmodulin-dependent kinase
CaRRE	Calcium/calmodulin-dependent kinase IV-responsive RNA element
ChIP	chromatin immunoprecipitation
CTD	C-terminal domain
DAPI	4', 6-diamino-2-phenylindole
DC	dendritic cell
EMSA	electrophoretic mobility shift assay
FBS	fetal bovine serum
GA	Golgi apparatus
GM-CSF	granulocyte macrophage colony-stimulating factor
H4R3me2s	histone 4 symmetric di-methyl R3
hnRNP	heterogeneous ribonucleoprotein particle
iCLIP	individual-nucleotide resolution UV-crosslinking and immunoprecipitation
IRES	internal ribosome entry site
LC-MS/MS	liquid chromatography-tandem mass spectrometry LC-MS/MS
ncRNA	non-coding RNA

NMD	nonsense-mediated mRNA decay
PAR-CLIP	Photoactivatable-Ribonucleoside-Enhanced crosslinking and immunoprecipitation
pre-mRNA	precursor messenger RNA
PRMT5	protein arginine methyl transferase 5
PSG	penicillin/streptomycin/glutamine
PTB	polypyrimidine tract-binding protein
Py	polypyrimidine
REPA	regulatory elements between the Py and 3′AG
RRM	RNA recognition motif
RT-PCR	reverse transcription-polymerase chain reaction
SELEX	systematic evolution of ligands by exponential enrichment
snoRNA	small nucleolar ribonucleic acid
snRNA	small nuclear ribonucleic acid
snRNP	small nuclear ribonucleic protein
SR	serine/arginine-rich
SS	splice site(s)
STREX	stress hormonal axis-regulated exon
Sxl	sex lethal
Tra	transformer
U2AF	U2 auxiliary factor

UTR	untranslated region

Thesis Chapters

Chapter I: This chapter is a review manuscript providing an overview of the regulation of 3′ splice site usage. 3′ splice site regulatory RNA elements and their underlying mechanisms are reviewed and categorized. The roles of chromatin modifications, transcription elongation and posttranslational modifications of splicing factors in splicing regulation are also discussed. This manuscript is intended to be submitted for publication in CMLS journal.

Chapter II: A brief overview of the thesis work is given in this chapter. It describes the theme of the thesis and specific points addressed in the thesis research with their overall implications.

Chapter III: This chapter constitutes the first part of the thesis research. It has been published in Molecular & Cellular Biology.

Chapter IV: This chapter is the second part of the thesis research. It has been revised after peer review and submitted for publication in BBA-Molecular Cell Research.

Chapter V: This chapter is a discussion of overall thesis research.

Chapter VI: Future directions are pointed out in this chapter.

Chapter VII: In this chapter overall conclusions of the thesis research are summarized.

CHAPTER I

Diverse Regulation of 3′ Splice Site Usage

Abstract

The regulation of splice site (SS) usage is important for alternative pre-mRNA splicing and thus proper expression of protein isoforms in cells; its disruption causes diseases. In recent years, an increasing number of novel regulatory elements have been found within or nearby the 3′SS in mammalian genes. The diverse elements recruit a repertoire of *trans*-acting factors or form secondary structures to regulate 3′SS usage, mostly at the early steps of spliceosome assembly. Their mechanisms of action mainly include: (1) competition between the factors for RNA elements, (2) steric hindrance between the factors, (3) direct interaction between the factors, (4) competition between two splice sites, or (5) local RNA secondary structures or longer range loops, according to the mode of protein/RNA interactions. Beyond the 3′SS, chromatin remodeling/transcription, posttranslational modifications of *trans*-acting factors and upstream signaling provide further layers of regulation. Evolutionarily, some of the 3′SS elements seem to have emerged in mammalian ancestors. Moreover, other possibilities of regulation such as that by non-coding RNA remain to be explored. It is thus likely that there are more diverse elements/factors and mechanisms that influence the choice of an intron end. The diverse regulation likely contributes to a more complex but refined transcriptome and proteome in mammals.

1. Introduction

Alternative pre-mRNA splicing allows the generation of up to thousands of variants from a gene, greatly contributing to the proteomic diversity of metazoans [1, 2]. In humans, about 92-95% of multi-exon genes undergo alternative splicing [3, 4], with functional consequences ranging from altered mRNA stability and fine-tuning of protein functions to antagonistic effects. Aberrant splicing, resembling alternative splicing but in abnormal conditions, causes diseases [5, 6]. Therefore, it is important to understand how the splice sites (SS) are alternatively or aberrantly used in cell physiology or diseases.

In mammals, the two splice sites at the start and end (5′ and 3′) of introns have distinctly different arrangements of particular motifs. The 5′SS is mostly comprised of a GU dinucleotide within a short consensus sequence. In contrast, the 3′SS consists of tripartite consensus motifs: the branch point (BP), polypyrimidine tract (Py) and 3′ AG dinucleotide (Fig. 1, upper diagram). The three motifs are differently conserved, required or spaced among different metazoan species *S. cerevisae* [7], *S. pombe* [8], *C. elegans* [9-11], and mammals [12], as summarized by Hollins and colleagues [9]. Particularly, the mammalian tripartite motifs have more permissive in-between spaces and different requirements for the 3′ AG [12, 13], likely providing more targets for splicing regulation than the 5′SS, as well as more than the 3′SS in other species.

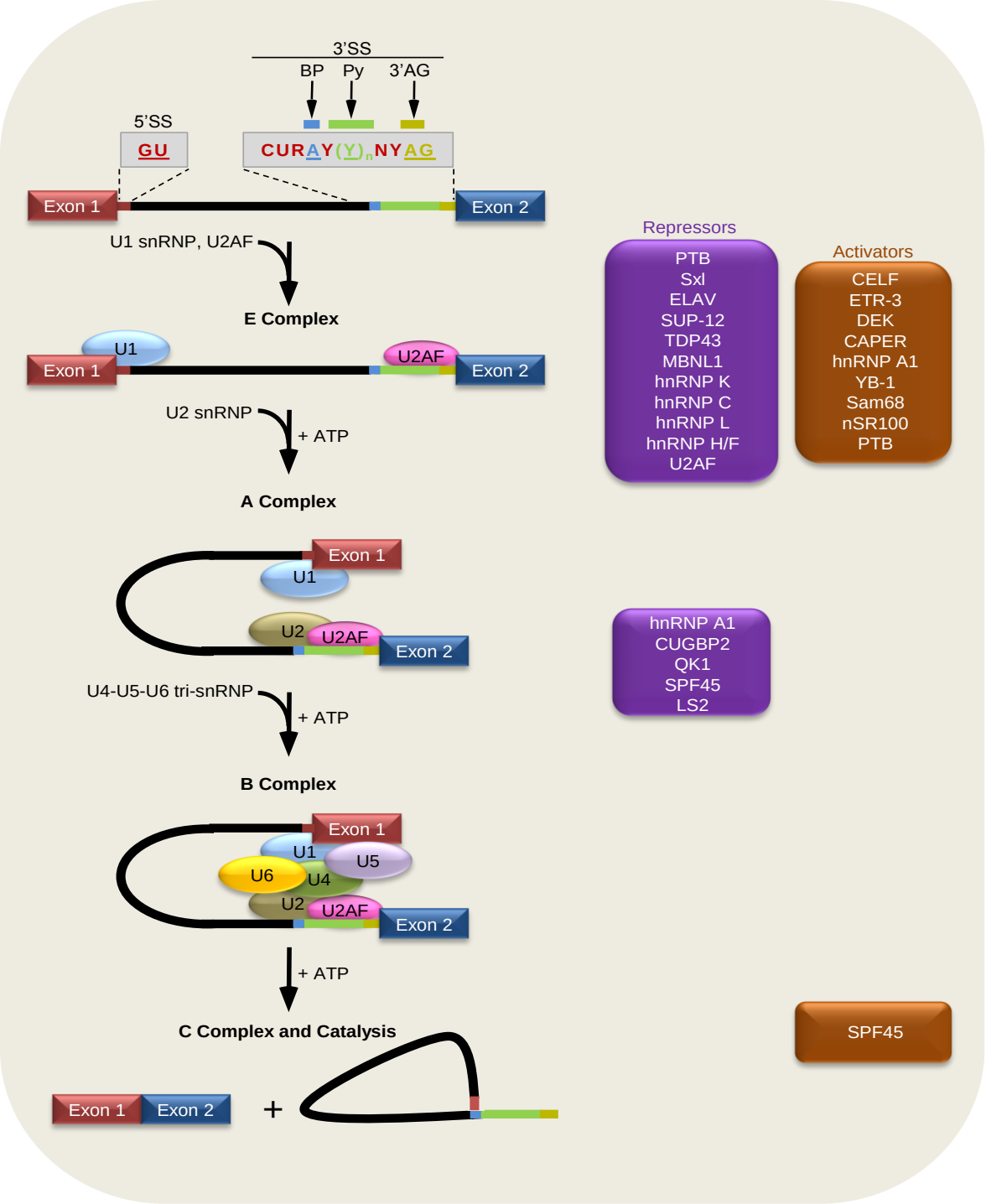


Figure 1. Diagram of pre-mRNA splicing steps and *trans*-acting 3′SS regulators. Shown at the top are splice sites motifs shown in different colors with consensus sequence and represented in the intron with respective colors. Below is the stepwise assembly of constitutive splicing factors (ovals) for excision of intron between exon 1 (red) and exon 2 (blue). Boxes on the right list *trans*-acting alternative splicing factors that regulate corresponding steps of spliceosome assembly for exon 2 splicing.

The splice sites are recognized sequentially by a set of small nuclear ribonucleoproteins (snRNPs) during spliceosome assembly on pre-mRNA introns. This process has been reviewed by Will and Luhrmann [14], to which and references therein readers are referred for more details. Briefly, after the 5′SS is bound by the U1 snRNP, the 3′SS Py and 3′ AG are bound by the U2 auxiliary factor (U2AF) heterodimer to form an E (early) complex (Fig. 1). Then the BP is base-paired with the U2 small nuclear ribonucleic acid (snRNA) of the U2 snRNP to form the A complex in an ATP-dependent manner. Subsequent recruitment of U4-U5-U6 tri-snRNP and complex rearrangements form the catalytic C complex where the intron is removed and exons are joined through two transesterification steps. In total, a spliceosome has over 170 proteins that assemble on pre-mRNA introns in *in vitro* systems using HeLa nuclear extract. It should be noted that, in the above systems relatively short introns have been used for intron definition and spliceosome formation; for longer ones, exon definition seems to be a more reasonable approach for recognition of splice sites and splicing [15].

The spliceosome assembly or splicing can be affected by a wide variety of *cis*-acting regulatory RNA elements called splicing enhancers or silencers [16, 17]. Many elements overlap with or are in-between the consensus motifs of splice sites though others are away from the splice sites. The elements often recruit *trans*-acting splicing factors and/or form RNA secondary structures to regulate splicing [16-18]. In addition to the commonly known regulatory factors, more and more so-called “constitutive splicing factors”, which were identified mostly from the HeLa nuclear extract system and model pre-mRNA transcripts, have also been demonstrated to regulate alternative splicing, for example the U2AF65 [19, 20], SmB/B’ [21] or U1C [22]. The list is probably going to extend to more factors in the future. Thus, before a complete survey of

these factors is available, here we will still follow the tradition to call them regulatory or constitutive factors respectively.

Recent studies on the regulation of 3′SS usage have uncovered a diverse group of elements and corresponding factors previously unidentified in this region and suggested novel mechanisms beyond the earlier competition model. Here we review the recently characterized 3′SS RNA elements, their binding factors and try to categorize their mechanisms of action into several models according to their common features. Other layers of splicing regulation including chromatin modification and rate of transcription elongation, as well as posttranslational modification of splicing factors and upstream signaling are also discussed regarding their effects on 3′SS usage. Focus is mainly on progresses in the last decade in mammalian systems.

2. Regulation of 3′ SS usage by RNA elements/splicing factors

Table 1 lists representative 3′SS regulatory RNA elements (silencers or enhancers) and *trans*-acting factors in the control of alternative splicing. Their regulatory targets during spliceosome assembly are shown in Fig. 1. Here we describe the elements/factors according their mechanisms of actions.

Table 1. List of 3′SS RNA elements and protein factors that regulate 3′SS usage, grouped according to their target constitutive factors.

Target Constitutive Factor	Location	Regulatory Element	Trans-acting Factor	Regulated Exon	Gene	Mechanism of Regulation	Ref.
Silencers/Repressors							
U2AF	Py	UCUU and UCUCU	PTB	6B	*β-tropomyosin*	i	[23]
		UCUU	PTB	3	*α-tropomyosin*	i	[24-26]
		CUCUCU	PTB	N1	*C-src*	i	[27-29]
		UCUU	PTB	SM	*α-actinin*	i	[30, 31]
		CUCU like	PTB	24nt exon, EN and 5	*GABA receptor, Clathrin light chain B* and *NMDAR1*	i	[32]
		C/U-rich motifs	PTB	# Group	# Group	i*	[33]
		UUUUUGUUGUUU UUUUU	Sxl	2	*Tra*	i	[34, 35]
		8/10U	ELAV	Terminal	*Neuroglian*	i*	[36]
		UG repeats	SUP-12	2A	*ADF/Cofilin*	i*	[37]
		UG	TDP43	9	*CFTR*	i*	[38, 39]
		YGCU(U/G)Y	MBNL1	5	*cTNT*	i	[40, 41]
		UCCCU	hnRNP K	5a, 6	*Snap25, Runx1*	i	[42]
		UUUs	hnRNP C	10	*CD55*	i	[43]
	Between Py and 3′ AG	CA repeats	hnRNP L	V10	*CD44*	ii	[44]
		CaRRE	hnRNP L	STREX	*Slo gene*	ii	[45-47]
		CA repeats	hnRNPL	20	*TJPI*	ii	[48]
		GGG	hnRNP H/F	# Group	#Group	ii	[49]
U2	Near BP	UAGGG(A/U)	hnRNP A1	3	*Tat (HIV)*	ii	[50]
		UGUGU and GU	CUGBP2	N1	*NMDA R1*	ii	[51]
		UGU motifs	CUGBP2		*CUGBP2*	ii	[51]
		AUUAAC	QK1	12	*NUMB*	i	[52]
		ULM of SF1	SPF45	6	*Fas*	ii*	[53]
		G rich	LS2	Minigene	*Ftc* and *PEP* Constructs	ii*	[54]
U2AF of weaker 3′SS*	Upstream of 3′SS	UUUCUU, UUUUUC, UUUUCU	U2AF	#Group	#Group	iii	[20]
Constitutive factors e.g. U1, U2AF	Py	Stem loop	MBNL1	5	*cTNT*	iv	[40, 41]
		U runs	Sxl	3	*Sxl*	iv	[55-57]
		CUCUCU	PTB	N1	*C-src*	iv	[27-29]
		UUUs	hnRNP C	10	*CD55* (#Group)	iv*	[58]
Enhancers/Activators							
U2AF	Py	CUG and UG motifs	CELF	SM, NM	*Actinin*	I	[59]
		U/G motifs	ETR-3	5	*cTNT*	I	[60]
		U2AF interaction (Py)	DEK	Minigene	*IgM-AdML* Constructs	II	[61]
		Py	CAPER	6	*VEGF*	II*	[62, 63]
		Py with 3′ CG or AG	hnRNP A1	#Group	#Group	II	[64]
		CAUC	YB-1	V5	*CD44*	II	[65]
	Near BP	AUUAAA	Sam68	V5	*CD44*	II	[66]
	Between Py and 3′ AG	UGC	nSR100	#Group	#Group	II	[33]
U2AF of competing 3′SS*	Py	CU rich	PTB	#Group	#Group	III	[67]
Constitutive factors e.g. U1, U2AF	Near 3′SS	Docking/Selector	hrp36*	Minigene	*Dscam* Constructs	IV	[68-70]
		SR binding	U1, U2AF35	Minigene	*dsx* Constructs	IV	[71-74]
		Artificial binding sites	hnRNP A/B or hnRNP F/H	Minigene	*hnRNPA1* Constructs	IV	[75-78]

Notes: *: predicted; #: a group of regulated exons as detailed in the text. For mechanisms of silencing/repression, i: competition, ii: steric hindrance, iii: competing 3′SS, or iv: local RNA secondary structure/looping out. For enhancement/activation, I: competition with a repressor for the same or overlapping RNA element, II: interaction with U2AF, III: competing 3′SS, or IV: local RNA secondary structure or longer range RNA loops.

2.1. Repression of 3′SS usage

There is a diverse group of silencers and repressors, but they mostly interfere with the early spliceosome factors (Fig. 2).

2.1.1. Inhibition of U2AF binding to the Py

Inhibition of U2AF65 binding was observed upon repressor binding to RNA silencers overlapping with the Py (Fig. 2A). Well-studied early examples include the inhibition of exon 6B of beta-tropomyosin [23, 79], exon 3 of alpha-tropomyosin [25, 80], N1 exon of Rouse sarcoma oncogene (*C-src*) [27-29], SM exon of alpha-actinin [31], and gamma amino-butyric acid (*GABA*) receptor, Clathrin light chain B and N-mehthyl-D-aspartate receptor subunit NR1 (*NMDAR1*) exons [32], by polypyrimidine tract-binding protein (PTB) binding to elements such as CUCU to compete with U2AF65 binding to the Py in mammalian systems. A more recent example of PTB-mediated repression of U2AF65 binding is its inhibition of neuron-enriched exons in non-neuronal cells [33]. A similar mechanism represses the male-specific alternative 3′SS of drosophila transformer (*Tra*) exon 2 by sex lethal (sxl) [34, 35].

Other elements include the U-rich and UG repeat elements that are bound by embryonic lethal abnormal visual system protein (ELAV) and SUP-12, respectively, to repress splicing of the *neuroglian* 3′ terminal exon and *ADF/Cofilin* exon 2A [36, 37], or UG motifs by TAR DNA-binding protein 43 (TDP43) to repress cystic fibrosis transmembrane conductance regulator (*CFTR*) exon 9 [38, 39]. These probably target U2AF65 as well. A similar YGCU(U/G)Y motif also inhibits U2AF65 when bound by muscleblind-like 1 (MBNL1), to repress the splicing of cardiac troponin T (*cTNT*)exon 5 [40, 41]. In this case, it also involves a mutually exclusive

secondary structure encompassing the Py, to which U2AF65 binding is competed by MBNL1 [41].

A recently reported regulatory element within the Py is the UCCCU motif that is bound by heterogeneous ribonucleoprotein particle K (hnRNP K) to repress the exon 5a of neural SNARE complex component synaptosomal-associated protein 25 (*Snap25*) and exon 6 of runt-related transcription factor 1 (*Runx1*) [42]. Mutation of the hnRNP K consensus motifs abolished its binding and addition of hnRNP K reduces U2AF65 binding to the 3′SS in a dosage-dependent way in UV crosslinking assays. The hnRNP K target motif appears to be enriched in the 3′SS of a group of neuronal alternative exons, implying a wider impact of this regulation [42].

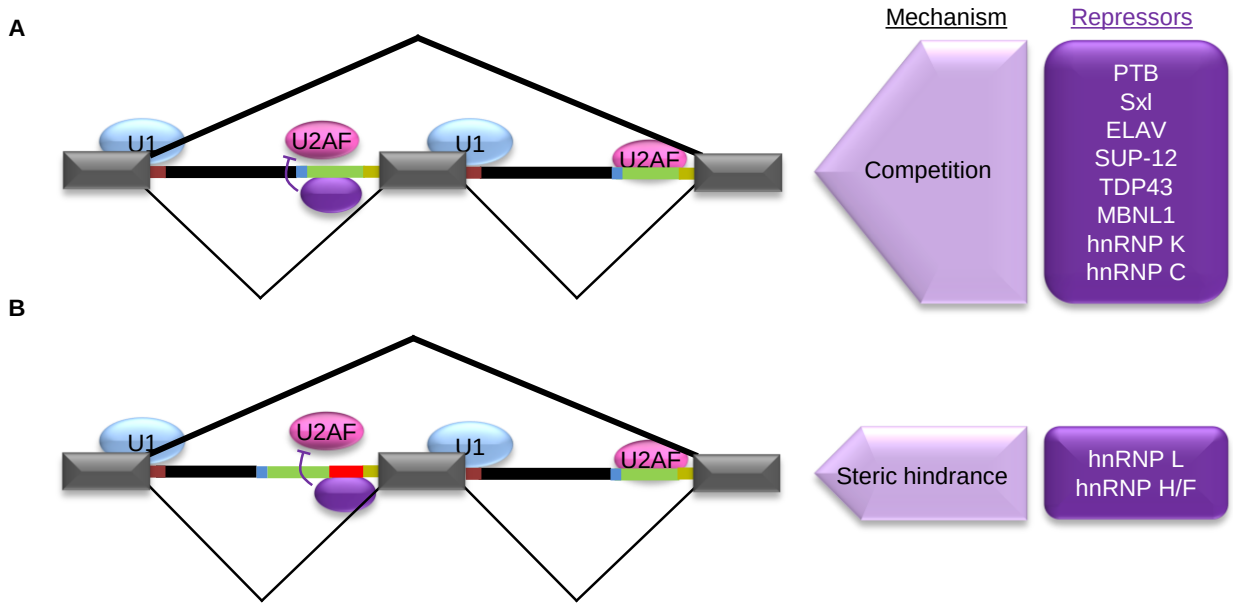
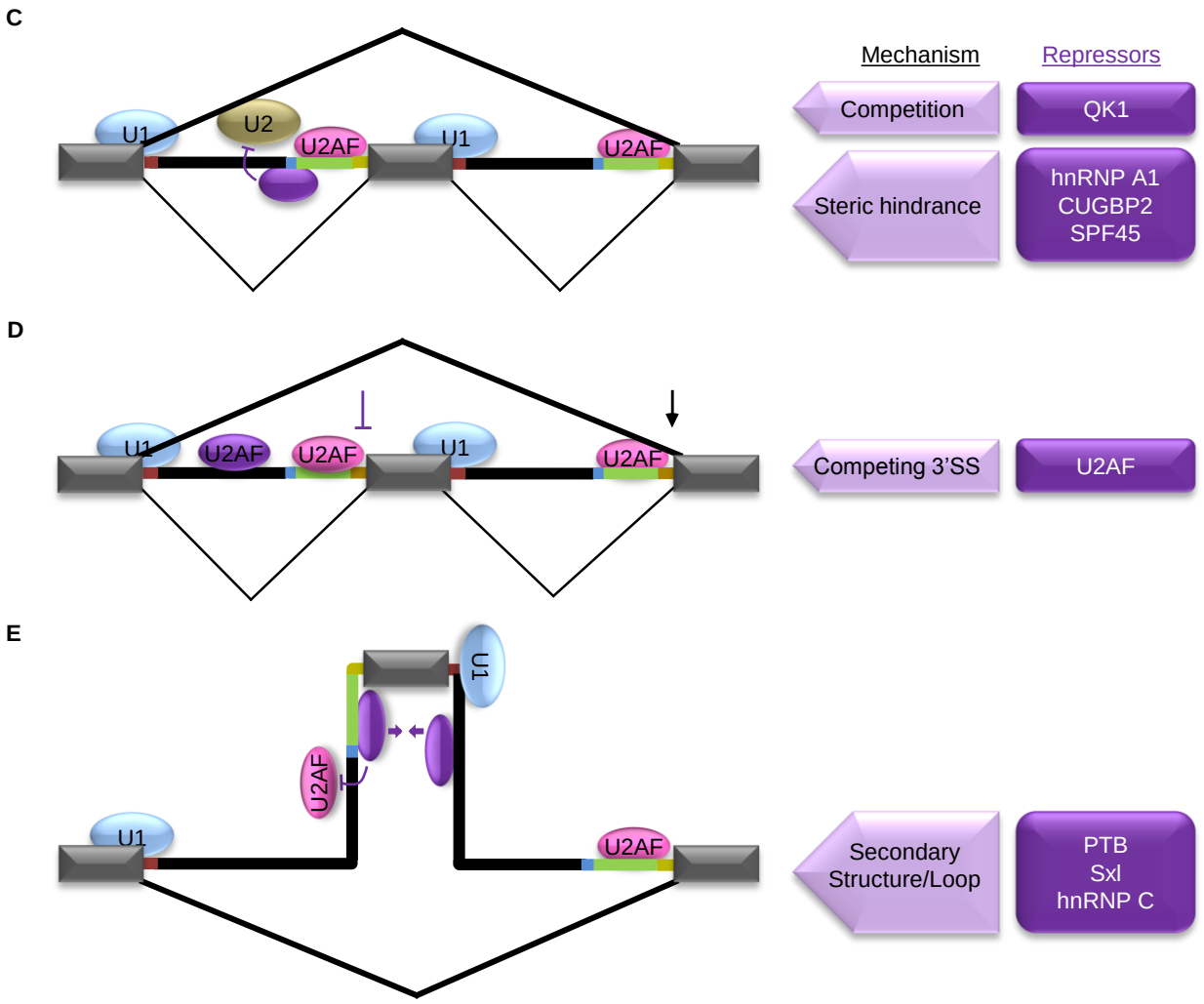


Figure 2. Mechanisms of splicing repression at 3′SS. The regulated exon shown in the middle is flanked by introns and exons. Splicing patterns in bold lines indicate the outcome of regulation. Purple oval represents one of the repressors listed in purple boxes on the right side that are involved in the regulation through respective mechanisms indicated on the left side. **A,** Splicing repression through *cis*-acting elements within Py. The location of a major class of silencers within 3′SS overlaps with Py. These elements recruit *trans*-acting factors to inhibit U2AF binding to Py through **competition (i)**. **B,** Splicing repression through *cis*-acting element insertions between Py and 3′ AG. A group of 3′SS silencers is uniquely positioned between Py and 3′ AG (red line in intron). Such elements also inhibit U2AF mainly through **steric hindrance (ii)** caused by specific binding of repressors.



C, Splicing repression through *cis*-acting elements near BP. Some 3′SS silencers are located near or in overlap with BP. The binding of *trans*-acting factors to these elements results in inhibition of U2 snRNP interaction with BP either through **competition (i)** or **steric hindrance (ii)**. **D,** Inhibition of weakened 3′SS by competition of downstream stronger 3′SS. The weakening of upstream 3′SS by intronic binding of U2AF (purple oval) results in inhibition by **competing 3′SS (iii)**. **E,** Splicing repression through '**RNA secondary structures or loops'**. In some cases Py binding *trans*-acting alternative splicing factors interact cooperatively with another RNA binding protein bound to downstream intron (purple ovals and arrows). This interaction results in **looping out (iv)** of the regulated exon and interferes with U1 and U2AF binding to its splice sites leading to inhibition of exon definition. RNA secondary structures also trap 3′SS motifs in a similar way but without proteins to repress splicing.

Another recent study has found that hnRNP C binds a U-rich element within Py, to compete with U2AF65 causing complement decay-accelerating factor 55 (*CD55*) exon 10 skipping [43, 58]. The binding of U2AF65 to Py and exon inclusion increased upon hnRNP C knockdown as shown in individual-nucleotide resolution UV-crosslinking and immunoprecipitation (iCLIP) assays and RNA sequencing analysis. Direct competition of U2AF65 binding by hnRNP C was observed in UV cross-linking assays using U_{10} RNA oligonucleotides. This hnRNP C competition with U2AF65 similarly inhibits cryptic 3′ splice sites to avoid the exonization of Alu elements [43].

Besides the Py, a second location of 3′SS elements and their binding factors that interfere with U2AF65 binding is between the Py and 3′ AG (Fig. 2B). We have called these elements REPA (regulatory elements between the Py and 3′AG) [44-49]. These include CA repeats and purine-rich elements particularly G-tracts.

CA repeat elements at this location are bound by hnRNP L to inhibit U2AF65 binding to upstream Py and splicing of a group of exons [44-48, 81]. The binding of hnRNP L and inhibition of U2AF65 was demonstrated in UV cross-linking/immunoprecipitation assays [44, 45, 48, 81]. Moreover, hnRNP L loss of function/rescue assays confirmed its role in the regulation of stress hormonal axis-regulated exon (STREX) [45].

One may argue that the boundary of the CA repeats actually could not be clearly distinguished from the Py, but the recently identified G tracts between the Py and 3′AG in a group of genes should have defined the REPA elements without any doubt [49]. G tracts are recognized by the quasi RNA recognition motifs (RRMs) of hnRNP H/F [82-84]. REPA G tract elements have been identified in the 3′SS of a large group of human genes. Moreover, hnRNP

H/F knockdown increased the endogenous exon inclusion for tested REPA G tract host genes [49]. Therefore, unlike the enhancer effect of G tracts at other intronic locations [85-87], the uniquely positioned REPA G tracts between the Py and 3′ AG are splicing silencers. The presence of similarly positioned G tract elements in over a thousand of human 3′SS suggests a wider impact of these elements in diversifying the protein products and functions [49].

The 3′SS inhibition by the REPA elements could be better explained by steric hindrance between the *trans*-acting repressor and U2AF65/35 heterodimer, which could be co-purified stably from nuclear extracts [88]. The inhibition could be due to the space constraint imposed by the element insertion between the Py and 3′ AG, which are often bound by the heterodimer for splicing of AG-dependent 3′SS [89-92]. The insertion causing the steric hindrance is opposite to the shortened distances between splice site motifs, such as BP and 3′AG or 5′SS in repressing the 3′SS of the (mutually exclusive) SM exon of alpha-actinin [26, 31, 93].

It should be noted that a REPA might not necessarily always inhibit splicing. In the case of a UGC motif at the REPA location [33], it is bound by nSR100 in neurons to directly interact with U2AF65 and enhance its binding, which will be detailed in the enhancement section. Therefore, whether a REPA causes steric hindrance between the splicing regulators might depend on the particular factor. Moreover, since such *trans*-acting factors can also receive upstream signaling [45], the effect by even the same factor could be changed depending on the cell environment.

Taken together, these recent cases demonstrate that: (1) the py-rich elements and bound factors could be diverse, and (2) location of the elements can also be out of the Py: “inserted”

between the Py and 3′ AG (REPA) to recruit more *trans*-acting factors, to interfere with U2AF65 binding.

2.1.2. Inhibition of U2 snRNP interaction with the BP

The interaction of U2 snRNP with the BP is a critical step in spliceosome assembly. It is regulated by a number of *trans*-acting repressors binding to RNA motifs near or directly to the BP (Fig. 2C). An earlier example showed that the U2 snRNP could be interfered sterically by the serine/arginine-rich (SR) protein ASF/SF2 (SRSF1) binding several nucleotides upstream of the BP [94]. In the splicing repression of HIV Tat exon 3, hnRNP A1 binding to the -26nt does not interfere with U2AF binding but inhibits the subsequent U2 snRNP interaction with the BP [50]. The N1 exon of NMDA receptor subunit *NMDAR1* and exon 6 of CELF RNA binding protein family member CUG triplet repeat RNA-binding protein 2 (*CUGBP2*) are also regulated in this manner [51, 95-97]. *NMDAR1* pre-mRNA contains UGUGU and GU motifs around the predicted BP of N1 exon upstream 3′SS and UGU motifs are similarly located for *CUGBP2* exon 6 [51]. Such UG-rich motifs are one of the two groups of sequences recognized by CUGBP2 to regulate alternative splicing, the other group being CUG triplet repeats [96-105]. CUGBP2 bound to NMDAR1 and its own 3′SS motifs in chemical modification footprinting assays [51], and its inhibition of lariat formation was confirmed by using recombinant CUGBP2 or mutagenesis of its binding sites under *in vitro* splicing conditions [51]. Nuclear extract complementation and mutagenesis analyses revealed that CUGBP2 binding interferes with U2 snRNP association [51]. These studies support the model of branch site-perimeter-binding and interference with U2 snRNP for exon skipping.

A more recent example is QK1 (quaking homolog, KH domain) competition with the BP binding constitutive factor SF1 (splicing factor 1) to repress the exon 12 of *NUMB* [52], a Notch pathway regulator [106-109]. QK1 is a member of STAR (signal transduction and activation of RNA) family of proteins and binds ACUAAY motif [110, 111], which is similar to the consensus ACUNAC of SF1 [112]. Both QK1 and SF1 bind the AUUAAC motif immediately upstream of Py in *NUMB* and QK1 competes with SF1 in UV crosslinking and mutagenesis assays [52].

In case of *FAS* (TNF receptor superfamily member 6) exon 6, the splicing repression requires the interactions between U2AF-homology domain (UHM) of SPF45 (splicing factor 45) with UHM-ligand motifs (ULMs) of constitutive splicing factors including SF1 and U2AF65 [53]. SPF45 crosslinked to RNA in a U2AF dependent manner [113]. The interactions of SPF45 with SF1 and U2AF likely compete with UHM-ULM interactions required for *FAS* exon 6 splicing, possibly by inhibiting the interaction of U2AF65 with SF1 and subsequent U2 snRNP binding to BP [53].

There are also G rich motifs that are enriched in the upstream 60nt of 3′SS to repress the exon usage of some testis-specific genes in Drosophila [54]. These motifs are bound by the LS2 protein, a homologue of the U2AF larger subunit emerged through gene duplication but has diverged in binding specificity for G-rich motifs, as determined by systematic evolution of ligands by exponential enrichment (SELEX) and electrophoretic mobility shift assays (EMSA) [54]. This change in binding specificity switched the protein function to a repressor. The motifs appear to be around the usual location of BP but whether it represses U2 snRNP remains unknown.

Together, the current evidence for the splicing repression of a number of these exons indicates that regulatory motifs overlapping or near BP can specifically inhibit the interaction of SF1 or U2 snRNP with the essential 3′SS motif. The observations in the above examples strongly support this model, particularly *NUMB* exon 12 which is repressed by a direct competition between QK1 and SF1 to inhibit BP recognition.

2.1.3. Competition between 3′ splice sites: inhibition of weaker 3′SS by enhancement of downstream 3′SS

Competition between 3′SSs of two consecutive exons has been described recently [20]. In this scenario, the downstream 3′SS is enhanced by the binding of U2AF65, thereby weakening the upstream 3′SS and its immediate downstream exon. This mechanism is proposed for a group of exons for which U2AF binding events in the intronic regions upstream of 3′SS were identified by CLIP-seq experiments. U2AF binding to these sites inhibits the usage the 3′SS of the upstream intron (Fig. 2D).

3′SS scanning and competition have been proposed for more than two decades but on the basis of a common BP between two competing 3′ AGs [93]. The above-mentioned cases do not seem to share the same BP between the two sites, therefore likely regulated through a different mechanism. For instance, it could be through simple competition of the 3′SS for limited local 3′SS factors, scanning by the spliceosome for stronger 3′SS/exons or crosstalk between the competing sites/exons. In this regard, a very interesting study is the biochemical characterization of the transition from an exon to intron definition complex during the PTB regulation of C-src N1 exon [29, 114]. Whether such transition is also shifted toward intron definition and middle

exon skipping by the stronger downstream 3′SS in these U2AF65-regulated transcripts would be interesting to know.

2.1.4. Inhibition of 3′ splice site usage by local RNA secondary structure or longer range loops

Another possible mechanism of splicing repression is trapping a splicing motif inside the double stranded region of a stem loop or 'looping out' an entire exon through interaction between complementary regions of pre-mRNA, RNA binding proteins or different domains of a protein bound to the flanking introns (Fig. 2E).

The effect of pre-mRNA secondary structure on splicing was first observed three decades ago followed by more examples later on [115-118]. Large scale analyses indicate that stable secondary structures based on free energy prediction are enriched at alternative 3′ or 5′ splice sites and are conserved across species [119, 120]. Complementary mutagenesis analyses of particular sites indicate that RNA stem loops can trap motifs including the Py or 3′ AG, or loop out the target exon, thereby inhibiting 3′SS usage [41, 116, 121, 122]. Regulation of *cTNT* exon 5 is a clear example of the formation of a splicing inhibitory stem loop in 3′SS [41].

For those secondary structures/loops bridged by RNA binding proteins, two well-studied examples involving 3′SS elements are in the repression of Drosophila *Sxl* exon 3 and *C-src* exon N1 [29, 55-57].

In the case of *Sxl* exon 3, an autoregulation through binding of SXL in the upstream 3′SS and downstream intron mediates the skipping. SXL exists as a monomer in solution but it can form higher order complexes through RNA binding-dependent multimerization [55]. In

regulation of *Sxl* exon 3, it binds to the two sites and interacts cooperatively through N terminus to ensure the splicing repression [55-57]. The amino terminus of SXL is essential for cooperative protein-protein interaction and plays a vital role in splicing regulation. Therefore, the multimerization of SXL suggests the formation of protein-protein bridges in the looping out model.

For *C-src* exon N1, PTB binds to both CU elements within the upstream 3′SS and in the downstream intron to repress exon inclusion, likely by looping out the N1 exon [28]. Whether PTB binds as a monomer or multimer remains unknown though PTB is a monomer in experimental solutions [123]. A model with different RRM domains of PTB contacting different target sites within the pre-mRNA has also been proposed [124]. However, they do self-interact in yeast two hybrid assays [125]. Therefore, it remains undetermined how the loop is formed if it does.

A more recent study proposed the looping out mechanism for the regulation of *CD55* exon 10 by hnRNP C binding to upstream Py as described above [43], and downstream intron [58]. The binding of hnRNP C at these positions with 165 nucleotides in-between was determined by iCLIP. The spacer sequence between the two binding sites of hnRNP C flanking *CD55* alternative exon is proposed to allow the cooperative particle formation and splicing repression through a looping out mechanism.

Although observations in the above cases suggest the protein-RNA or protein-protein interactions between the splicing factors bound to the flanking introns but direct physical evidence for the “looping out” of an exon in the pre-mRNA remains to be seen. Moreover, for the secondary structures formed by complementary RNA regions, how the structures are

regulated in cells to determine the usage or skipping of a 3′SS is unclear. It could be imagined that *trans*-acting factors (protein or RNA) might be involved to receive the upstream signals.

From the above examples, it is clear that the elements and factors now appear to be quite diverse in sequence and location in the repression of 3′SS usage. Their mechanisms of action mainly involve RNA-RNA, RNA-protein or protein-protein interactions. The binding of *trans*-acting factors inhibits the recruitment of constitutive splicing factors (Fig. 2 and Table 1). Such inhibition could possibly be caused by protein competition (i), steric hindrance (ii), 3′SS competition (iii) or RNA secondary structure/looping out (iv), for which the first one is the simplest and best characterized.

2.2. Enhancement of 3′SS usage

The enhancement of 3′SS usage mainly involves the inhibition of repressor binding, binding of *trans*-acting activator to enhancers or direct interaction with U2AF, to promote the early steps of spliceosome assembly (Fig.s 1 & 3, and Table 1).

2.2.1. Inhibition of repressor binding

The binding of a *trans*-acting factor to enhancer element within Py to interfere with the binding of a *trans*-acting repressor is a well-known mechanism of splicing activation at 3′SS (Fig. 3A). The CUG and UG enhancer motifs in the upstream 3′SS of *actinin* exons SM, NM and *cTNT* exon 5 are such examples [59, 60]. In this regulation, these motifs are bound by CELF proteins leading to competitive displacement of splicing repressor PTB that binds to adjacent sites, thereby enhancing the usage of the downstream exon. A recent example involves the displacement of hnRNP L by hnRNP LL due to a genetic mutation in the *CHRNA1* (cholinergic

receptor, nicotinic, alpha polypeptide 1) gene [126]. The latter, unlike hnRNP L, does not interact with PTB, thereby allowing U2AF65 interaction with a cryptic 3′SS and inclusion of an intronic sequence.

2.2.2. Interaction with U2AF by activators

Other regulatory elements at different positions of 3′SS are recognized by RNA binding proteins that interact with constitutive splicing factors for splicing activation (Fig. 3B-D). A well-known example of such regulation is the splicing activation of mini-gene constructs of IgM and AdML pre-mRNAs through interaction of U2AF subunits with DEK (DEK proto-oncogene) [61]. The mini-gene activation is achieved by the interaction of a chromatin- and RNA-associating protein DEK with U2AF35 (Fig. 3B). However, the binding of DEK to the pre-mRNA is not demonstrated. Moreover, a similar mechanism is likely involved in the splicing activation of vascular endothelial growth factor (*VEGF*) exons 6 and 7 by CAPER proteins that share sequence homology with U2AF65 [62, 63, 127]. Although the molecular details for this mechanism are not known, it was postulated that interaction of CAPER with EWS/FLI-1, which interacts with some 3′SS constitutive factors, is involved [63].

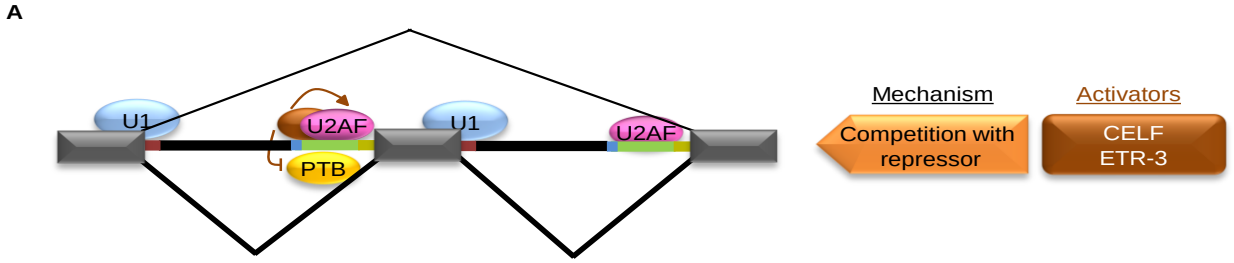
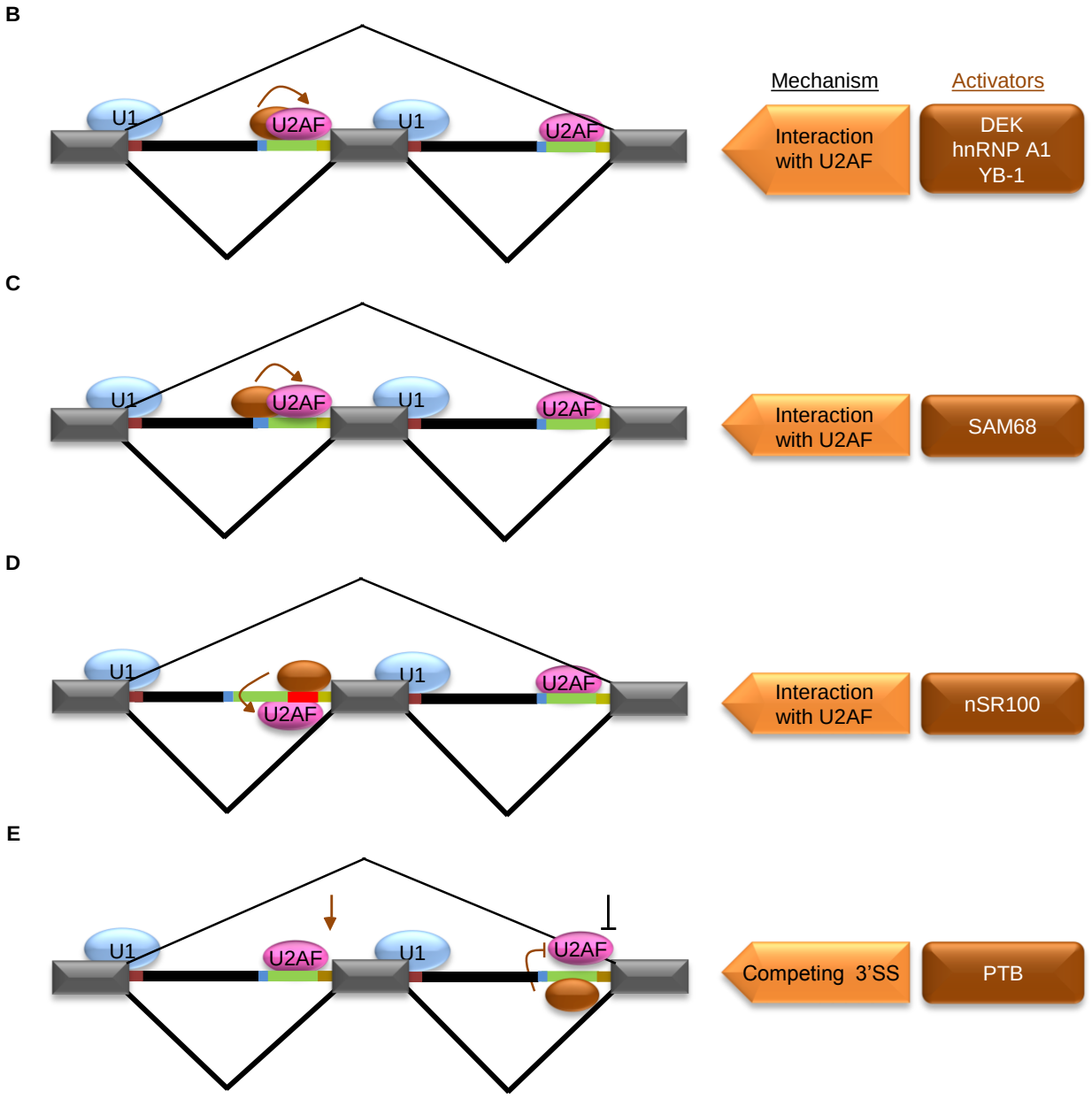


Figure 3. Mechanisms of splicing activation at 3′SS. Individual activators are shown as brown ovals and they are listed on the right side in brown boxes along with respective mechanisms for the regulation shown on the left side. **A,** Splicing activation through **i**nhibition of a repressor by enhancer motifs in Py. The binding of a *trans*-acting factor to an enhancer motif within Py results in **competition with a repressor (yellow oval/PTB) binding (I)** and facilitates U2AF binding to Py to promote 3′SS usage.



B-D, Splicing activation through cis-acting elements located at different positions within 3′SS. The locations of these elements are within Py **(B),** near BP **(C)** or between Py and 3′ AG (red line) **(D)**. A common feature of 3′SS regulation by these elements is their recognition by respective *trans*-acting factors for **interaction with U2AF (II)** to facilitate its binding to Py and 3′ AG leading to activation of splicing. **E,** Activation of 3′SS by repression of downstream 3′SS. The repression of a downstream 3′SS through binding of PTB to Py inhibiting U2AF results in activation of a **competing 3′SS (III)**.

A more recent example of similar regulation is the interaction between RNA bound hnRNP A1 with U2AF to proofread the splice site and promote its recognition [64]. HnRNP A1 is often observed as a splicing repressor but this study explained a novel role in proofreading of 3′SS. The pre-mRNAs with pyrimidine rich sequences followed by AG or CG were utilized to study the ability of constitutive splicing factor U2AF to discriminate between the two dinucleotides. HnRNP A1 formed a complex with U2AF heterodimer facilitating it to discriminate between these 3′SS (Fig. 3B). Moreover, upon knockdown of hnRNP A1 in cells, the binding of U2AF to U rich sequences of an intronless gene *c-jun* as well as to 5′UTRs (5′ untranslated region) of three other transcripts increased, which suggests a role of hnRNP A1 for preventing the binding of U2AF to U rich sequences beyond 3′SS in cells. However, whether hnRNPA1 has an enhancing effect on U2AF65 binding to the 3′SS of an endogenous exon in cells is not known.

A similar mechanism may also contribute to the activation of *CD44* (CD44 molecule) exon V5 splicing [65], in addition to a previously known model [66]. A recent study reported CAUC splicing enhancer motifs within the weak Py of V5 upstream 3′SS that are bound by the Y box-binding protein-1 (YB-1) to recruit U2AF65 and U2AF35 through protein-protein interaction [65] (Fig. 3B). The YB-1 interacts through its C-terminal domain with the RS domain of the U2AFs and enhances the recruitment of U2AF65 to Py, which the U2AF65 alone does not bind [65]. Another RNA motif (AUAAA) immediately upstream of Py is bound by Sam68 (Src-associated in mitosis, 68 kDa protein) and enhances the 3′SS usage by facilitating U2AF65 binding to Py [66] (Fig. 3C).

The recently identified UGC motifs between Py and 3′ AG also promote the inclusion of a group of exons specifically in neuronal cells [33]. nSR100 (neural-specific Ser/Arg repeat-related protein of 100 kDa) binds to these motifs and interacts with early spliceosome components including U2AF65 to promote exon inclusion (Fig. 3D). The nSR100 binding motifs were identified in the upstream introns of a group of regulated exons. The *in vivo* binding of nSR100 to these motifs was determined by photoactivatable-ribonucleside-enhanced crosslinking and immunoprecipitation (PAR-CLIP). Mutagenesis of mini-genes in splicing reporter assays confirmed their critical role in splicing activation. Furthermore, immunoaffinity purification coupled to mass spectrometry analysis showed that nSR100 interacts with a number of constitutive splicing factors including U2AF subunits that bind in a close proximity to nSR100 binding motifs. Moreover, higher nSR100 expression was able to restore exon inclusion after U2AF65 knockdown, which suggests a compensatory effect when U2AF65 is rate limiting. This provides an interesting example of tissue specific splicing activation through the binding of a *trans*-acting factor to motif uniquely positioned between Py and 3′AG, where locates the REPA elements.

SR proteins are known splicing enhancers and one example that involves the 3′SS region is the activation of *C-src* N1 exon by ASF/SF2 binding to an exonic enhancer in the 5′ half with two nucleotides overlapping with the 3′SS [128]. Another protein that binds at a similar position (3′ AG) of proximal 3′SS is SPF45 in regulating *Sxl* exon 3, which enhances the second catalytic step of splicing [113]. However, the details of molecular mechanisms underlying the regulation of these exons remain unknown.

2.2.3. Competition between 3′SS

In this case, weakening of a downstream site activates an upstream site (Fig. 3E). This was observed in the control of alternative splicing by the commonly known repressor PTB [67]. This mechanism was inferred from the genome-wide PTB binding map based on CLIP-seq that allowed the identification of a group of regulated exons in the context of PTB binding. A group of exons are regulated through this mechanism but the molecular details of this regulation remain unknown.

2.2.4. Enhancement of weaker 3′SS by local RNA secondary structure or longer range looping

Besides the direct involvement of 3′SS elements and *trans*-acting factors in its regulation, another mechanism is the enhancement of an otherwise weaker 3′SS by RNA secondary structures or looping, to bring motifs or splice sites closer, particularly in splicing of long introns [78].

The usage of a distal 3′SS could be facilitated by a hairpin structure that brings it closer to the branch point in an E1A intron [117], and more examples have been found recently in yeast introns [129]. A similar hairpin structure also blocks cryptic 3′ AG to facilitate correct 3′ AG usage [130]. Besides the 3′SS motifs, the 5′ and 3′ splice sites can also be brought closer by intra-molecular base pairing, for instance between two complementary regions near the 5′ splice site and branch point to aid the early spliceosome assembly in the yeast rp5l B pre-mRNA [131], as supported by both *in vitro* and *in vivo* evidence [131].

Studies on drosophila *Dscam* pre-mRNAs suggest a similar model of interaction between two intronic regions to bring splice sites closer in selecting a downstream 3′SS of mutually exclusive exons [69, 70] (Fig. 4A). The selective enhancement of 3′SS usage in the exon 6 cluster of *Dscam* requires base pairing interaction between one of the selector sequences near 3′SS and a docking site [69, 70] (Fig. 4A). However, how a specific 3′SS among that of the other mutually exclusive exons is chosen in a particular cell or cell state remains unknown. It likely involves *trans*-acting factors such as hrp36, an hnRNP A1 homologue [68].

Besides RNA intramolecular base-pairing, protein-protein interaction may also bring the 5′ and 3′ splice sites closer (Fig. 4B). Proteins such as SRSF1 (ASF/SF2) and SRSF2 (SC35) specifically interact with both the U170K, a component of the 5′SS-binding U1 snRNP, and the 3′ AG-binding U2AF35 [71, 72]. Moreover, SC35 interacts with Tra and Tra2 which also shows interaction with U2AF35 and suggested to facilitate interaction between splice sites in regulation of *dsx* sex specific splicing [71-74]. Studies with hnRNP A/B and F/H binding sites near introns ends indicate that their protein-RNA interactions could bring the splice sites closer as well [75-78] (Fig. 4C). Insertion of the binding sites near intron ends significantly increased the splicing of long introns suggesting looping out of a portion of pre-mRNA to bring two splice sites in close proximity. This model was further explained *in vivo* by hybridizing the ends of introns with oligonucleotides containing hnRNP A/B binding sites within their non-hybridizing tails. Computational analysis of hnRNP A/B and F/H binding sites at the end of introns supports this model as well [78].

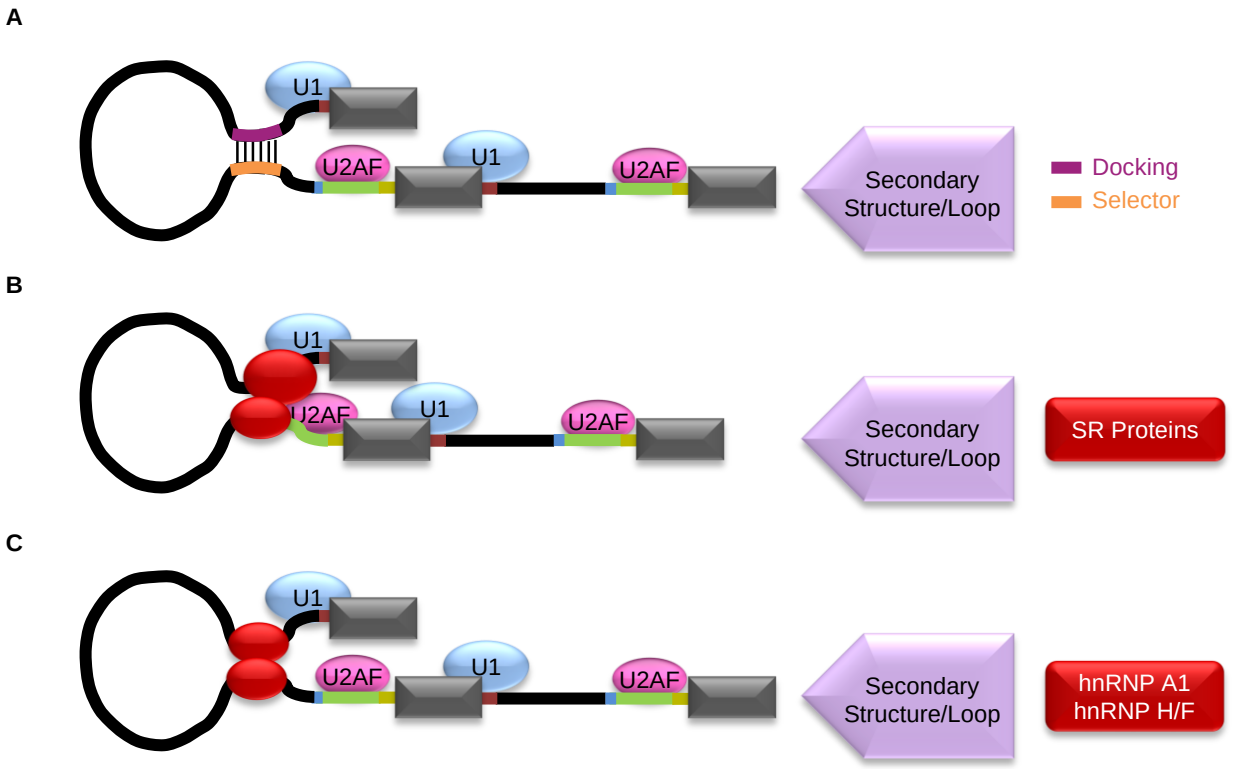


Figure 4. Mechanisms of weaker 3′SS activation through RNA secondary structures/loops. **A,** Base-pairing interaction between intronic RNA motifs. The activation of a 3′SS through looping of pre-mRNA can also result from base pairing of an intronic motif (purple line) with a second motif located near regulated splice site (orange line). These motifs are referred to as docking or selector sites as indicated on the right side. The formation of such **secondary structures or loops (IV)** promotes 3′SS usage. **B,** Interaction between a complex of splicing factors and 3′SS spliceosome components. In this mechanism a protein complex (red oval) binds to an intron upstream of regulated exon and interacts with spliceosome proteins of 3′SS to promote its usage. SR proteins shown in red box on the right side form such known complexes. **C,** Cooperative interaction of a splicing factor. This mechanism of pre-mRNA looping involves the binding of a splicing factor (red oval) near intron ends. Examples of such splicing factors are shown in red box on the right side. The cooperative interaction between the factors bound near intron ends creates a loop of spanning intronic sequence that brings its splice sites along with their spliceosomal components in close proximity to activate 3′SS.

Taken together, the current evidence on 3′SS activation by different elements/factors suggests that the mechanisms mainly involve the inhibition of a repressor (Fig. 3A), protein-protein interactions between an RNA binding protein and U2AF (Fig. 3B-D), activation of a 3′SS by weakening of competing 3′SS (Fig. 3E), or local RNA secondary structures, longer range

RNA loops through protein-protein bridge or RNA base pairing to bring closer two splice motifs or sites to promote early spliceosome assembly (Fig. 4). However, many questions still remain, for instance, how does the activator protein help 3′SS recognition by interacting with the U2AF? Does it need U2AF65/35 binding to the RNA in such a tight space or simply serve to bring the U2AFs close, like that by YB-1[65], to help SF1 binding to the BP? Moreover, how do two splice sites compete and is it limited only to nearby 3′SSs?

Overall, according to the mode of protein/RNA interaction in the above examples, the mechanisms of 3′SS regulation (repression or activation) can be divided into: (1) protein competition for a common RNA element/sequence (repression i; enhancement I), (2) steric hindrance between protein factors (repression ii), (3) direct interaction between activators and U2AF (enhancement II), (4) competition between splice sites (repression iii, enhancement III), and (5) local RNA secondary structures or RNA loops (repression iv; enhancement IV).

3. Regulation of 3′SS usage by chromatin configuration/transcription rate

Chromatin remodelling and the rate of transcription elongation also modulate exon usage with enhancing or repressing effect depending on the target exon [132-137]. However, the positional role of chromatin/nucleosome modifications and remodelling with respect to the precise splice sites remains largely unknown. Studies based on computational models revealed that the nucleosome occupancy at approximately 25 bp upstream of 3′ AG is lowest as compared to other intronic or exonic regions [138], a position generally corresponding to BP [138, 139].

SWI/SNF is a chromatin-remodelling complex that uses ATP to facilitate conformational changes for transcriptional regulation [140-142]. The ATPase activity lies in the Brahma (Brm) or Brahma-related protein (Brg1). Overexpression and knockdown assays of Brm strongly

suggest its role in splicing regulation [143]. Its most profound specific effect was observed for *CD44* exon V5. However, the chromatin remodeling activity of Brm does not seem to be required for this regulation as mutations or deletions of the ATPase domain did not substantially change splicing whereas deletion of its chromatin binding C-terminal part resulted in decreased exon inclusion. RNA immunoprecipitation assays of Brm revealed that it interacts with U1 and under conditions that favor the inclusion of *CD44* exon V5 it also interacts with U5 snRNA [143].

The splicing of *CD44* exon V5 is enhanced by Sam68 binding in the exon [144], as well as near BP as described above [66]. Co-immunoprecipitation assays revealed that Brm also associates with Sam68 and they cooperatively enhance splicing of this exon [143]. Moreover, ChIP (chromatin immunoprecipitation) walking using antibodies against phospho-CTD (C-terminal domain) of PolII (RNA polymerase II) revealed higher occupancy in the CD44 variant region including V5, which suggests a slower transcription rate for enhanced exon usage of CD44 V5 [143]. These data support a role for a factor of the chromatin remodelling complex in the regulation of splicing, likely involving its interaction with those of the 3′SS.

4. Role of posttranslational modifications and upstream signaling in 3′SS regulation

Besides the *cis*-acting RNA elements, *trans*-acting factors and chromatin/transcriptional machinery, posttranslational modifications of *trans*-acting splicing factors by cellular signaling pathways add another layer of regulation. Table 2 lists the posttranslational modifications of 3′SS *trans*-acting factors that influence splicing regulation. The effects on splicing regulation through these factors are described below.

Table 2. Posttranslational modifications of 3′SS factors

Splicing Factor	Modification	Residue	Effect on Splicing	Ref.
U2AF65	Hydroxylation	Lys-15, Lys-276	Activating	[145]
SF1	Phosphorylation	Ser-80, Ser 82	Activating/Repressing	[146-149]
hnRNP K	Phosphorylation	Ser/Thr (predicted)	Repressing	[42]
hnRNP L	Phosphorylation	Ser-513	Repressing	[45]
DEK	Phosphorylation	Ser-19, Ser-32	Activating	[61]
Sm proteins	Methylation	Arg rich motifs	Activating	[150]
SRSF2	Acetylation	Lys-52	Activating	[151, 152]
U2AF35	Ubiquitination	Within residues 43-146	Repressing	[153]
Prp3	Ubiquitination	Two Lys residues	Repressing	[154]
SF3b*	Spliceostatin binding	NA	Repressing	[155]

*:non-covalent binding. NA: not applicable.

4.1. Hydroxylation

U2AF65 undergoes posttranslational hydroxylation by the Fe(II) and 2-oxoglutarate–dependent dioxygenase jumonji domain-6 protein (JMJD6) in the presence of oxygen [145]. The hydroxylation of Lys15 and Lys276 of the endogenous U2AF65 was detected by liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis of HeLa cells. U2AF65 co-immunoprecipitates with JMJD6 in an RNase-sensitive manner. Knockdown of JMJD6 resulted in increased exon 19 skipping of endogenous *MGEA6* pre-mRNA and a similar effect was observed for a α-tropomyosin minigene construct comprised of exons 1, 3 and 4. This indirectly shows the role of U2AF65 hydroxylation in splicing activation of these exons as JMJD6 catalyzes it. The effect of hydroxylation on *MGEA6* splicing is further supported by treatment of cells with oxygenase inhibitor desferrioxamine (DFO), which showed a similar effect as JMJD6 knockdown. These data demonstrate a unique post-translational modification of a splicing factor in the presence of oxygen, likely related to splicing regulation under hypoxia conditions. However, a direct effect of U2AF65 hydroxylation on splicing, the underlying molecular

mechanism and the hydroxylation status of other splicing factors under similar conditions remain unknown.

4.2. Phosphorylation

Phosphorylation of 3′SS *trans*-acting factors such as SF1 also contributes to alternative splicing regulation [146-148]. The protein-protein interaction of SF1 and U2AF65 and their cooperative binding to 3′SS motifs is a critical step for further spliceosome assembly. *In vitro* kinase assays followed by mass spectrometry analysis revealed the phosphorylation of SF1 at serines 80 and 82 [148]. *In vivo* metabolic labelling of overexpressed SF1 or its mutants confirmed this phosphorylation. The phosphorylation of SF1 facilitates its ternary complex formation with U2AF65 and slightly enhances the RNA binding in electrophoretic mobility shift assays [146, 148], hence likely to promote the spliceosome assembly and splicing. Contrary to the activating effect of Ser80 and Ser82 phosphorylation, previous studies of SF1 Ser20 phosphorylation by protein kinase G (PKG) suggest a negative effect on protein-protein interaction as well as spliceosome assembly [147, 149]. Thus, phosphorylation of different residues of SF1 may have different effects on spliceosome assembly and interaction of SF1 with U2AF.

Two other examples of splicing factor phosphorylation are of hnRNPs K and L [42, 45]. The hnRNP K is phosphorylated in the PKA-regulated splicing of SNAP25 exon 5a. Its binding to the 3′SS Py is abrogated by pre-treatment with lambda protein phosphatase, suggesting an essential role of its phorsphorylation in its regulation of the 3′SS. The serines 284 and 353 of hnRNP K are phosphorylated by the MAP/ERK pathway [156], but their role in the splicing regulation is unknown. In case of hnRNP L, depolarization activated CaMKIV

(Calcium/calmodulin-dependent kinase IV) phosphorylates serine 513 to regulate splicing of STREX [45-47, 157]. The phosphorylation enhances its binding to the CaRRE1 (Calcium/calmodulin-dependent kinase IV-responsive RNA element) pre-mRNA and inhibition of U2AF65 [45].

The phosphorylation of DEK is required for intron removal [61]. This protein interacts with U2AF35 to form a complex that is disrupted upon phosphatase treatment. Mutagenesis analysis revealed that the complex formation is dependent on the phosphorylation of serines 19 and 32.

Other regulatory factors acting on the 3′SS are also phosphorylated but the phosphorylation effect on 3′SS usage has not been reported. For instance, hnRNP A1 is phosphorylated upon activation of the p38 MAP kinase pathway by osmotic or UV stress, which caused its redistribution from nucleus to the cytoplasm and accompanied enhanced proximal 5′SS usage of E1A reporter minigenes [158]. The redistribution requires phosphorylation of a number of serines within a C-terminal peptide that interacts with transportin 1 [159]. PTBP1 is phosphorylated by protein kinase A upon forskolin stimulation [160]. The target Ser16 is essential for its redistribution from the nucleus to the cytoplasm and neurites [160, 161]. Moreover, the phosphorylation of SR proteins regulates their subcellular redistribution, RNA binding and effect on splicing as well [162-166]; their specific kinase yeast Sky1 is implicated in 3′AG recognition [167]. However, their phosphorylation effect on the 3′SS factors in particular has not been shown.

4.3. Methylation

Many constitutive and alternative splicing factors are methylated including the spliceosome core Sm proteins [150, 168]. The reduction in methylation of Sm proteins leads to aberrant splicing of specific pre-mRNAs. For instance, the splicing of *Mdm4* (MDM4, p53 regulator) exon 7 is repressed in the absence of PRMT5 (protein arginine methyltransferase 5) that methylates Sm proteins [150].

4.4. Acetylation

Acetylation inhibitors also affect various steps of spliceosome assembly, suggesting that acetylation of multiple components/steps of the spliceosome assembly are involved in this process [152]. SRSF2 acetylation on lysine 52 in its RRM destabilizes the protein through proteosomal degradation in the regulation of the cell growth/apoptosis variants of caspase-8 gene [151]. A large number of splicing regulator proteins including most of hnRNPs mentioned above are also acetylated in cells [133]. Particularly lysines 87 98, 224 of hnRNP F that are potential targets of acetylation and ubiquitination are critical for increased protein stability by a deacetylase inhibitor trichostatin A (TSA) [169]. Their acetylation effect on 3′SS usage remains to be examined.

4.5. Ubiquitination

Ubiquitination of U2AF35 contributes to splicing regulation. One of the *Shigella* factors IpaH9.8 specifically interacts with U2AF35 to interfere with splicing of IgM pre-mRNA [170]. IpaH9.8 catalyzes U2AF35 ubiquitination *in vitro* [153].

U4 snRNP component Prp3 is modified by Prp19 complex with nonproteolytic ubiquitin chains [154]. This ubiquitination facilitates the interaction of Prp3 with U5 snRNP and therefore

promotes the stability of U4-U5-U6 tri-snRNP. Usp4 deubiquitinates Prp3 and disrupts the U4-U5-U6 tri-snRNP in HeLa splicing extracts to inhibit splicing of *ftz* pre-mRNA. Moreover, knockdown of Usp4 decreased the levels of a group of correctly spliced mRNAs. This suggests the role of Prp3 ubiquitination and its reversal in spliceosomal rearrangements and inhibition of splicing. Moreover, a component of U5 snRNP called Prp8p is also ubiquitinated and this modification has been suggested to be involved in the maintenance of U4-U5-U6 tri-snRNP assembly [171, 172]. Therefore, more than one spliceosomal factors are ubiquitinated during splicing.

4.6. Spliceostatin binding to SF3b

Spliceostatin does not modify splicing factors posttranslationally but shows highly specific binding to the 3′SS factor SF3b; therefore, it is included here.

Spliceostatin is a methylated derivative of a natural product FR901464 [155, 173]. Originally this product was identified for its anti-cancer effect through enhanced transcription activity of SV40, which promotes cell cycle arrest [174]. Later on it was discovered that spliceostatin binds to SF3b [155], an important component of U2 snRNP critical for its interaction with U2AF [175-178]. The binding of spliceostatin to SF3b results in inhibition of U2 snRNP recognition of the branch point, in complex formation and *in vitro* splicing assays [155, 179, 180]. Spliceostatin treatment of HeLa cells lead to nonsense-mediated mRNA decay (NMD) of splice variants and downregulation of genes important for cell division, such as cyclin A2 and Aurora A kinase [179].

Taken together, current evidence have demonstrated that various modifications of *trans*-acting factors regulate protein level, protein-protein or protein-RNA interactions to control the 3′SS usage. However, the modifications of many more 3′SS splicing factors remain unknown.

5. Evolutionary emergence of regulatory elements and factors

Global insights into alternative exon creation revealed that over the course of mammalian evolution the abundance of alternative splicing significantly increased as compared to lower vertebrates [181]. This is further consolidated by independent alternative splicing data sets from different species [182, 183]. However, the evolutionary changes that contributed to the splicing regulatory mechanisms involved in the emergence of alternative exons remain largely unknown.

Evolutionary mechanisms for the *de novo* creation of alternative exons include tandem exon duplication, exonization of transposable elements and transition of a constitutive to alternative exon [184-188]. The transitions are achieved through accumulation of mutations at splice sites or splicing regulatory elements [188-191]. The splice sites of lower vertebrates such as fish show remarkably low prevalence of regulatory elements such as GGG [192]. This implies that the 3′SS regulatory elements perhaps are largely characteristic of mammals.

Sequence alignment of the upstream 3′SS of *Snap25* exon 5a from different species showed that the hnRNP K binding motifs are mostly conserved in vertebrates and expanded into multiple copies in mammals whereas they are absent in invertebrates [42]. Similar sequence alignment of STREX upstream 3′SS of different species indicated that the CA contents of CaRRE1 increased from lower to higher vertebrates [45, 193]. For the examined G tract element host genes [49], it mainly emerged in mammals over the course of evolution. Moreover, the

critical role of the G tracts in splicing repression is confirmed by mutagenesis and splicing assays [49]. The evolutionary emergence of this splicing regulatory element suggests its role in the transition of constitutive to alternative exons. Together, these evolved regulatory elements may have contributed to the proteomic diversity in mammals [193].

6. Further considerations

The competition model is perhaps by far the simplest and most well supported for the 3′SS regulation. However, many cases have emerged that cannot be simply explained by this model. Therefore, other mechanisms such as steric hindrance, splice site competition and local RNA secondary structure or looping have also been proposed. Yet, more direct evidence and molecular details are still needed for some of the models proposed later. Also how the inhibition or activation happens after recruitment of the *trans*-acting factors remains unknown in most cases.

Despite the progresses in understanding the splicing regulation of 3′SS usage, further research is required for deeper insights into the regulatory mechanisms. The regulation of splicing takes elaborate interplay of multiple factors to determine the fate of an exon. The examples discussed in this chapter mainly focus on the role of a particular element/factor within 3′SS in splicing regulation. How they collaborate with other elements/factors or splice sites to determine the choice of a particular splice site among the many other sites remains largely unknown. For example in the regulation of STREX splicing, hnRNP L is a repressor through binding to CaRRE1 [45-47], but its interaction with other factors such as PTBP1, hnRNP K and LL in the regulation remains unknown [45, 194]. Morevoer, the regulation of *trans*-acting factors by upstream signals has just started to be unveiled; more effort is needed to fully understand the

impact of this regulation on 3′SS choice and splicing in general. A number of other modifications of splicing factors such as acetylation and methylation in the regulation of alternative splicing have just emerged recently [132, 133, 150-152, 169, 195, 196]. Furthermore, chromatin modification and the transcriptional machinery also have important impact on splice site choice; however, their detailed mechanisms of action remain to be answered though models based on transcription rate and splice site strength have been proposed.

The functions of novel non-coding RNAs (ncRNA) have just started to be revealed as well but already shown to impact many steps of gene expression [197]. Their role in splicing regulation particularly through 3′SS remains unexplored. One study showed that small nucleolar RNA (snoRNA) HBII-52 regulates the splicing of serotonin receptor $5HT_{2C}$ by base pairing with a silencer element in its exon to influence the splicing of Vb through alternative usage of a 5′SS [198]. In another study, a natural antisense transcript (NAT) that overlaps with 5′SS is associated with intron retention in the 5′ UTR of zinc finger E-box binding homeobox 2 (*Zeb2*) pre-mRNA [199]. This intron retention preserves the internal ribosome entry site (IRES) hence allowing the translation of Zeb2 protein. Several other studies also suggest a link between anti-sense RNAs and posttranscriptional processing particularly splicing [200-202]. Such mechanisms for 3′SS regulation remain to be explored.

Beyond the expected regulatory elements/factors, how do the constitutive (or even "core") spliceosome factors regulate alternative splicing? Besides differences in the splice site sequences and strengths of target transcripts, the role of the homologous factors such as PUF60 and CAPER to U2AF65 [127, 203], U2AF26 to U2AF35 [204] in this process remains to be intensively explored. Moreover, many splicing factors themselves are alternatively spliced.

Though some of their variant transcripts including those of the core spliceosome factors are subject to NMD [205], in other cases variant protein products of splicing regulators do have differential effects on splicing [206]. Furthermore, pseudogenes have also come into play with gene expression [207]. As analysis of the genome/transcripome accelerates, more such homologous factors, splice variants, pseudogene or even non-homologous factors but with overlapping functions might be identified beyond the previously known constitutive or "core" spliceosome factors.

In addition to studying the details of molecular mechanisms regulating individual exons, genome/transcriptome-wide analyses among different species have identified conserved splicing regulatory elements within/nearby splice sites and reveal their evolutionary emergence [192, 208]. A human genome search identified several potential splicing regulatory elements including evolutionarily emerged splicing silencer G tract element between Py and 3′ AG [49]. Application of these approaches to more elements/factors will help reveal the diverse regulation of 3′SS choice and evolutionary changes that contributed to the increased proteomic diversity and organismal complexity in mammals. The knowledge of large-scale analysis of splicing regulation will also help us understand the pathological conditions that arise from aberrant splicing due to genetic mutations. For instance, the mutations of several 3′SS spliceosome factors are associated with different clinical outcomes [209-211], suggesting their cooperation with other genetic lesions [212], or different target exon profiles.

In conclusion, the diversity of characterized 3′SS regulation has been increasing rapidly in the past decade. With more possibilities of regulation remaining to be explored, this diversity is likely to increase. The diverse regulation contributes to the higher transcriptome and proteome

diversity in mammals than lower vertebrates. Further studies will help us understand not only this important step of gene regulation in normal cell physiology but also the pathophysiology and development of therapeutic approaches for aberrant splicing-caused diseases.

CONTRIBUTIONS: I wrote the draft of this manuscript and revised with Dr. Jiuyong Xie.

ACKNOWLEDGEMENTS: This work is supported by the Natural Sciences and Engineering Council of Canada (NSERC), the Canadian Institutes of Health Research (CIHR, FRN_106608) and Research Manitoba.

CHAPTER II

Overview of the Thesis Research

The theme

Among the 3′ splice site (3′SS) RNA elements reviewed in Chapter I, the REPA (regulatory elements between the Py and 3′AG) G tracts are of particular interest because of their unique location and potential role in evolutionary emergence of novel splice variants in mammals [49]. The goal of this study is to understand the role of these evolutionarily emerged REPA G tract elements in the regulation of alternative pre-mRNA splicing and contribution to protein diversity. The REPA G tract host cell cycle promotor gene *PRMT5* (protein arginine methyltransferase 5) was used as a model to explain the molecular basis of their splicing regulation and functional consequences of resulting isoforms [49].

PRMT5 catalyzes the symmetric dimethylation of arginines in many proteins to control key cellular processes such as gene expression, cell cycle progression, differentiation and maintenance of Golgi structure [213-228]. The control of these diverse processes often involves the methylation of histones to regulate the expression of relevant genes including some cell cycle regulators. The direct methylation of some cell cycle regulators by PRMT5 has also been reported to promote proliferation and tumorigenesis. Moreover, PRMT5 has been found to be upregulated in a number of cancers. Because of the cell cycle accelerating and the tumorigenesis promoting functions of PRMT5 it has been regarded as an oncogene. In the maintenance of Golgi structure, PRMT5 colocalizes with it and methylates its component. Moreover, PRMT5 shows dynamic patterns of intracellular localization that are associated with pluripotency or

differentiation properties of embryonic cells. However, the molecular basis of the versatility of regulation and localization patterns remains largely unexplored.

A splice variant of PRMT5 (PRMT5S) is produced upon skipping of exon 3 and partial segment of exon 4 [49]. The upstream 3′SS of exon 3 contains a splicing silencer REPA G tract element. Similar G tract elements are found in a large group of human genes. These G tract elements appeared mainly in mammals over the course of evolution and likely contributed to the increased occurrence of alternative splicing and protein diversity in higher species. However, the molecular mechanism underlying this regulation and functional consequences of the resulting isoform remain unknown.

Similarly positioned CA-rich elements between Py and 3′ AG are bound by hnRNP L to repress the splicing of downstream exon through inhibition of U2AF 65 binding to Py [45-47, 229]. G-tract elements at other locations are known to be recognized by hnRNP H/F to regulate splicing, particularly as splicing enhancers in introns. However, the molecular mechanism underlying REPA G tract mediated splicing repression remains to be explored.

The alternative splicing of PRMT family members is known to change its localization and activity [230]. However, whether the skipping of PRMT5 exon 3 and partial exons 4 confers distinct properties to the shorter isoform or leads to differential effects on cells remains unknown. Investigation of the effects of this isoform on cell cycle progression, gene expression, Golgi structure and differentiation will help to understand the diverse roles of PRMT5 in cells.

Therefore, the focus of this thesis project is on elucidating the molecular mechanism underlying the REPA G tracts mediated splicing repression that contributed to the evolutionary emergence of shorter PRMT5 splice variant. The differential properties of resulting splice variant

as well as its distinct functional consequences in key cellular processes including cell cycle, gene expression, maintenance of Golgi structure and cell differentiation are also explored.

Thesis work

The first part of this thesis shows functional clustering of genes containing the $(G)_{5-8}$ class of G tracts, evolutionary emergence of such G tracts in PRMT5, its mechanism of splicing regulation and differential effects of PRMT5S on cell cycle and gene expression. This study revealed that PRMT5 G tract element is bound by hnRNP H/F which interferes with U2AF65 binding to Py and inhibits splicing. This splicing repression creates a shorter isoform with intact methyltransferase domain but with opposite effects on cell cycle progression as compared to the full length PRMT5. It preferentially regulates the expression of a group of genes involved in cell cycle arrest at interphase. This study explains the mechanism of splicing regulation by the evolved PRMT5 G tract element and its role in creating a novel isoform with differential functional consequences.

The differential effects of PRMT5S on cell cycle and gene expression prompted us to further elucidate its distinct properties and to explore its effects on other cellular processes controlled by PRMT5.

Therefore, the second part of this thesis is focused on expression and intracellular localization patterns of PRMT5S as well as its effects on maintenance of Golgi structure and cell differentiation. This study revealed that PRMT5S shows differential expression among different cell types and tissues and exhibits distinct localization patterns. It also shows differential effects on the maintenance of Golgi structure and opposite effects on differentiation of dendritic cells. Moreover, the two PRMT5 isoforms regulate expression of a large number of genes involved in

various processes particularly apoptosis and differentiation. This study shows the detailed distinct properties of PRMT5S and its differential effects on cells.

Overall implications

This study provides a direct link between evolutionary emergence of a novel class of G tract elements between Py and 3′ AG and splicing regulation leading to differential functional consequences. It provides the molecular basis for the splicing regulation by G tract elements that contribute to the emergence of a PRMT5 splice variant with distinct effects on cells. This study suggests a similar mechanism of splicing regulation for a large group of human genes containing similar G tract elements which likely contribute to the overall protein diversity.

CHAPTER III

Evolutionary Emergence of a Novel Splice Variant with an Opposite Effect on the Cell Cycle

Integration into the theme

A large group of human genes contain evolutionarily emerged splicing silencer REPA (regulatory elements between the Py and 3′AG) G tracts and several examined exons were found to be regulated by these elements [49]. However, the molecular mechanism underlying this regulation and its contribution to the proteomic diversity over the course of evolution has not been reported.

Therefore, in this part of the thesis research we decided to elucidate the molecular mechanism underlying the REPA G tract mediated skipping of exon 3 and partial exon 4 of our model gene *PRMT5* (protein arginine methyl transferase 5) which leads to the evolutionary emergence of a shorter splice variant [49]. We also sought to study differential functional consequences of this shorter splice variant specifically in cell cycle regulation.

This part of my thesis research has been published in MOLECULAR & CELLULAR BIOLOGY [231].

ABSTRACT

Alternative splicing contributes greatly to the diversification of mammalian proteomes, but the molecular basis for the evolutionary emergence of splice variants remains poorly understood. We have recently found a novel class of splicing regulatory elements between the polypyrimidine tract (Py) and 3′ AG (REPA) at intron ends in many human genes, including the multifunctional *PRMT5* (protein arginine methyl transferase 5). The PRMT5 element is comprised of two G tracts that arise in most mammals and accompany significant exon skipping in human transcripts. The G tracts inhibit splicing by recruiting hnRNP H/F to reduce U2AF65 binding to the Py, causing exon skipping. The resulting novel shorter variant PRMT5S exhibits similar histone H4R3 methylation effect as its original longer isoform PRMT5L, but distinct localization and preferential control of critical genes for cell cycle arrest at interphase in opposite to PRMT5L. This study thus provides a molecular mechanism for the evolutionary emergence of a novel splice variant of opposite functions in a fundamental cell process. The presence of REPA elements in a large group of genes implies their wider impact on different cellular processes for increased protein diversity in humans.

INTRODUCTION

Alternative precursor messenger RNA (pre-mRNA) splicing greatly increases the proteomic diversity in metazoans [1, 2, 232]. Particularly, in humans and other primates splice variants have reached the highest complexity [182, 183], with about 90% of human genes alternatively spliced [3, 4]. Aberrant splicing causes a large fraction of human genetic diseases [5, 233]. However, the molecular basis for the evolutionary emergence of alternative exons that impact protein functions and cellular processes remains largely unknown, although several models have been proposed [188].

The 3′ end of introns between the polypyrimidine tract (Py) and 3′AG is highly constrained in sequence and length, with a consensus of PyNYAG (Y: pyrimidine, N: any nucleotide) [89]. However, we have found a CA-rich splicing regulatory element called CaRRE1 at this location [45-47, 229], suggesting relaxation of the constraint in some transcripts and the potential existence of other similar elements. Particularly, a purine-rich element (G- or A-rich) such as a G-tract at this location is expected to strongly disrupt the 3′ splice site (3′SS).

G-tracts with a minimal functional motif GGG are splicing regulatory elements bound by hnRNP H (H1) or its paralogues including hnRNP F [82-84, 86, 87, 234-237]. They are enhancers or silencers of splicing depending on their location in the pre-mRNA [85, 87, 234, 237-240]. We have identified G tracts between the Py and 3′ AG in more than a thousand human genes including *PRMT5* (protein arginine methyl-transferase 5). We call elements at this location REPA (regulatory elements between the Py and 3′AG) [49]. These REPA G tracts appear to have mostly emerged in mammalian ancestors to act as splicing silencers involving hnRNP H/F. However, critical questions regarding the molecular mechanisms for their role in the emergence

of splice variants as well as their functional consequences on cellular processes remain to be answered.

PRMT5 catalyzes the symmetrical dimethylation of protein arginines [168], including histones 2A, 3 and 4 (H2A, H3 and H4) [217, 219, 222, 226], spliceosomal Sm proteins [168, 241], and tumor suppressors PDCD4 (Programmed cell death 4) and p53[215, 220]. It controls gene transcription/RNA processing [150, 196, 216, 217], maintains circadian rhythm and stem cell pluripotency or proliferation [196, 214, 218, 222, 223], accelerates cell cycle progression [221, 224], and promotes tumorigenesis [213, 220]. However, the molecular basis of these diverse roles in cellular processes has not been fully understood.

Here we use PRMT5 as an example to elucidate a molecular mechanism for the emerged REPA G tracts to mediate human exon skipping and its consequence on cell cycle.

MATERIALS AND METHODS

Plasmid construction: Splicing reporters with 3′SS of different species or human G-tract mutant were constructed as described previously [49, 229] (Fig. 6B). To make PRMT5 exon 3 splicing reporters, the ApaI and BglII fragment of DUP175 was replaced by human PRMT5 exon 3 with partial upstream (62nt) and downstream (43nt) introns. To make pET28a-hnRNP H for bacterial expression of His-hnRNP H, the RT-PCR product of HeLa RNA amplified with upstream 5′-TTGGATCCATGATGTTGGGCACGGAAGGTGGAGAG-3` and downstream 5′-GGCTCGAGCTATGCAATGTTTGATTGAAAATCACTG-3′ primers was cloned into pET28a between BamHI and XhoI restriction sites. The Myc-PRMT5L or -PRMT5S expression plasmids were made by PCR amplification of full length or short open reading frames of PRMT5 using Phusion High-Fidelity DNA polymerase, from a PRMT5 cDNA clone (ID 3833019, OPEN Biosystems), and cloned into the pCMV-Myc vector between EcoRI and BglII restriction sites.

For lentivirus-based expression, Myc-PRMT5L or S was sub-cloned into lentiviral vector cppt2E [242]. All constructs were confirmed by sequencing. The lentiviral plasmid pLKO.1 containing shRNA against 3′ untranslated region (3′UTR) of human PRMT5 (shPRMT5, clone ID TRCN0000107085, mature antisense sequence TATTCCAGGGAGTTCTTGAGG) was purchased from OPEN Biosystems.

Cell culture and transfection: The cells were grown and transfected as described previously [49]. We used 0.4 µg of expression plasmids for HeLa cells in 24 well plates. For virus preparation, we carried out calcium phosphate-mediated transfection of HEK293T cells as described [243].

RNA interference and rescue: Vesicular Stomatitis Virus Glycoprotein-pseudotyped lentiviral vectors were produced and transduced as described previously [243]. HeLa cells were transduced with shPRMT5-containing virus for 3 hours in 24-well plates. After 36 hours, the cells were transduced again with shPRMT5 virus only or with shPRMT5 virus and either of the rescue protein-expressing viruses for 12 hours followed by splitting. Five days after first transduction, cells were harvested for western blot, RNA extraction or immunostaining. The knockdown hnRNP H/F was carried out as described previously [49, 84].

RT-PCR and primer extension: We performed semi-quantitative RT-PCR of endogenous PRMT5 exon 3 splicing and for validation of differential expression in RNA-Seq analysis, as previously described [49, 229]. For human or zebra fish PRMT5, we used upstream primers 5′-CAGGAACCTGCTAAGAATCG-3′ or 5′-GACCCTGCAAAGTCACGTCCTG-3′ and downstream γ-^{32}P-ATP-labelled primers 5′-GCCAGTGTGGATGTGGTTG-3′ or 5′-GATTCGAGCCAGGTTGGCACAG-3′ respectively (Fig. 6A), and amplified for 24 cycles. Primer extensions were carried out as described [229], but with a γ-^{32}P-ATP-labelled reverse

primer DUP10 (5′-CAAAGGACTCAAAGAACCTCTG-3′) annealed at 50 ^{0}C. RT-PCR or primer extension products were resolved on 6% urea PAGE gels and exposed to phophorimager plates after drying. We quantified band intensities using ImageJ (National Institutes of Health). The splicing efficiency or exon inclusion/skipping level is expressed as the variant intensity relative to the total level of transcripts (included + skipped product).

UV cross-linking and immunoprecipitation: For UV cross-linking, HeLa nuclear extract was prepared as described previously [47, 244]. RNA probes were *in vitro* transcribed in the presence of [α-^{32}P] UTP with T7 RNA polymerase as described [236], from PCR products of human wild type or G tract mutant splicing reporter minigenes using T7 promoter-tagged upstream primers 5′-TAATACCGACTCACTATAGGGAAGACTCTTGGGTTTCTG-3′ or 5′-TAATACGACTCACTATAGGGCTCAAACAGACACCATGCATGG-3′ and downstream primer 5′-CATGGTGTCTGTTTGAGGTTG-3′. UV cross-linking and immunoprecipitation were carried out as described previously [47]. For immunoprecipitation, we used protein G-Sepharose beads (Pierce) coated with 1.5 µg of anti-hnRNP H/F (1G11-Santa Cruz Biotechnology) or 2.0 µg of anti-U2AF65 (MC3-Sigma Aldrich). Immunodepletion of hnRNP H/F from HeLa nuclear extract containing 0.5M NaCl was carried out based on a published procedure [29]. Immunodepleted HeLa nuclear extract was dialyzed 3 times against DG buffer [45] (20 mM HEPE-KOH pH 7.9, 20% glycerol, 80 mM potassium glutamate, 0.2 mM EDTA, 0.2mM PMSF, 1.0 mM DTT).

Purification of recombinant His-hnRNP H: To purify the recombinant N-terminal histidine-tagged hnRNP H, *Escherichia coli* Rosetta-gami 2 (DE3) pLysS (Novagen) transformed with pET28a-hnRNP H were grown overnight in LB medium containing 50 µg/ml kanamycin and induced with 0.3 mM isopropyl 1-thio-β-D-galactopyranoside at 37 ^{0}C for three hours. Protein

purification was carried out as described [85] except that PBS was used instead of 50 mM Tris-HCl (pH 8.0) containing 0.1 M NaCl.

Deep sequencing of RNA transcripts (RNA-Seq): HEK293T cells expressing Myc-PRMT5L or -PRMT5S were used for RNAseq due to their high transfection efficiency (~90%, confirmed by immunostaining of Myc-tag epitopes). The total RNA of control (non-transfected), PRMT5L- or PRMT5S-expressing groups (each in triplicate) were reverse transcribed using random primers for library preparation and subject to Illumina Hi-Seq protocols at the McGill University and Genome Quebec Innovation Centre using a Hiseq 2000/2500 sequencer. We obtained on average 67±1.9 million uniquely mapped paired HiSeq reads for each sample, with a total of ~2 million reads and 23342.33±168.56 (42.41±0.3%) mapped genes per treatment group. The differential gene expression analysis was done using the DEseq and edgeR Bioconductor package [245, 246]. The edgeR-adjusted p values (<0.05) were used to filter for the significantly changed transcripts.

Western blot analysis and immunostaining: Western blots were performed as previously described [161]. Anti-U2AF65, anti-hnRNP F/H (1G11), anti-hnRNP H, anti-nucleolin, anti-β-actin and anti-cMyc were purchased from Santa Cruz Biotechnology Inc. Anti-PRMT5 (EPR5772) and anti-H4 symmetric dimethyl R3 (ab5823) were purchased from Abcam, and rabbit anti-U2AF65 and FITC conjugated anti-α-tubulin from Sigma. Immunostaining was carried out as described [161], using rabbit anti-cMyc or mouse FITC conjugated anti-α-tubulin primary antibodies at dilutions of 1:100 or 1:1000 respectively. Texas-Red conjugated anti-rabbit IgG secondary antibody was used at a dilution of 1:1000. We used 4’,6-diamidino-2-phenylindole (DAPI) to visualize the DNA at 1:5000 dilution.

Human genome search and pathway analysis: Human genome search and following analysis were carried out as described previously [49].

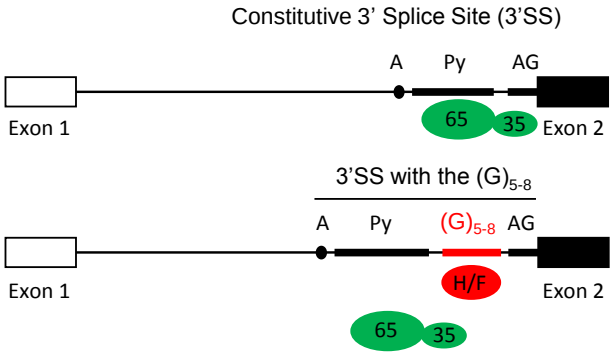
RESULTS

Evolutionary "invasion" of $(G)_{5-8}$ into the upstream 3′SS of a group of human exons

Of the about one thousand REPA G tracts identified from the human genome [49], 130 contain $(G)_{5-8}$ tracts that are significantly associated with alternative splicing (Fig. 5A, and Supplementary Table S1, p=9E-7 in hypergeometric test). Most of them (82%) are G pentamers and 12 of the 14 randomly examined pentamers (86%) were present only in mammalian genes in sequence alignment of the 3′SS of multiple vertebrate species [49] (Fig. 5B and Supplementary Table S1). Thus, the G pentamers appear to have emerged mainly in mammalian ancestors.

The host genes of these G tracts are significantly enriched in several categories of cellular functions (Fig. 5B). The biggest functional cluster contains 30 genes that control cell growth and proliferation, including the *PRMT5* gene that is known to promote cell cycle progression (Fig. 6). We chose the G tracts upstream of its exon 3 for mechanistic and functional studies.

A



B

Clustered Functions of 3'SS G_{5-8}-Containing Human Genes

Top 5 Molecular and Cellular Functions	Number of Genes	Gene Names*	p-value
Cellular Growth and Proliferation	30	ACAN, ARHGEF6, BPI, CGREF1, CUL1, FES, GLUL, GREM1, GRK5, HADHB, HLA-C, IRAK1, ITGB4, LAMB1, LONP1, MICA, PAK4, PDZK1, PFN2, PPAN, **PRMT5**, PXDN, RAPGEF1, RGL3, S100A1, STARD10, TCF19, TNFRSF25, TRPV1, VASP	2.25E-03 - 3.00E-02
Cellular Assembly and Organization	22	ACAN, ARHGEF6, ARPC2, BAZ1B, CAMK1, CROCC, CUL1, FES, GREM1, GRK5, IRAK1, ITGB4, LAMB1, LONP1, MICALL2, PAK4, PFN2, RUFY3, S100A1, TRPV1, VASP, FMNL1	4.26E-05 - 4.83E-02
Cell Morphology	19	ARHGEF6, ARPC2, BPI, CAMK1, FES, FMNL1, GRK5, IRAK1, ITGB4, LAMB1, MICALL2, PAK4, PAX8, PFN2, RAPGEF1, RUFY3, S100A1, TNFRSF25, VASP	3.92E-04 - 4.47E-02
Cellular Movement	18	ACAN, ARHGEF6, ARPC2, CUL1, FES, FMNL1, GREM1, ITGB4, LAMB1, NISCH, PAK4, PAX8, RAPGEF1, S100A1, STAB1, TNFRSF25, TRPV1, VASP	3.92E-04 - 4.83E-02
Cell Cycle	7	CUL1, FES, FMNL1, ITGB4, PAX8, RAPGEF1, VASP	3.92E-04 - 4.83E-02

(*: in red are 7 of the 12 G_{5-8} in mammals only, from 14 verified)

Figure 5. Identification and functional clustering of human genes containing REPA $(G)_{5-8}$ between Py and 3′ AG of 3′SS. A, Diagram showing the position of evolved REPA $(G)_{5-8}$ (Red) between Py and 3′AG, in comparison with the corresponding location of the constitutive 3′SS. Possible locations of the potential *trans*-acting factors U2AF65/35 (green ovals) and hnRNP H/F (red oval) binding to the 3′SS are indicated. **B,** Functional clustering of genes that contain alternative exons with REPA $(G)_{5-8}$. In red within the clustered functions are 7 of the 12 genes with their REPA G tracts found in mammalians only but not lower vertebrates, from 14 genes that can be verified by sequence alignment. *PRMT5* is highlighted in bold.

The exon 3 is skipped in some human transcripts in the UCSC Genome Browser (e.g. GenBank accession # AK302240.1). In all 15 cases of the skipping, a 5′ fragment of exon 4 is also skipped due to alternative usage of a 46nt-downstream 3′SS within the exon (Fig. 6A), producing a previously uncharacterized shorter isoform, PRMT5S. The involvement of exon 4 suggests competition or cooperative interactions between the splice sites, as observed by others

[20, 67, 73, 74, 78]; however, this mode of regulation is not common among the 130 exons. Therefore, we focused on exon 3 skipping to illustrate the role of the G tracts and their mechanism of action.

The human 3′SS upstream of exon 3 harbours a $(G)_3T(G)_5$ element (Fig. 6B). Similar G-tracts of various lengths are present in most mammals but absent in chick or other lower chordates including zebra fish, as shown in the sequence alignment of the corresponding 3′SS of 29 species (G nucleotides within the element in red). Intermediate $(G)_{0-2}[T(A/C)T]T(G)_{2-6}$ sequences are also present in a group of mammalian species within the region corresponding to REPA G tract (Fig. 6B). The alignment thus suggests that the G tracts emerged in a mammalian ancestor but have since evolved differently through accumulated insertions/deletions/point mutations.

Semi-quantitative low cycle number polymerase chain reaction after reverse transcription (RT-PCR) showed that the PRMT5S isoform comprises about 30% of transcripts in human HeLa, MDA-231 and HEK293T cells while it was barely detectable (only about 3%) in zebra fish muscle, egg and brain tissues (Fig. 6C). This dramatic difference in exon skipping levels between human and fish is consistent with the splicing silencer role of the REPA G tracts in PRMT5 and other genes [49]. We thus went on to further characterize its effect on exon usage and mechanism of action.

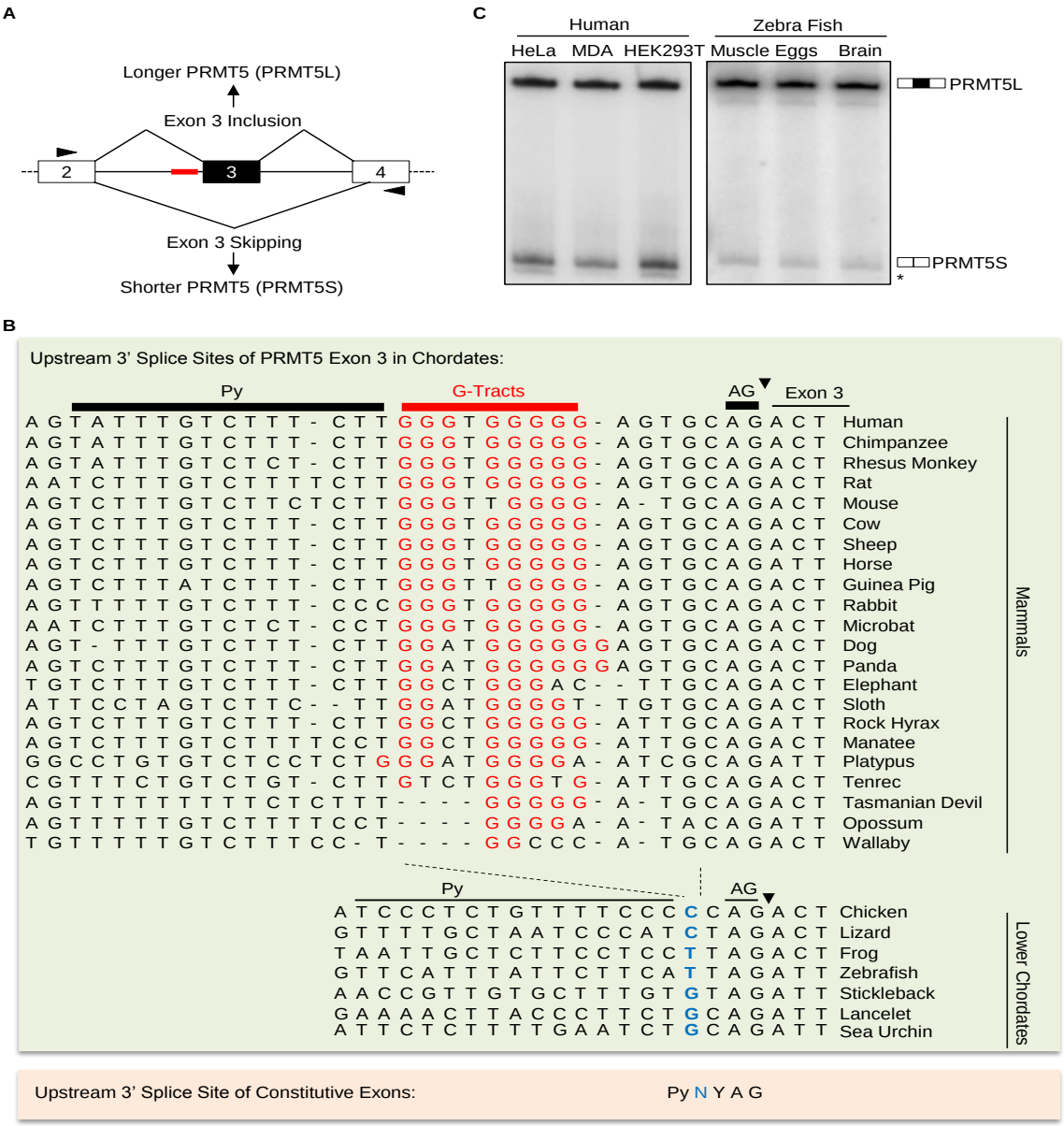


Figure 6. Alternative splicing of PRMT5 exon 3 and evolutionary emergence of REPA G tracts between the Py and 3′AG. A, Diagram of pre-mRNA around human PRMT5 exon 3 (not to scale). Alternative splicing of exon 3 with concomitant usage of an alternative 3′SS within exon 4 creates longer or shorter PRMT5 splice variants, PRMT5L or PRMT5S. Red bar: location of the G tracts; arrowheads: positions of PCR primers. **B,** Alignment of the upstream 3′SS sequences of PRMT5 exon 3 of 29 species. The G tracts (red bar) between Py and 3′ AG (black bars) are indicated above the aligned sequences. Dotted black lines and nucleotides in blue color: corresponding position of the G tracts in birds and other lower chordates. At the bottom is the consensus sequence of constitutive 3′SS for comparison. **C,** Usage of PRMT5 exon 3 in human cell lines and zebra fish tissues. Shown is a representative denaturing PAGE gel of semi-quantitative RT-PCR products with exon 3 included (PRMT5L) or excluded (PRMT5S), as confirmed by sequencing. Asterisk: a product from a potential cryptic splice site.

The evolved PRMT5 G-tracts inhibit U2AF65 binding to Py and splicing by mainly recruiting hnRNP H

We first assessed the effect of the PRMT5 REPA G tracts of different species on splicing by transferring each corresponding 3′SS to the upstream of a heterologous constitutive exon of the splicing reporter DUP175 [229] (Fig. 7A). The vector itself showed almost no exon skipping (lanes 1-2) in primer extension assay of the transiently expressed transcripts in HEK293T cells. Replacing the vector 3′SS with the zebra fish 3′SS, which contains no G tracts, did not change this splicing pattern (lane 3). The opossum 3′SS, containing a G tetramer, caused 15% exon skipping, plus substantial amount of cryptic splicing (*) at a downstream 3′SS in the middle exon (lane 4), a product verified in our previous reports [47, 247]. In contrast, the 3′SS of platypus, dog and humans, each containing a $(G)_4$ plus a $(G)_3$; $(G)_6$ plus a $(G)_2$; or $(G)_5$ plus a $(G)_3$ respectively (Fig. 6B), caused almost complete skipping of the exon along with substantial level of cryptic splicing (Fig. 7A, lanes 5-7, >90% skipping). Importantly, mutating the Gs of the human element to As strongly reduced the percentage of exon skipping as well as cryptic splicing and increased the full length product (lane 8, 43% skipping), as in RT-PCR assays [49]. Therefore, the 3′SS of different species exhibit different effects on splicing, and the splicing inhibition is detected only in the presence of REPA G tracts.

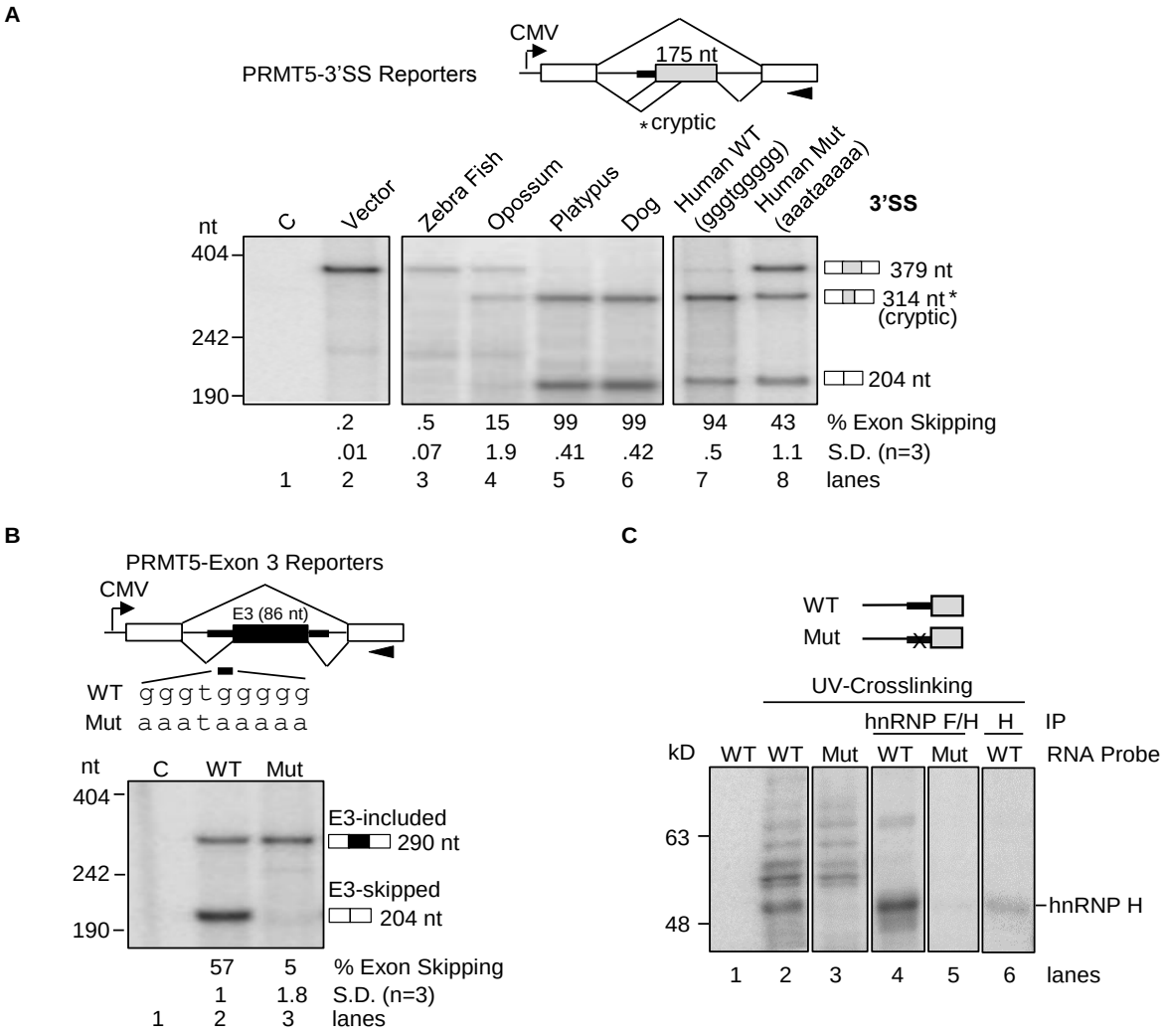
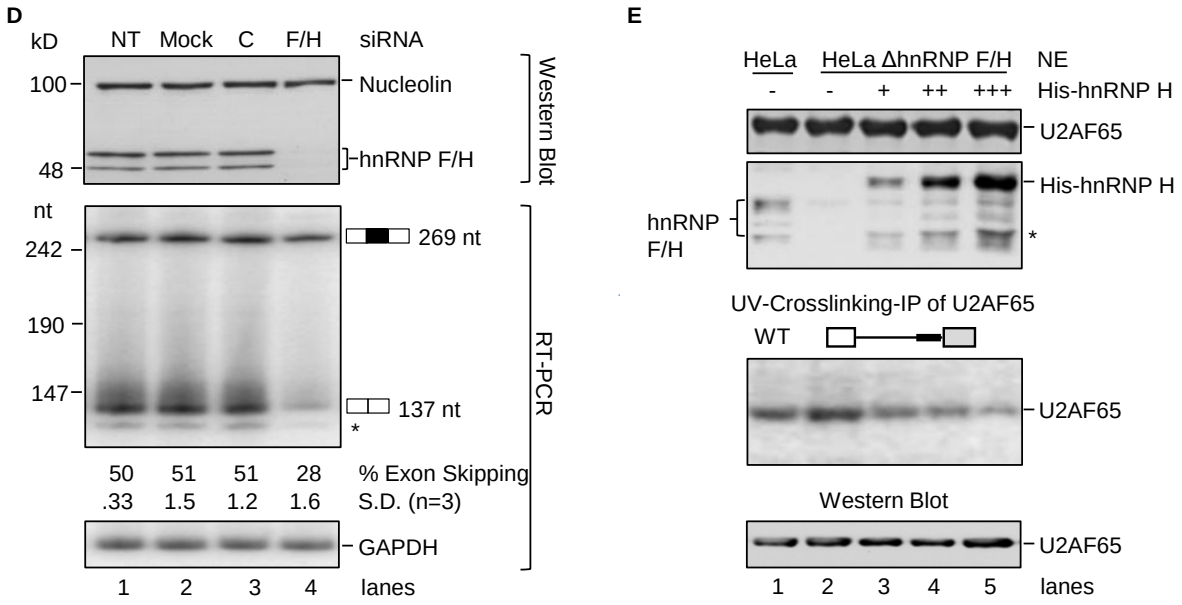


Figure 7. Role of the PRMT5 G tracts as a splicing silencer. A, Splicing of the PRMT5-3′SS splicing reporter minigenes, containing 3′SS of different species upstream of middle exon in the vector DUP175. Upper panel, diagram of the minigenes. Black bar: upstream 3′SS of PRMT5 exon 3 of different species that replaced the corresponding 3′SS of DUP175 (not to scale). Splicing patterns are indicated by slanted lines linking exons (boxes). *: cryptic splicing; arrowhead: reverse primer for primer extension. Lower panel, representative denaturing PAGE gels of primer extension assays of minigenes containing the upstream 3′ SS of vector (lane 2) or of PRMT5 exon 3 of different species (lanes 3 to 8). C: mock transfected sample. The percentages (%) of exon skipping are calculated as the intensity of the 204nt product relative to the total of the 204nt and full length (379nt) products in each lane. **B,** Upper panel, diagram of the minigene for human PRMT5 exon 3 (E3, 86nt) and its splicing products (not to scale). Black box and long bars: human PRMT5 exon 3 and partial flanking introns, respectively; Short bar: wild type G tracts or mutant; Arrow head: primer. Lower panel: primer extension assay as in **A**. **C,** UV crosslinking assay of the binding of hnRNP H/F to the human PRMT5 G tracts. Upper panel, probes used in UV cross-linking. Lower panel, phosphorimages of PAGE gels of proteins from HeLa nuclear extract cross-linked to wild type or mutant RNA probes (lanes 2, 3) and immunoprecipitated hnRNP H(H1) from the reactions (IP, lanes 4-6).



D, Effect of hnRNP H/F knockdown on endogenous PRMT5 exon 3 skipping in HeLa cells. Upper, western blot of non-treated, mock, control siRNA (scrambled)- or hnRNP H/F siRNA-transfected HeLa cell lysates (lanes 1-4), showing knockdown of hnRNP H/F. Lower panel are denaturing PAGE gels of semi-quantitative RT-PCR of endogenous PRMT5 variant products and GAPDH. **E,** Effect of hnRNP H/F immunodepletion and of His-hnRNP H add-back on U2AF65 binding to the Py of the upstream 3′SS of PRMT5 exon 3. Upper, western blot of normal (lane 1), hnRNP H/F-depleted (lane 2) or His-hnRNP H add-back HeLa nuclear extracts (lanes 3, 4 and 5). *: His-hnRNP H fragment. Lower panel, diagram of RNA probe (top), phophorimage of immunoprecipitated U2AF65 cross-linked to the G-tract-containing PRMT5 RNA probe in HeLa nuclear extracts (middle) and western blot of the same gel showing equal loading of U2AF65 (bottom).

To confirm that the human G tracts also inhibit splicing of their original downstream exon, we cloned the PRMT5 exon 3 with partial flanking introns into DUP175 and tested its splicing with wild type or G tract mutant 3′SS in HEK293T cells. Fifty seven percent (±1.8%, n=3) of the transcripts from wild type reporter showed exon 3 skipping in primer extension assays (Fig. 7B, lane 2). Importantly, mutating the G tracts nearly abolished the exon skipping (lane 3, 5% skipping). Thus, the evolved human G-tracts are indeed splicing silencers of PRMT5 exon 3.

To elucidate the molecular mechanisms underlying the splicing inhibition by the G tracts, we first verified their *trans*-acting factors using UV crosslinking assays with wild type or mutant RNA probes and HeLa nuclear extracts. In these assays, a crosslinked protein band of about 50kD was specifically abolished by the G to A mutations (Fig. 7C, lanes 2-3), reminiscent of the G tract-binding proteins hnRNP H/F [49, 82-84, 86, 87, 235-237]. Therefore, we immunoprecipitated the cross-linked products with anti-hnRNP F/H (1G11) or anti-hnRNP H (H1) antibody. The result indicates that both hnRNP H and F are crosslinked to the wild type but not mutant probes (lanes 4-5), and the major band contains hnRNP H (lane 6).

Previously we observed a slight increase of the exon 3 inclusion upon hnRNP H/F knockdown in a screening of several exons using agarose gels [49]. To more accurately measure the molar changes of the splice variants, we carried out the more sensitive assay using ^{32}P-labelled primers for RT-PCR and denaturing polyacrylamide gels. The result showed that the percentage of exon 3-skipped transcripts was about 50% in non-treated cells under this culture condition, which decreased significantly to 28% upon hnRNP H/F knockdown (Fig. 7D, lane 4, compared to lanes 1-3, p=7.6E-05, n=3). Taken together with the mutagenesis and crosslinking data (Fig. 7A-C), this result supports that the G tract-binding hnRNP H/F are indeed splicing repressors of endogenous PRMT5 exon 3.

To determine the mechanism of splicing inhibition by the element/*trans*-acting factors, we examined the binding of U2AF65 to wild type or G tract mutant RNA probes in HeLa nuclear extracts. G to A mutation of G tracts increased the U2AF65 immunoprecipitated after UV crosslinking of nuclear extract to the RNA probe (Supplementary Fig. S1), suggesting an inhibitory effect of G tracts on U2AF65 binding. Since the inhibition of splicing by G tract

requires the presence of hnRNP H/F in cells [49], we also examined the effect of the *trans*-acting factors on U2AF65.

We examined U2AF65 binding to the wild type RNA probe in hnRNP H/F-immunodepleted, or hnRNP H add-back HeLa nuclear extracts (Fig. 7E). The depletion of both hnRNP H and F from HeLa nuclear extracts enhanced (2.13 fold) cross-linking of U2AF65 to the RNA (lanes 1 and 2). Importantly, adding back bacterially expressed His-hnRNP H reversed the depletion effect in a dose-dependent manner (lanes 3-5). Thus, hnRNP H binding to the G tracts inhibits U2AF65 binding to the Py.

Taken together, these data demonstrate that the evolved human REPA G tracts recruit mainly hnRNP H to inhibit U2AF65 binding and splicing of PRMT5 exon 3. This provides a molecular basis for the mechanism of splicing inhibition at an early stage of spliceosome assembly.

The splicing inhibitory effect of these G tracts is consistent with their atypical location disrupting the 3′SS consensus but is completely opposite to their reported enhancing effect when positioned at other regions of mammalian introns [85-87, 238]. The evolutionary emergence of the alternative splice site by "invasion" of a *de novo* element between the Py and 3′AG is also distinct from the creation of other alternative splice sites by simple point mutations [191].

The G tract-mediated exon skipping creates a shorter PRMT5 protein (PRMT5S) that exhibits distinct localization and promotes cell cycle arrest at interphase

To determine the functional consequences of the emerged splice variant of PRMT5, we examined the resulting protein PRMT5S. Skipping of exon 3 and concomitant usage of the alternative 3′SS within exon 4 result in deletion of 44 amino acids from the NH_2 terminal region of the original full-length (637aa) PRMT5 protein (Fig. 8A, PRMT5S and PRMT5L). However,

PRMT5S (593aa) still contains the intact C-terminal methyl transferase domain, suggesting that it may also methylate proteins as does PRMT5L.

To determine if the PRMT5S has an effect on the methylation status of target proteins in cells, we looked for such an effect by expressing either PRMT5L or PRMT5S in HeLa cells in which the total PRMT5 proteins were knocked down by lentivirus-mediated RNA interference. The PRMT5 proteins were efficiently knocked down and the exogenous isoforms expressed in the knockdown cells respectively (Fig. 8B). We probed the cell lysates with an antibody against Histone 4 symmetric di-methyl R3 (H4R3me2s), a known target of PRMT5 [217, 222, 226, 227]. The antibody recognizes two bands of ~12kD and 15kD in SDS-PAGE gels consistent with the known sizes of the H4 protein [226, 227] (Fig. 8C). The H4R3me2s decreased upon PRMT5 knockdown. Importantly, it was restored to the original level upon expression of either PRMT5L or PRMT5S. Therefore, PRMT5S is capable of maintaining H4R3 methylation, as is PRMT5L. This effect is consistent with the intact methyl-transferase domain in PRMT5S (Fig. 8A).

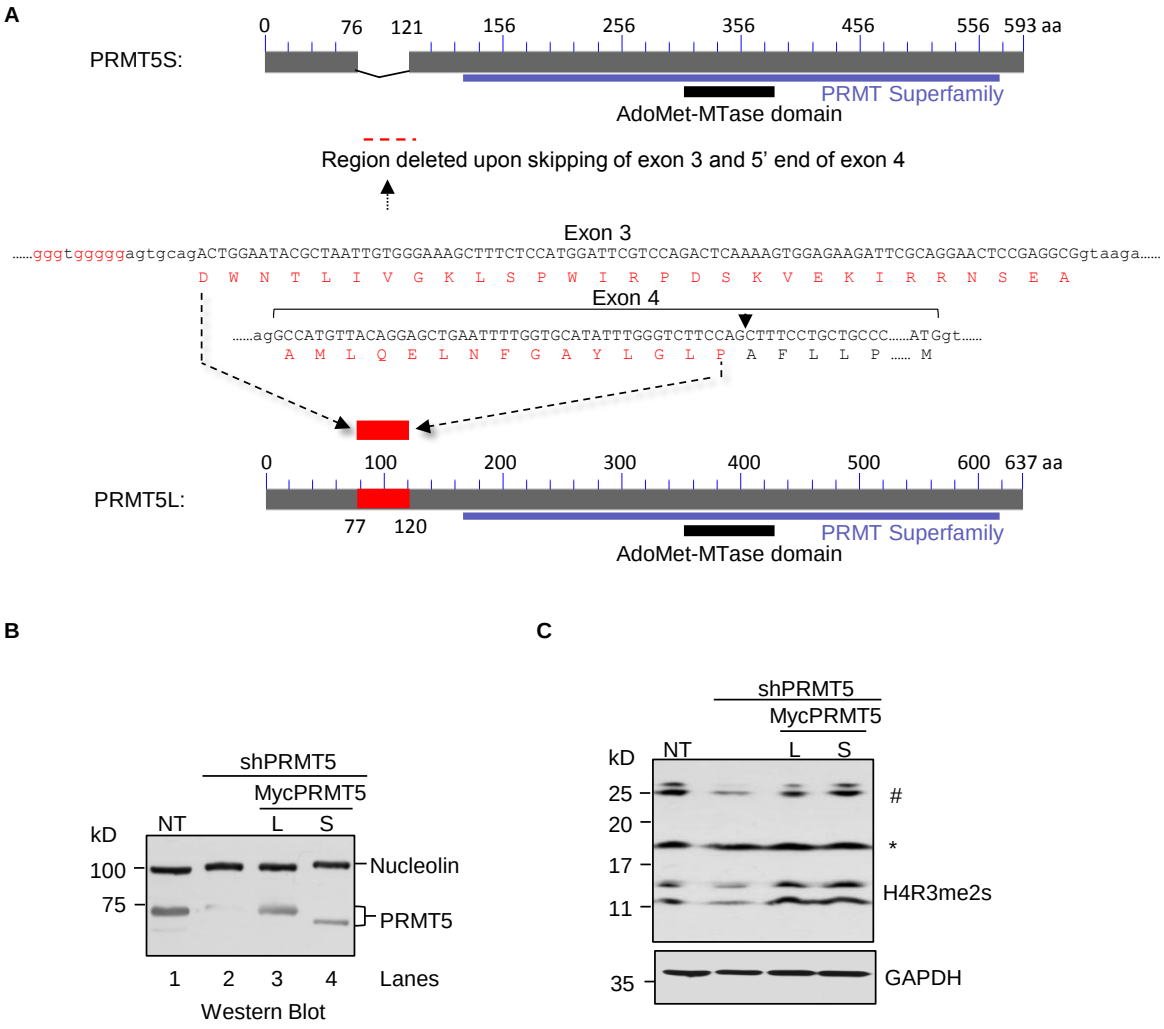


Figure 8. Domains of PRMT5S and PRMT5L proteins, and rescue of H4R3 methylation by either of them. A, Diagram of the domains of human PRMT5S (Upper) and PRMT5L (Lower) and the region of the G tract regulated exon 3 as well as exon 4 (middle). Red nucleotides: G tracts; aa: amino acids; AdoMet-MTase: S-adenosylmethionine-dependent methyltransferases; red letters of amino acids and red bar: the peptide deleted in PRMT5S because of exon 3 and partial exon 4 skipping; arrowhead in exon 4: alternative 3′ splice junction used for PRMT5S. **B,** Western blot showing knockdown with lentiviral short hairpin shPRMT5 and rescue with MycPRMT5L or MycPRMT5S in HeLa cells. **C,** Western blot showing the histone H4R3 methylation in HeLa cells after PRMT5 knockdown/rescue. # and *: bands of unknown identity, likely symmetrically dimethylated arginine epitopes of unknown proteins that are also recognized by the antibody.

To determine the effect of the PRMT5S protein on cells, we examined it during cell cycle, where PRMT5L plays an important role [213, 214, 221, 224, 227]. We differentiated the

interphase and mitotic phases by the typical tubulin and DAPI staining patterns in HeLa cells untreated (control), knocked down of PRMT5 (shPRMT5) or rescued with either one of the splice variants (Fig. 9A, and 8B). Immunostaining of the endogenous PRMT5 in control cells showed diffuse localization to the nucleus and cytoplasm, as well as enrichment in cytoplasm near nucleus (Fig. 9A, upper panel, red), consistent with a previous report [248]. The immunosignals were abolished by shPRMT5 expression as expected (second panel from top, row of PRMT5). The expressed MycPRMT5L mainly localized near the nucleus, a pattern similar to its colocalization with the Golgi apparatus [225], whereas MycPRMT5S localized more diffusely throughout the cell (lower two panels, Red). A Flag-tagged PRMT5L showed similar localization as the MycPRMT5L (page 80). Moreover, though their localization patterns changed during cell cycle but the major difference of diffusing or punctate distribution remained. We thus conclude that PRMT5S exhibits localization patterns distinct from those of PRMT5L in cells.

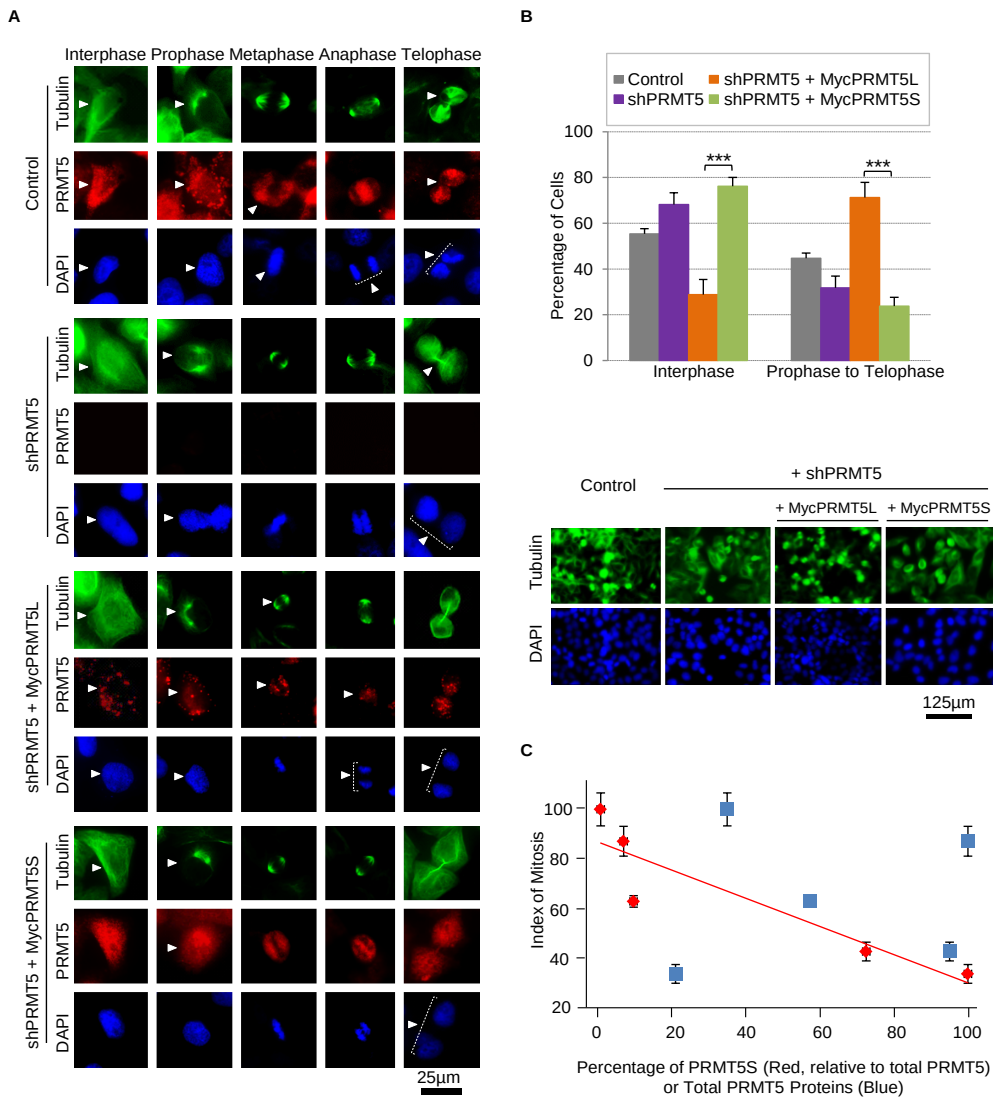


Figure 9. Opposite effect of PRMT5S to that of PRMT5L on cell cycle. A, Representative images of immunostained HeLa cells at different stages of cell cycle, subjected to total PRMT5 knockdown or rescued with MycPRMT5L or MycPRMT5S as described for Fig. **8B**, using anti-α-tubulin or anti-PRMT5. DAPI staining of the nuclei is also shown below each sample. Arrowheads indicate representative cells in the images with multiple cells; dotted lines indicate the diving nuclei. **B,** Upper panel, percentages of HeLa cells in interphase or prophase to telophase (mitosis) after knockdown of PRMT5/rescued as described for panel **A** and in the Fig. **8B** legend. ***: $p < 0.001$. Lower panel: representative lower-magnification images of PRMT5 knockdown/rescue HeLa cells. The patterns of tubulin or DAPI staining typical of cell cycle stages as described for panel **A** indicate that the effect of PRMT5S on the cell cycle is opposite to that of PRMT5L. **C,** Index of mitosis vs the relative level of PRMT5S protein. The relative levels of PRMT5S [PRMT5S/(PRMT5S+PRMT5L)], calculated from western blots of HeLa cells, are shown in red and total PRMT5 protein in blue. For the index of mitosis, the percentage of cells in mitotic phases (prophase to telophase) of the MycPRMT5L rescue group is taken as 100; all others are normalized to this group. The upper and lower bands of PRMT5 in Fig. 8B were measured as PRMT5L and PRMT5S, respectively.

To determine the effects of the two variants on cell cycle, we examined the number of cells at interphase or mitotic phases of cell cycle among these cells (Fig. 9A-B). Upon knock-down of endogenous total PRMT5, the percentage of interphase cells increased from 55.3% in non-transduced controls to 68.2% (±5.09%, p = 0.01, n=3 groups, 300 cells/group, same in the following cell phase analysis in this panel) (Fig. 9B). This change is consistent with the predominant expression of PRMT5L in HeLa cells (Fig. 6C, and 8B) and its mitosis-promoting effect [213, 214, 221, 224, 227]. Importantly, as anticipated, rescue with MycPRMT5L significantly reduced the percentage of interphase cells to 28.7% (±6.6%, p = 0.002) (Fig. 9B). In contrast, rescue with MycPRMT5S could not reverse the effect of knockdown on mitosis, but further enhanced it by increasing the percentage of interphase cells to 76.2% (±3.8%, p = 0.001). Furthermore, the level of PRMT5S (relative to total PRMT5 protein, PRMT5S/(PRMT5S+PRMT5L)), calculated from western blots of HeLa cells, is inversely correlated with the mitosis index whereas the relative level of the total PRMT5 protein alone (PRMT5S+PRMT5L) did not exhibit such correlation (Fig. 9C). This suggests that the relative level of the PRMT5S isoform, but not the total level of the PRMT5 protein, corresponds to the levels of mitosis inhibition in HeLa cells. Therefore, the human PRMT5S has an effect on cell cycle progression opposite to that seen with its full-length counterpart, PRMT5L.

PRMT5S preferentially regulates the expression of a group of genes involved in cell cycle arrest at interphase

To obtain insight into why the evolved PRMT5S has an effect on the cell cycle opposite from that seen with PRMT5L, beyond their obvious differences in subcellular localization, we carried out deep sequencing of mRNA transcripts to identify target genes that PRMT5S may regulate differently from PRMT5L.

The transcriptomes of HEK293T cells overexpressing either of the PRMT5 variants were sequenced using paired-end RNA HiSeq (see Materials and Methods). We identified 50 genes that are preferentially regulated by PRMT5S (regulated specifically by PRMT5S or more strongly by PRMT5S than by PRMT5L, $p<0.05$). We were able to validate the trends of changes observed in RNA-Seq analysis by semi-quantitative RT-PCR of all 22 genes that gave expected products, out of 25 genes examined (Supplementary Fig. S2).

To determine the impact of preferential gene expression regulation by PRMT5 variants on potential cellular functions, we carried out gene ontology analysis. Consistent with the known function of PRMT5 in promoting cell proliferation [213, 214, 221, 224, 227], PRMT5L-preferred genes were significantly enriched in functional clusters of cell growth and proliferation. Strikingly, the functions of PRMT5S-preferred 50 genes clustered most significantly in cell death and cell cycle arrest pathways (Fig. 10A, $p < 0.0001$). Particularly, 8 of them are known to arrest cell cycle at different stages of interphase (G0/G1, S or G2).

A

Functional classification of PRMT5S regulated genes

Category	Function Annotation	p-Value	Genes	# Genes
Cell Death	cell death, apoptosis or necrosis	2.18E-09	BAG3, CRYAB, CTGF, CXCR1, CYR61, DNAJB1, DUSP1, EGR1, EGR2, EGR3, FOS, FOSB, GADD45B, HMOX1, NEK11, PIGR, PTAFR, SMAD7, SOCS3, BAG3, HSPH1, MAFB, PIK3IP1, ARC, HUNK	25
Cell Cycle	**arrest in interphase**	**6.96E-06**	**CXCR1(G1, G1/S); CYR61 (G1); DUSP1 (G1, G2); EGR1,FOS (G0/G1); GADD45B(G1, G2/M);NEK11(S, G2);SMAD7(G1).**	**8**

B

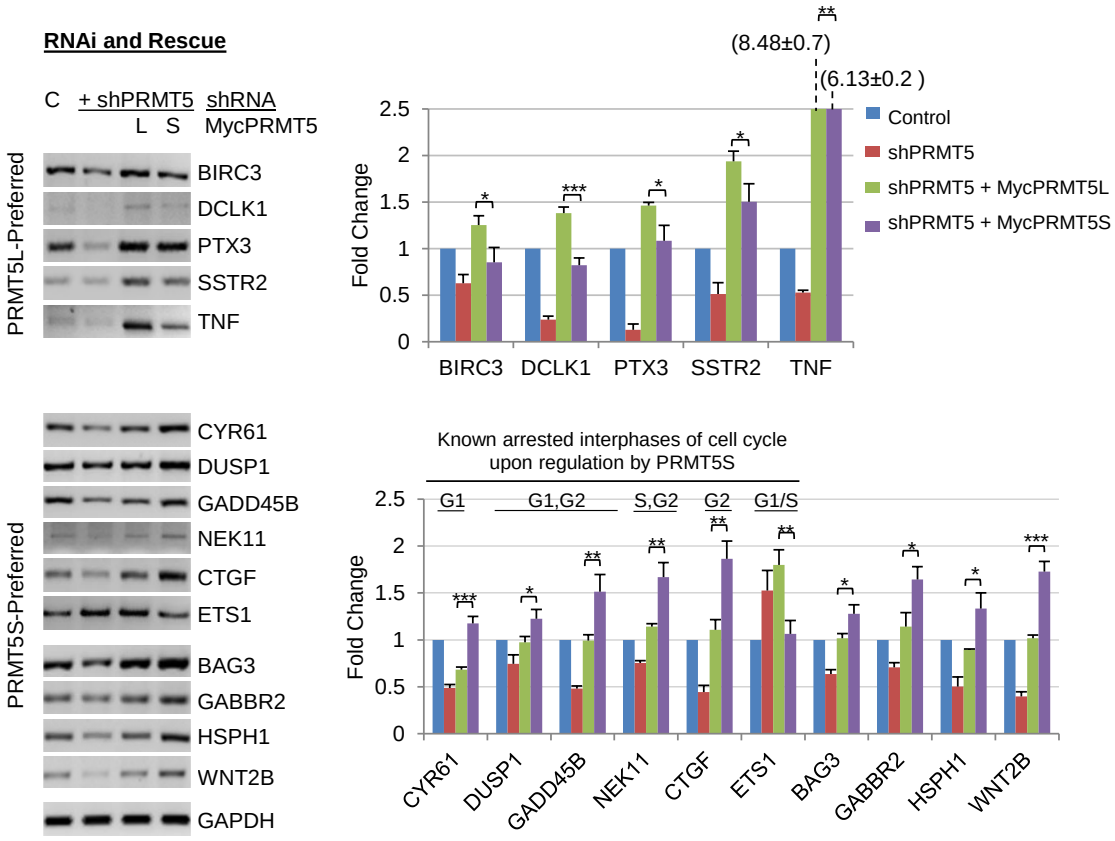


Figure 10. Identification of a group of cell cycle-arresting genes preferentially regulated by PRMT5S in RNAi and rescue assays. A, Functional clusters with most significant enrichment of PRMT5S-preferred genes identified by RNA-Seq analysis. **B,** Preferentially regulated genes by PRMT5L or PRMT5S in PRMT5 knockdown and rescue assays. Left, representative agarose gels of the RT-PCR products of genes that are confirmed to be preferentially regulated by PRMT5L or PRMTS in RNAi/rescue assays in HeLa cells, in alphabetical order for the PRMTL targets and in order of their description in the text for the PRMT5S targets. Four genes not reported to affect cell cycle but also regulated by PRMT5S are shown below the cell cycle-regulating ones. GAPDH: RNA loading control. Right, bar graphs (average ± SD, n = 3) of the changes of the corresponding genes in the same order as in the left panel. Numbers above the dotted lines on top of bars are the fold changes beyond the vertical axis scale. *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$.

To confirm the preferential regulation of cell cycle-arresting genes by PRMT5S, we again used HeLa cells with RNAi-mediated knockdown of total PRMT5 followed by rescue with either PRMT5S or PRMT5L, to examine the expression of 19 genes identified in RNA-Seq analysis as preferentially regulated targets. The changes in the expression of 16 (84%) of these genes upon PRMT5 knockdown were rescued by either one or both PRMT5 isoforms, consistent with the trends observed in RNA-Seq analysis, and 15 of them showed significantly preferred regulation of expression (Fig. 10B). Of these, 5 genes were preferentially rescued by PRMT5L and 10 by PRMT5S ($p<0.05$). Consistent with the gene ontology analysis, the PRMT5S-preferred group includes 6 genes that are known to induce cell cycle arrest when regulated in this manner. The products of these genes function at different phases of cell cycle. Specifically, cysteine-rich angiogenic inducer 61 (*CYR61*), dual specificity phosphatase (*DUSP1*) and growth arrest and DNA-damage inducible beta (*GADD45B*) act at G1 [249-251]; NIMA (never in mitosis gene A)-related kinase 11 (*NEK11*) at S [252]; and *DUSP1*, *GADD45B*, *NEK11* and connective tissue growth factor (*CTGF*) at G2 phase [251, 253-257]. Moreover, PRMT5S particularly reversed the elevated transcripts level of proto-oncogene *ETS1* (transcription factor E-26 transforming sequence-1), which was observed in the PRMT5 knockdown cells. *ETS1* is also a G1/S phase transition enhancer [258]. Therefore, PRMT5S preferentially controls the expression of multiple genes that cause cell cycle arrest at different stages of interphase, providing a reasonable explanation for its different effect cell cycle as compared to PRMT5L.

Taken together, our data demonstrate that the REPA G tracts emerged in mammalian ancestors to recruit mainly hnRNP H for the inhibition of U2AF65 binding to the Py, resulting in splicing repression. A novel PRMT5 isoform was thus generated to have a different subcellular localization and effects on cell cycle progression, possibly through preferential control of cell

cycle-regulating genes, from those of the original isoform. This report therefore provides a molecular mechanism for the REPA G tracts to help generate functional diversity of a protein over the course of evolution.

DISCUSSION

Several models have been proposed for the evolutionary emergence of an alternative exon [188], including transition from constitutive to an alternative exon during evolution, which appears to be the case for PRMT5 exon 3. It transited from a mainly constitutive exon in fish and birds to an alternative exon in mammals during evolution. The existence of weakly detectable exon 3 skipped transcripts in fish (Fig. 6C) is perhaps through other elements/factors affecting the splice site strength. However, the critical factor for this transition is apparently the insertion of REPA G tracts in mammalian ancestors. This report provides a mechanism by which they help generate a novel splice variant that has a distinct functional consequence in cells.

The reported roles of PRMT5 are quite diverse but the molecular basis of its functional diversity has been largely unclear. The novel PRMT5S variant that emerged in mammals adds to the protein and functional diversity of PRMT5 by exhibiting different localization and differential effects on the expression of critical genes in cell cycle control (Fig.s 9 and 10). The western blot analysis of endogenous PRMT5 also showed a shorter protein corresponding to the expected size of PRMT5S (Fig. 8B), consistent with previous observations in different cell lines [49, 259]. Moreover, PRMT5S is differentially expressed among human tissues (page 77), perhaps as a tissue-specific modulator of cell cycles. Interestingly, PRMT5S enhanced the differentiation of dendritic cells while PRMT5L showed an inhibitory effect in ectopic expression assays (page 88). The different effects on cellular processes along with spatially and

temporally regulated alternative splicing of the two variants thus likely contribute to the diverse roles of PRMT5, including but not limited to the maintenance of stem cell pluripotency and promotion of cancer cell growth as well as cell differentiation [213, 214, 218, 221-224, 228]. Understanding these effects could be facilitated by revealing the potentially different profiles of methylation targets and enzyme activities of the PRMT5 isoforms in subcellular locations/compartments in future investigations.

The unusual invasion of the REPA G tract into the space between the Py and 3′AG and its mechanism of action revealed here clearly demonstrate that this region of mammalian 3′SS has been remarkably permissive for the evolutionary emergence of *de novo* regulatory RNA elements. It is possible that other REPA elements, particularly purine-rich ones, may have emerged similarly at this location as well. Thus, more REPA elements are expected to be discovered in the mammalian genome in future investigations. The evolved element allows *trans*-acting factors to bind (Fig. 7), and moreover, is likely to relay upstream signals for further regulation of alternative splicing of the mammalian genes, as in other cases [193]. The presence of such elements in a group of genes implies that the impact of this molecular mechanism likely goes beyond the genes involved in cell cycle regulation, to a wider range of cellular processes in different tissues (Fig. 5B, and Supplementary Table 1) [49].

CONTRIBUTIONS: I am the first author of this manuscript. Dr. Jiuyong Xie provided the conceptual basis of this study and did the G tract identification and functional clustering analysis. I designed and carried out all the experiments, wrote the first draft of manuscript and revised it with Dr. Jiuyong Xie.

ACKNOWLEDGEMENTS: We thank Doug Black for providing polyclonal hnRNP H/F and hnRNP F antibodies for initial tests, for insightful comments on and editing the manuscript; Vincent Lobo for zebrafish; Niaz Mahmood for help with some of the gene expression experiments; the McGill University and Genome Quebec Innovation Centre, particularly Pascale Marquis, for their RNA-Seq service and gene expression analysis; Mike Myschyshyn and Irene Xie for editing and Sika Zheng for helpful comments on the manuscript. This work is supported by a discovery grant from the Natural Science and Engineering Research Council of Canada (NSERC), and in part by a CIHR grant (FRN106608) and a Manitoba Research Chair fund to J.X..

Supplementary Information

Supplementary Table S1: Accessible with published manuscript.

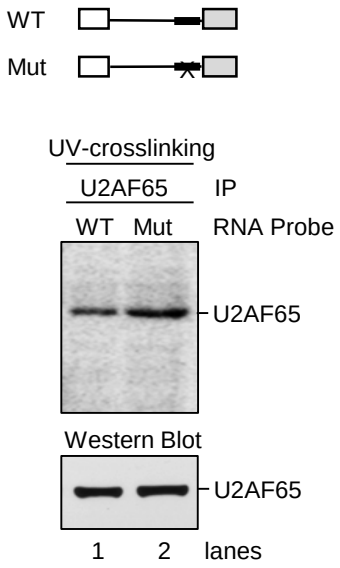


Figure S1. Effect of G tracts on U2AF65 binding to Py. UV crosslinking and immunoprecipitation assay of U2AF65 binding to Py upstream of human PRMT5 G tract in 3′SS. Diagram of RNA probes used in UV crosslinking (top), phosphorimages of PAGE gels of immunoprecipitated U2AF65 from HeLa nuclear extracts UV cross-linked to wild type or mutant RNA probes (lanes 1-2) (middle) and western blot of U2AF65 showing equal loading (bottom).

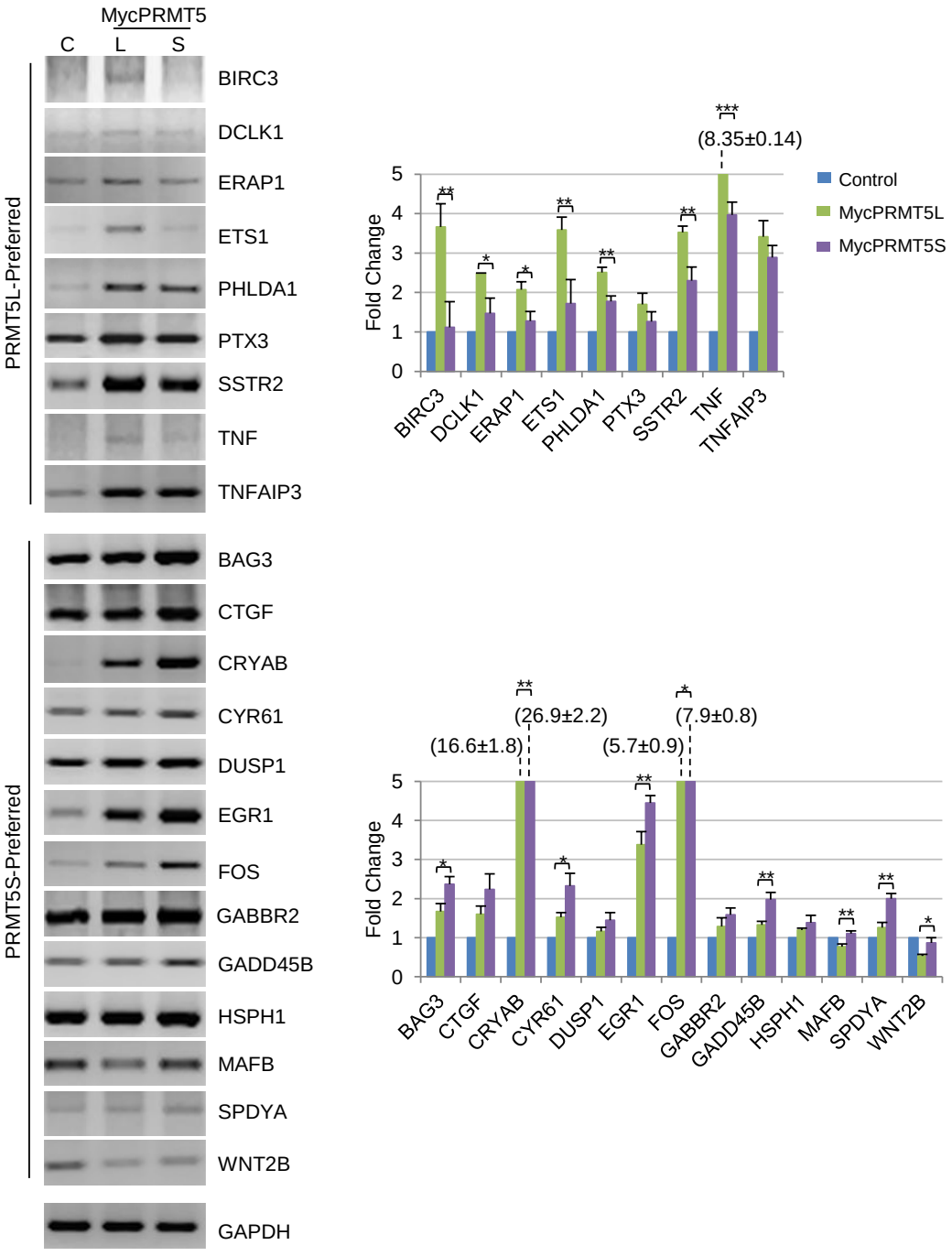


Figure S2. Validation of differential regulation of gene expression by PRMT5 isoforms. Preferentially regulated genes by PRMT5L or PRMT5S in RNA sequencing analysis. Left, representative images of agarose gels showing RT-PCR products of genes that showed differential regulation by PRMT5 isoforms, in alphabetical order. GAPDH: RNA loading control. Right, bar graphs (average ± SD, n = 3) showing the changes of corresponding genes in left panels. Numbers at the top indicate fold changes beyond the vertical axis scale. *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.001$.

CHAPTER IV

Differential Expression, Distinct Localization and Opposite Effect on Golgi Structure and Cell Differentiation by a Novel Splice Variant of Human PRMT5

Integration into the theme

The findings detailed in Chapter III clearly demonstrate the role of a novel class of REPA G tracts in the evolutionary emergence of a shorter PRMT5 isoform (PRMT5S) and its differential functional consequences. The distinct localization patterns and preferential regulation of gene expression by PRMT5S provide a reasonable explanation for its opposite effects on cell cycle progression as compared to full-length isoform. However, its detailed differential properties and effects on other cellular process remained largely unknown.

Therefore, we decided to further investigate the differential properties of PRMT5S particularly tissue-specific expression and localization patterns. We also studied its differential effects on maintenance of Golgi structure and differentiation of dendritic cells.

This part of my thesis research presented in this chapter has been revised after peer review and submitted for publication in BBA-MOLECULAR CELL RESEARCH.

ABSTRACT

Alternative splicing contributes greatly to the proteomic diversity of metazoans. Protein arginine methyltransferase 5 (PRMT5) methylates arginines of Golgi components and other factors exerting diverse effects on cell growth/differentiation, but the underlying molecular basis for its subcellular distribution and diverse roles has not been fully understood. Here we show the detailed properties of an evolutionarily emerged splice variant of human PRMT5 (PRMT5S) that is distinct from the original isoform (PRMT5L). The isoforms are differentially expressed among mammalian cells and tissues. The PRMT5S is distributed all over the cell but PRMT5L mainly colocalizes with Giantin, a Golgi marker. PRMT5 knockdown led to an enlarged Giantin pattern, which was prevented by the expression of either isoform. Rescuing PRMT5S also increased the percentage of cells with an interphase Giantin pattern compacted at one end of the nucleus, consistent with its cell cycle-arresting effect, while rescuing PRMT5L increased that of the mitotic Giantin patterns of dynamically fragmented structures. Moreover, the isoforms are differentially expressed during neuronal or dendritic cell differentiation, and their ectopic expression showed an opposite effect on dendritic cell differentiation. Furthermore, besides their differential regulation of gene expression, both isoforms also similarly regulate over a thousand genes particularly those involved in apoptosis and differentiation. Taking these properties together, we propose that their differential expression and subcellular localization contribute to spatial and temporal regulation of arginine methylation and gene expression to exert different effects. The novel PRMT5S likely contributes to the observed diverse effects of PRMT5 in cells.

INTRODUCTION

Alternative precursor messenger RNA (pre-mRNA) splicing greatly increases the proteomic diversity in metazoans [1, 16, 260]. It is tightly regulated in a spatial and temporal manner [16, 261, 262]. Mutations that lead to aberrant splicing cause genetic diseases [5, 6]. However, the differences between splice variants of many genes that contribute to the proteomic diversity in normal cell function and diseases remain to be characterized.

Protein arginine methyltransferase 5 (PRMT5) catalyzes the symmetric dimethylation of arginines of a diverse group of proteins [215-217, 220, 222, 226, 241], contributing to the regulation of gene expression, snRNP biogenesis, maintenance of stem cell pluripotency and promotion of cell proliferation [150, 216-218, 221-224]. The dynamic cytoplasmic or nuclear localization of PRMT5 during early embryonic development is essential for the maintenance of pluripotency or cell fate determination [222]. Particularly in the cytoplasm, PRMT5 is known to localize to the Golgi apparatus (GA) and methylates GM130, a critical factor for Golgi ribbon formation [225]. During mitosis, GA undergoes dynamic structural changes, with a compacted structure near one end of the nucleus in the interphase and various extents of fragmentation during mitosis [263-265]. PRMT5 co-localizes with GA markers and is thought to be required for the maintenance of Golgi structure at interphase through methylation of GM130 [225]. However, the cell cycle-specific structural changes of Golgi due to modulation of cell cycle by PRMT5 have not been considered.

Here we report the detailed properties of a recently identified shorter PRMT5 isoform (PRMT5S) that evolutionarily emerged through exon 3 and partial exon 4 skipping [231, 266]. These properties are different from that of the original isoform (PRMT5L), particularly

expression among cell lines/tissues, distinct subcellular localization during cell cycle and its effect on Golgi structure as well as cell differentiation and gene expression.

MATERIALS AND METHODS

RT-PCR: Semi-quantitative RT-PCR of endogenous *PRMT5* was performed as described previously [229, 266], using the primer pair hPRMT5F (5′-CAGGAACCTGCTAAGAATCG-3′) and hPRMT5R (5′-GCCAGTGTGGATGTGGTTG-3′). Human tissue-specific total RNA was purchased from Clontech Co.(Cat No. 636643), for which more information can be found at the product page of the company website (http://www.clontech.com). For PCR of *GAPDH* as an RNA loading control, we used the upstream GAPDHF (5′-CTTCATTGACCTCAACTACATGGTT-3′) and downstream GAPDHR (5′-GCTCCTGGAAGATGGTGATG-3′) primers. Both primer pairs are compatible with human, rat or mouse genes.

Plasmid construction: PRMT5 L or S open reading frames were amplified by PCR from PRMT5 cDNA clone (ID 3833019, OPEN Biosystems) using Phusion High-Fidelity DNA polymerase, and cloned between EcoRI and BglII restriction sites of pCMV-Myc vector. PRMT5L-Flag construct was obtained by cloning the open reading frame along with C-terminal Flag tag between XhoI and ApaI restriction sites of pcDNA3.1 vector. PRMT5 L or S were sub-cloned into vector cppt2E for lentivirus-based expression [267]. The constructs were confirmed by sequencing. We purchased plasmid pLKO.1 containing shRNA against human PRMT5 3′UTR (shPRMT5, clone ID TRCN0000107085, mature antisense sequence TATTCCAGGGAGTTCTTGAGG) from OPEN Biosystems.

Cell culture and ectopic expression of Myc-PRMT5 isoforms: HeLa and HEK293T cells were cultured and transfected with Myc-tagged PRMT5 plasmids for overexpression as described previously [47]. MDA-MB-231, BT20, U2OS and B35 cells were cultured in DMEM containing 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin/glutamine (PSG) solution (Invitrogen). PC12 cells were grown in DMEM containing 10% horse serum (HS), 2.5% FBS and 1% PSG. GH3 cells were grown in F-10 media containing 10% HS, 2.5% FBS and 1% PSG. H9 cells were obtained from Wicell and approved for use by the local ethics board and the Stem Cell Oversight Committee of the Canadian Institutes of Health Research. These cells were cultured and RNA extracted as described previously [268]. LA-N-5 and MAL (a lymphoblastoid cell line from an ataxia telangiectasia patient [269]) cells were cultured in RPMI containing 10% fetal bovine serum and 1% PSG. For LA-N-5 differentiation, cells were fed with fresh media containing 5μM 13-*cis*-retinoic acid (RA, R3255, Sigma-Aldrich) every 48 hours for 4 days [270-272]. Murine bone marrow-derived dendritic cells (BMDCs) were generated from bone marrow (BM) precursors as previously described [273]. Briefly, mouse BM cells collected from the femura and tibiae were cultured in RPMI 1640 (HyClone) medium containing PSG (Gibco), 2-ME, 10% FBS (HyClone), and 20 ng/ml GM-CSF (PeproTech). Culture medium was changed every 2 days and cells were collected at 7 time points over the course of differentiation for RNA analysis.

For ectopic expression of PRMT5L or S in BMDCs, we transduced precursor cells with either one of the protein-expressing viral vectors. Irrelevant EGFP-expressing lentiviral vector (Cppt2E) was used as control [273]. Differentiated DCs with ectopic expression of proteins were fixed and immunostained.

RNA interference and rescue: For splice variant specific knock down of PRMT5 we synthesized siRNAs targeting PRMT5S-specific sequence TGTCAGGAAGGGCTTTCCT or PRMT5L-specific sequence CTCCATGGATTCGTCCAGA. We transfected HeLa or MDA-MB-231 cells with 360 pmole of siRNA in 6 well plates. Cells were harvested after 60 hours of transfection for RNA analysis, immunoprecipitation or immunostaining.

Lentivirus transduction for PRMT5 knock down and rescue with either one of the isoforms was carried out as described [243]. HeLa cells were transduced with shPRMT5-containing virus followed by another transduction along with rescue protein expressing viruses after 36 hours. Cells were split and five days after first transduction, harvested for western blot and immunostaining.

Western blots, immunoprecipitation and immunostaining: Western blots and immunostainig were performed as previously described [161]. Anti-Myc, anti-β-actin, anti-Giantin, anti-PRMT5 (SC2232), anti-PRMT5 (SC136202) and anti-nucleolin were purchased from Santa Cruz Biotechnology Inc. Anti-PRMT5 (EPR5772) was purchased from Abcam, anti-Flag from Sigma and Cy3 conjugated anti-CD40 from Biolegend. For immunostaining, we used primary antibodies at dilution of 1:100 and Texas-Red or FITC conjugated secondary antibodies or Cy3 conjugated CD40 primary antibody at 1:1000. 4’,6-diamidino-2-phenylindole (DAPI) was used at 1:5000 dilution to visualize DNA. Immunoprecipitation was performed as described previously [45], using anti-PRMT5 (EPR5772) and whole cell lysates prepared in radioimmune precipitation assay (RIPA) buffer.

RNA sequencing: We transfected HEK293T cells with Myc-PRMT5L or –PRMT5S containing plasmids. Total RNA was extracted from triplicates of control (non-transfected) and transfected

cells, and submitted to McGill University and Genome Quebec Innovation Centre, where RNA samples were reverse transcribed using random primers for library preparation and subject to protocols for sequencing using Hiseq 2000/2500 sequencer. We obtained 67±1.9 million uniquely mapped reads on average for each sample. The gene expression analysis was done using DEseq and edgeR [245, 246]. We used edgeR-adjusted p values of less than 0.05 for further analysis. For functional clustering of differentially expressed genes we used online tool DAVID 6.7 with the highest stringency and p value <0.05 [274, 275].

RESULTS

Differential expression of PRMT5S in mammalian cells and tissues

In PRMT5S, exon 3 together with the 5′ 46nt of exon 4 is skipped (Fig. 11A). We measured the PRMT5 variants among cell lines using low cycle number semi-quantitative polymerase chain reaction after reverse transcription (RT-PCR, Fig. 11B). In mRNA transcripts from human cell lines U2OS, HEK293T, MDA-MB-231, HeLa, BT20, H9 and MAL, PRMT5S levels range from 10% (±2, n=3, in molar percentage, same as following levels) to 43% (±3, n=3) (lanes 2-8). In rat cell lines PC12, B35 and GH_3, no PRMT5S was visible in the agarose gels (lanes 9-11). It was only visible when a more sensitive RT-PCR assay using ^{32}P-labelled primer and denaturing PAGE gels was used (data not shown). Therefore, the PRMT5 variants are differentially expressed among these mammalian cell lines.

To determine the distribution of the novel isoform PRMT5S in humans, we obtained total RNA from 20 different tissues and measured its abundance using semi-quantitative RT-PCR. The result indicates that PRMT5S is differentially expressed among the tissues with varying molar percentages ranging from 8% in the spinal cord to 37% in the kidney (Fig. 11C, lanes 2-

21). Cerebellum, testis, thymus and small intestine also have substantial amount of this variant (22~33%, lanes 4, 16, 17 and 20, respectively). This data thus suggests that the expression of PRMT5S is regulated in a tissue-specific manner.

We also examined the expression of PRMT5 variants similarly in ten different rat tissues (Fig. 11D). Consistent with our observations in the above rat cell lines, the PRMT5S in rat tissues was not visible either in agarose gels.

Together these data indicate that PRMT5S is differentially expressed, and is mainly in human cells and tissues.

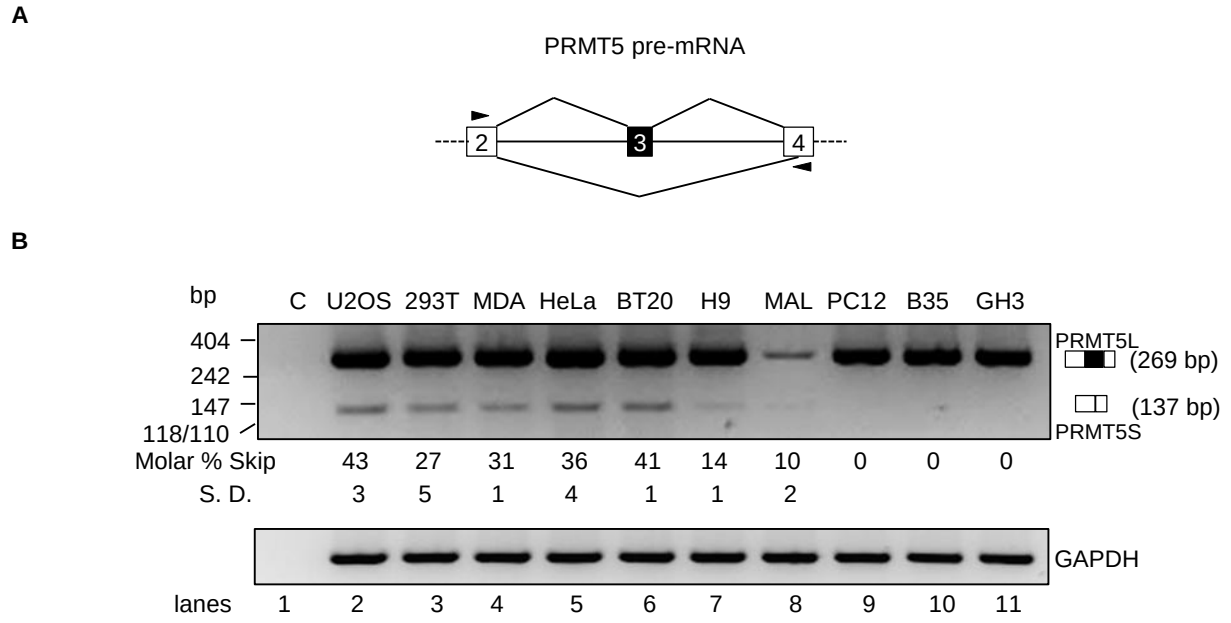
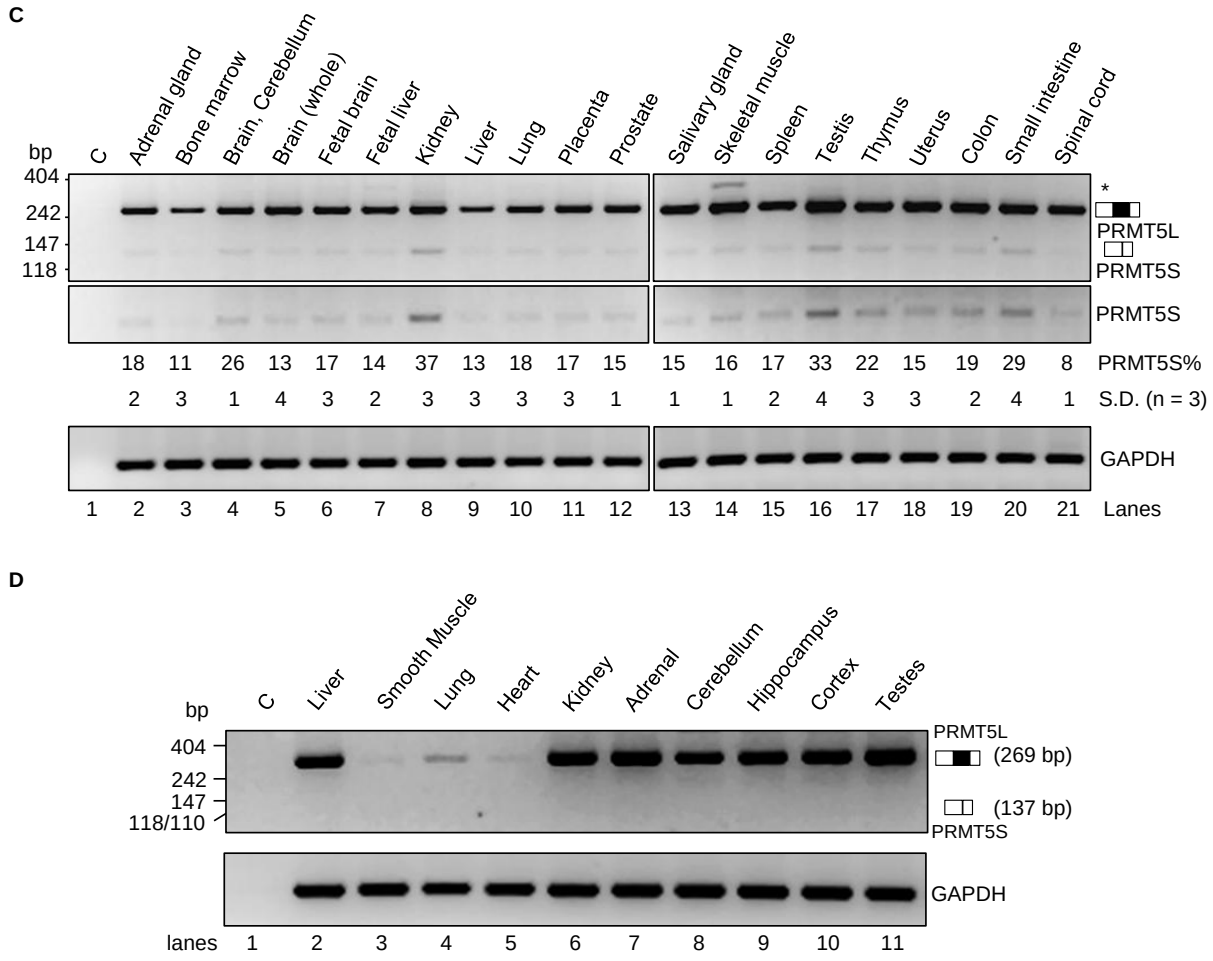


Figure 11. Differential alternative splicing of PRMT5 exon 3 and 5′ end of exon 4. A, Diagram of PRMT5 pre-mRNA around exon 3 indicating its alternative splicing (not to scale). Arrow heads: positions of PCR primers. **B,** Differential expression of PRMT5S in mammalian cell lines. Shown are representative images of agarose electrophoresis gels of exon 3 and partial exon 4-included or -excluded RT-PCR products as confirmed by sequencing (lanes 2-11). C (lane 1): PCR control. Human cells: U2OS (osteosarcoma), 293T (embryonic kidney tumor), MDA (breast cancer MDA-MB-231), HeLa (cervical cancer), BT20 (mammary carcinoma), H9 (embryonic stem cells) and MAL (lymphoblastoid cell). Rat cells: PC12 (pheochromocytoma), B35 (neuroblastoma) and GH_3 (pituitary epithelial-like tumor).



C, Differential expression of PRMT5S among 20 human tissues. Upper panel, representative images of agarose electrophoresis gels of semi-quantitative RT-PCR products of PRMT5L or PRMT5S from the 20 human tissues as indicated. Images of the PRMT5S products with longer exposure and higher contrast of the same gels are shown below. The average molar percentages of PRMT5S from three separate PCR reactions are indicated below each lane. Lower panel, RT-PCR products of GAPDH, as an RNA loading control. C: RT-PCR control without reverse transcriptase. *: an unidentified product in skeletal muscle and fetal liver. **D,** Expression of PRMT5 variants in rat tissues. Shown are images as in B-C of PRMT5 spliced products and GAPDH from different rat tissues (lanes 2-11). C (lane 1): PCR control.

Distinct subcellular localization of PRMT5S from that of PRMT5L

To examine the localization of PRMT5 variants, we performed immunostaining experiments. In HeLa cells, Myc-tagged PRMT5L localized mainly around the nucleus, similar to that observed by others [225], while PRMT5S was more diffusely distributed in both the nucleus and cytoplasm (Fig. 12A-B). In HEK293T cells, similar distribution patterns and differences were observed for the expressed variants (Fig. 12B). Therefore their differential subcellular distribution is consistent between the two different types of cells.

The localization of exogenous Myc-PRMT5L also overlaps with a PRMT5L that is Flag-tagged at the COOH-terminus, when co-expressed and -immunostained with the respective tag antibodies in the same HeLa cells (Fig. 12C). Therefore, the punctate distribution of the Myc-PRMT5L is a property of the PRMT5L protein rather than that of the peptide tags used.

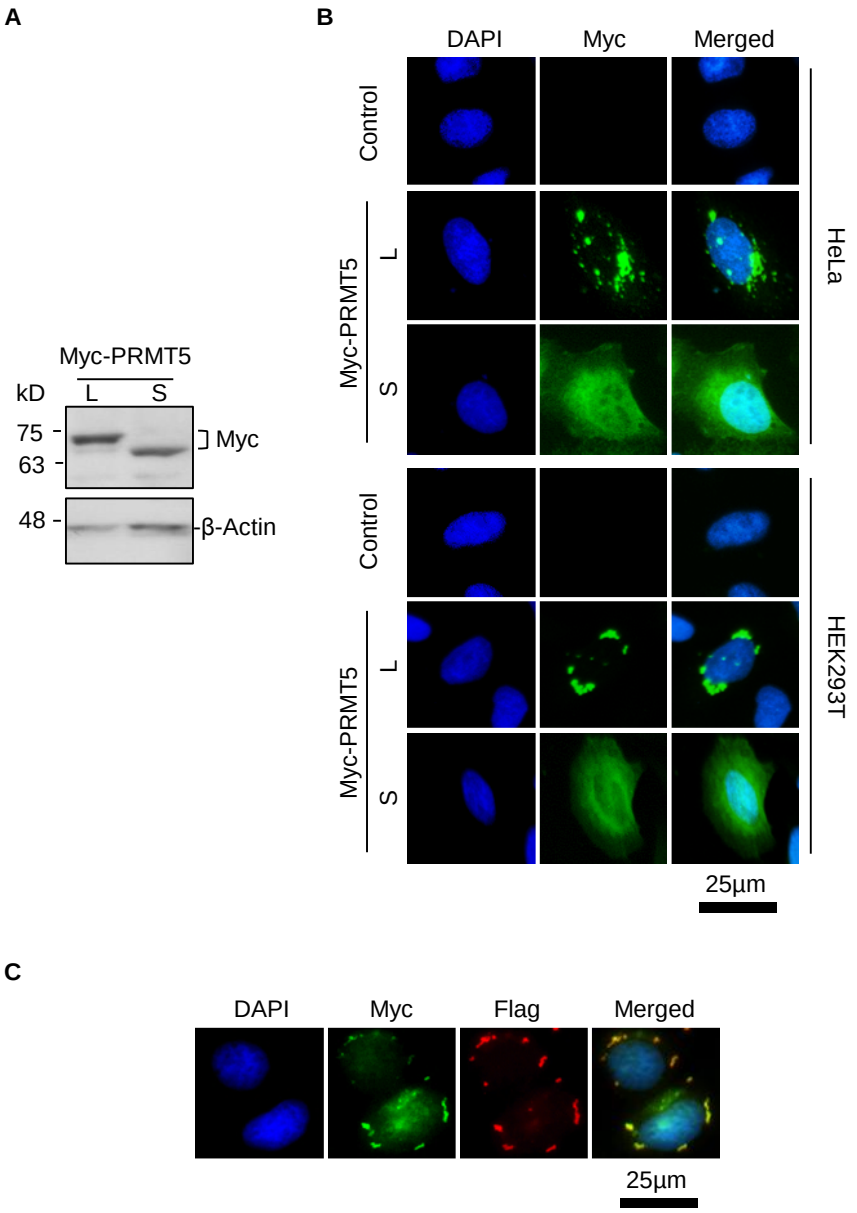


Figure 12. Expression and intracellular localization of exogenous PRMT5 isoforms. A, Expression of exogenous Myc-PRMT5L or S in HeLa cells. Western blot of HeLa cell lysates transfected with Myc-PRMT5L or S showing the expression of either one of the isoforms. **B,** Intracellular localization of exogenous PRMT5 isoforms. Shown are immunostained HeLa (upper panel) or HEK293T (lower panel) cells transfected as in **A**, using anti-Myc. DAPI staining of nuclei is merged with the Myc signals in the last column. Non-transfected control cells are shown in the top row of each panel. **C,** Subcellular localization of exogenous PRMT5 containing a different tag. Shown are HeLa cells co-transfected with Myc- or Flag-tagged PRMT5 and immunostained using anti-Myc or anti-Flag. DAPI staining of nuclei, Flag and Myc signal are merged in the last image.

To verify the localization of endogenous PRMT5 isoforms, we carried out splice variant-specific RNA interference using siRNAs targeting exon 3 (PRMT5L-specific) or the junction of exons 2 and 4 (PRMT5S-specific). We transfected HeLa cells with scrambled or either of the splice variant specific siRNAs. These PRMT5 siRNAs decreased the level of specific target transcripts as verified by RT-PCR, from 33% to 6% for PRMT5S (Fig. 13A, lower band, p=2.1E-6, n=3, lane 3 compared to lane 2) and from 67% to 39% for PRMT5L (upper band, p=1.5E-4, n=3, lane 4 compared to lane 2). Thus, the siRNAs specifically reduced either one of the endogenous PRMT5 variants.

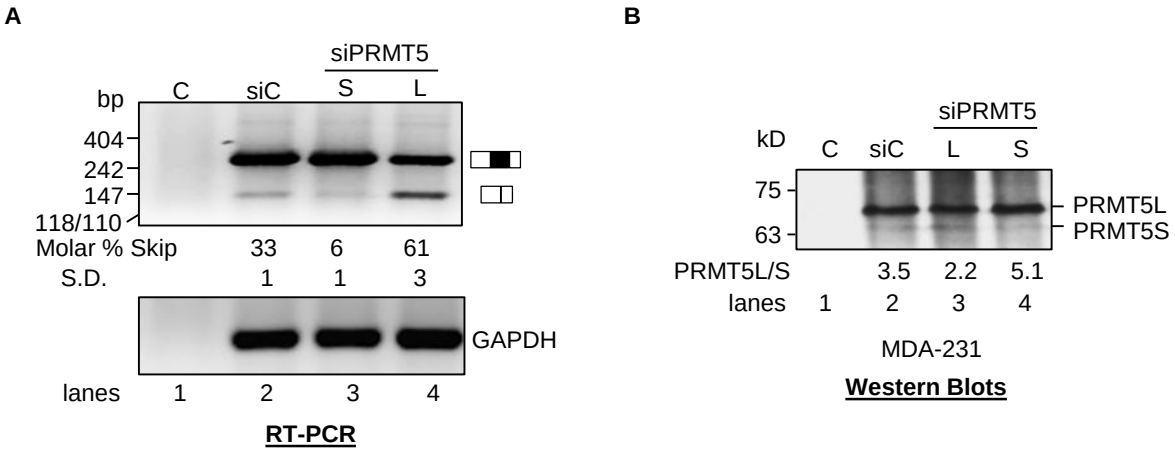
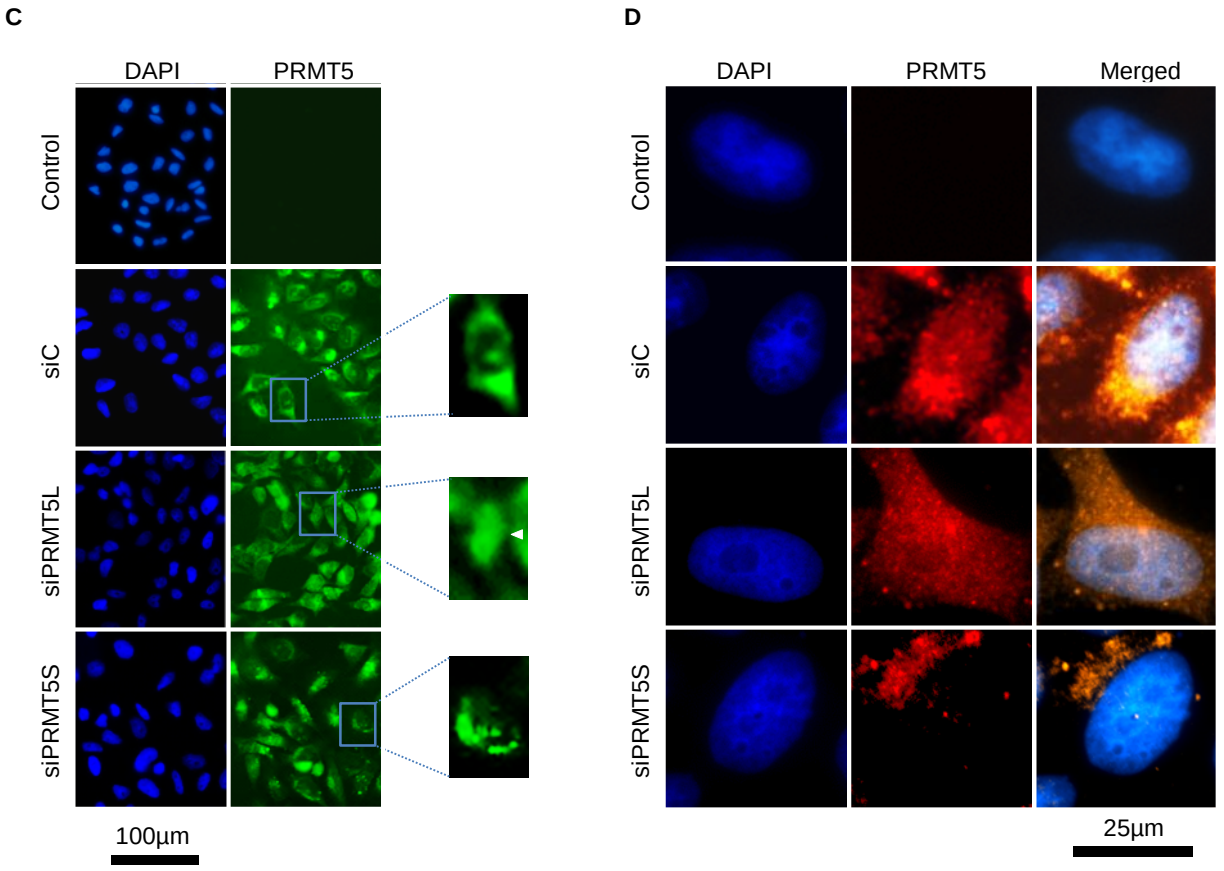


Figure 13. Differential subcellular localization of endogenous PRMT5 isoforms based on splice variant-specific RNA interference and immunostaining. A, Effect of splice variant-specific knock-down on endogenous PRMT5 variants. Shown are representative images of agarose electrophoresis gels of RT-PCR products of endogenous PRMT5 variants or GAPDH from HeLa cells transfected with scrambled control siRNA (siC) or splice variant-specific siRNA (lanes 2-4). C (lane 1): PCR control. **B,** Effect of splice variant-specific knockdown on endogenous PRMT5 protein isoforms. Shown is the western blot of immunoprecipitated PRMT5 from MDA-MB-231 cells transfected as in **A** (lanes 2-4). PRMT5 bands were not detectable in immunoprecipitation controls without PRMT5 antibody (lane 1).



C, Effect of splice variant-specific knock-down on endogenous PRMT5 localization. Shown are representative images of HeLa cells transfected with control or PRMT5 splice variant-specific siRNAs and immunostained using anti-PRMT5 (SC136202). Control cells without PRMT5 antibody are shown in the top row. Enlarged pictures of the PRMT5 signals of example cells are in the right column and pointed by a white arrow head in the panel with multiple cells. Nuclei stained with DAPI are in the left column. **D,** Representative higher magnification images of HeLa cells transfected as in **C** and immunostained with anti-PRMT5 (SC22132). DAPI-stained nuclei are merged with PRMT5 signal in the right column. Control cells without PRMT5 antibody are shown in the top row.

To confirm the variant-specific knockdown of endogenous proteins, we carried out western blots of immunoprecipitated PRMT5 from protein lysates of siRNA-transfected HeLa or MDA-MB-231 cells. Consistent changes were observed in both cell samples but the MDA sample showed a slightly higher relative level of PRMT5S-like band in the assay (Fig. 13B). The

PRMT5L-specific siRNA decreased the ratio of PRMT5L/S-like from 3.5 to 2.2 (lane 3 compared to lane 2) whereas PRMT5S-specific siRNA increased the ratio to 5.1 in the MDA cells (lane 4 compared to lanes 2 and 3). Thus, the ratio of the protein isoforms change consistently as the mRNA variants upon variant-specific siRNA knockdown, supporting the identity of the PRMT5S-like band as endogenous PRMT5S protein in cells.

We next examined the localization of endogenous PRMT5 in HeLa cells by immunostaining of the remaining isoforms (Fig. 13C). In most of the control siRNA-transfected cells (91% ±2.5%, n=3 groups, 100 cells each group, same in the following localization analysis in this figure), the endogenous PRMT5 staining pattern showed punctate spots with light diffusing backgrounds (Fig. 13C, siC panel), consistent with the presence of both the long and short isoforms. Transfection of PRMT5L-specific siRNA significantly reduced the number of strong spots and led to more diffuse distribution in about 59% (±4%, n=3 groups) of cells (Fig. 13C, siPRMT5L panel). In contrast, transfection of PRMT5S-specific siRNA created bright spots and simultaneously reduced the diffusing backgrounds in 68% (±3%, n=3 groups) of cells (Fig. 13, siPRMT5S panel). These changes were also confirmed by immunostaining with a different PRMT5 antibody under higher magnification (Fig. 13D), with the siPRMT5L reducing the punctate staining and siPRMT5S reducing the diffusing staining signals from the control siRNA samples. Thus, the localization changes of endogenous PRMT5 isoforms after splice variant-specific knockdown are consistent with the differences between the Myc-PRMT5 isoforms.

Together, these data support that the exogenous and endogenous PRMT5 isoforms exhibit consistently different subcellular localization. Specifically, the PRMT5S localizes all over the cell, distinct from the highly punctate localization of the PRMT5L.

PRMT5S has an opposite effect on Golgi structure during cell cycle

The distribution of PRMT5L mainly around the nucleus has been shown to colocalize with the Golgi markers P230, ManII, GM130, NSF, GRASP55 and Giantin, and it methylates GM130 to regulate Golgi structure [225]. The Golgi marker localization is cell cycle-dependent, changing from a compacted structure beside one end of the nucleus in interphase to dynamic fragmented structures during mitosis [263-265, 276, 277]. Moreover, PRMT5L itself also promotes mitosis [220, 221, 224]. We thus examined the localization of the PRMT5S and PRMT5L relative to Giantin during cell cycle in HeLa cells without or with PRMT5 knockdown/rescue to dissect the relationship between the PRMT5 isoforms and Golgi structures (Fig. 14A).

The endogenous PRMT5 exhibited cell cycle-dependent changes in subcellular localization (Fig. 14B, upper/control panel). Particularly the bright areas overlap with Giantin. Interestingly, PRMT5 knockdown resulted in a 1.6 fold (±0.1) increase in the area of the compact Giantin staining in the interphase as compared to that in control cells ($p = 2.6E-06$, n=3 experiments, 300 cells each experiment, Fig. 14B, 2^{nd}/shPRMT5 panel). The knockdown also promoted more diffusing distribution of Giantin at metaphase and to a lesser extent at subsequent stages. PRMT5 thus appears to be essential for the compactness of the Golgi apparatus during cell cycle.

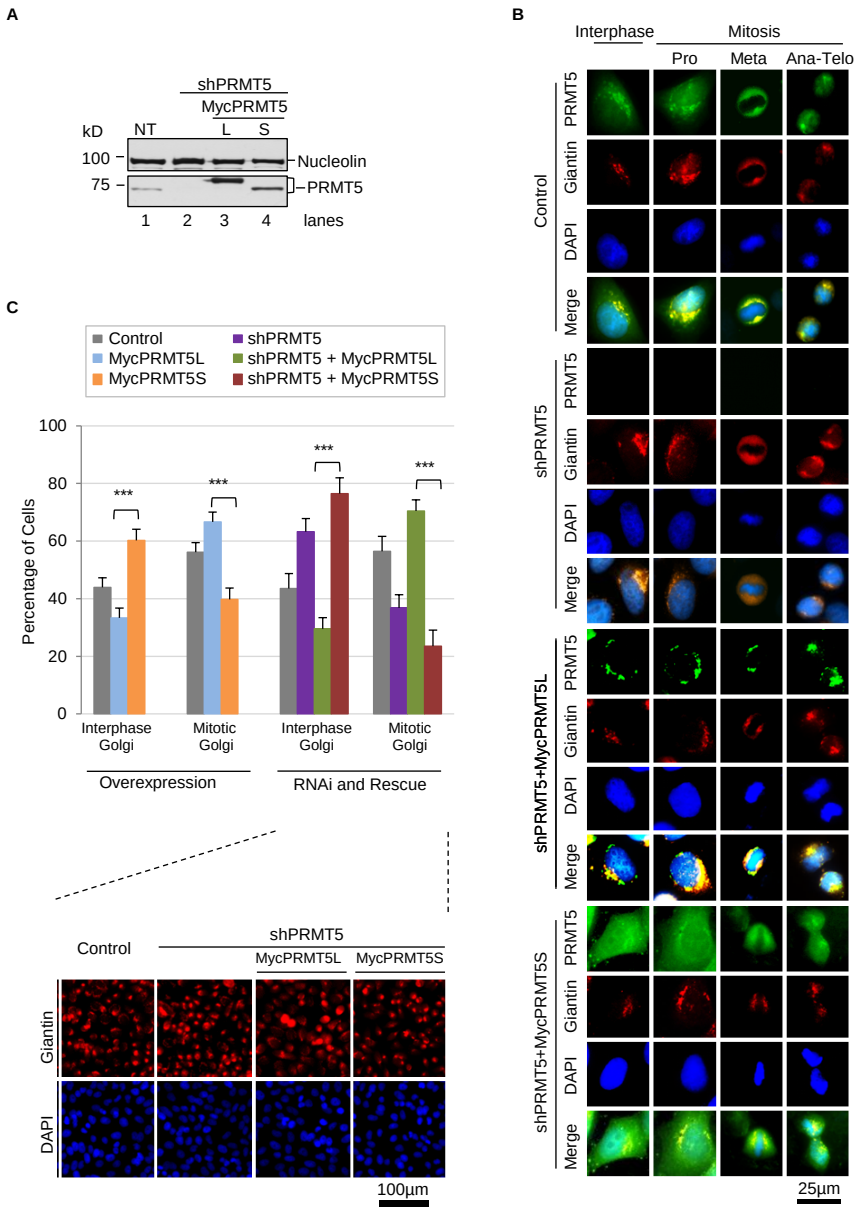


Figure 14. Colocalization of PRMT5 isoforms with cell cycle-specific Golgi patterns and effects on their structure and prevalence. A, Western blot showing the knock down of endogenous PRMT5 and rescue with PRMT5L or S in HeLa cells using lentiviral transduction of shRNA and expression plasmids. **B,** Representative images of HeLa cells treated as in **A** and immunostained using anti-PRMT5 (EPR5772) and anti-Giantin at different stages of cell cycle. DAPI-stained nuclei are merged with PRMT5 and Giantin signals in the bottom row of each panel. Pro: Prophase; Meta: Metaphase; Ana-Telo: Anaphase to Telophase. **C,** Bar graph showing the percentages of HeLa cells with interphase or mitotic Giantin patterns. Interphase: Polarized Giantin beside one end of nucleus; Mitotic: Half circle ribbon, diffused circle or bipolar Giantin representative of later mitosis stages. Left panel shows the percentages of HeLa cells overexpressing either one of the PRMT5 isoforms. Right panel: knockdown/rescue cells treated as in **A.** ***: $p < 0.001$. Below are the representative images of HeLa cells treated as in **A** and immunostained with Giantin antibody.

In PRMT5 knockdown cells expressing the rescue isoforms, Myc-PRMT5L showed dynamic patterns of localization during cell cycle (Fig. 14B, 3rd panel, shPRMT5+MycPRMT5L). This pattern overlaps mostly with that of Giantin. In contrast, the Myc-PRMT5S is more homogenously distributed all over the cell during cell cycle (Fig. 14B, bottom panel, shPRMT5+MycPRMT5S), with its brightest areas overlapping with Giantin as well. Moreover, both rescuing isoforms eliminated the slight changes by PRMT5 knockdown and restored the compactness of the Golgi structures similar to that of the control cells.

Together, the knockdown/rescue data indicate that the PRMT5 isoforms show distinct and dynamic co-localization with the Golgi marker Giantin during cell cycle and they are likely required for the compactness of Golgi structure.

We next examined the effect of PRMT5 isoforms on the prevalence of cells containing interphase compacted or mitotic dynamic Golgi structures. In overexpression experiments, MycPRMT5L increased the percentage of cells with mitotic Golgi whereas MycPRMT5S had the opposite effect (Fig. 14C, left panel). Knock-down of endogenous PRMT5 increased the percentage of cells with interphase compact Golgi from 43.5% to 63.2% (±4.6%, $p = 0.008$, n=3, 300 cells each, same in the following analysis of Golgi structures in this figure). Rescue with MycPRMT5S further increased this cell population to 76.4% (±5.5%, $p = 0.002$). In contrast, rescue with MycPRMT5L had the opposite effect: it decreased this cell population to 29.5% (±3.9%, $p = 0.02$) (Fig. 14C, right panel). These observations are consistent with the opposite effects of the PRMT5 isoforms on mitosis [220, 221, 224, 231]. Thus, PRMT5S has an opposite effect on the number of cells with interphase or mitotic Golgi structure compared to PRMT5L.

PRMT5 isoforms are differentially expressed during differentiation and PRMT5S has an opposite effect on the differentiation of dendritic cells

The differential expression, localization and effect on cell cycle and Golgi structure suggest that the two isoforms may have more downstream effects on cells. We then explored the role of the PRMT5 isoforms during cell differentiation.

We induced LA-N-5 human neuroblastoma cells differentiation using 13-*cis*-retinoic acid (RA), as reported by others [270-272]. Upon differentiation, the percentages of exon skipping significantly decreased from 46% to 34% (Fig. 15A-B, p=0.003, n=3). Thus, PRMT5 variants are differentially regulated during neuronal differentiation.

We also examined the expression of the variants in mouse dendritic cells (DCs). DCs are derived from hematopoietic stem cells, which undergo tightly regulated lineage commitment, differentiation and maturation processes in vivo [278-280]. *In vitro*, a well-established culture condition has been used to generate dendritic cells from adherent bone marrow cells in the presence of GM-CSF cytokine [273]. We used the *in vitro* differentiating DCs as a primary cell model to examine expression of the PRMT5 isoforms. The bone marrow precursor cells did not express any PRMT5 isoform (Fig. 15C, lane 2), while differentiating DCs started to express detectable PRMT5L only on day 3 and at higher levels during the following days (Fig. 15C, lanes 5-7). This suggests that PRMT5 is expressed in a variant-specific and differentiation time-dependent manner in DCs.

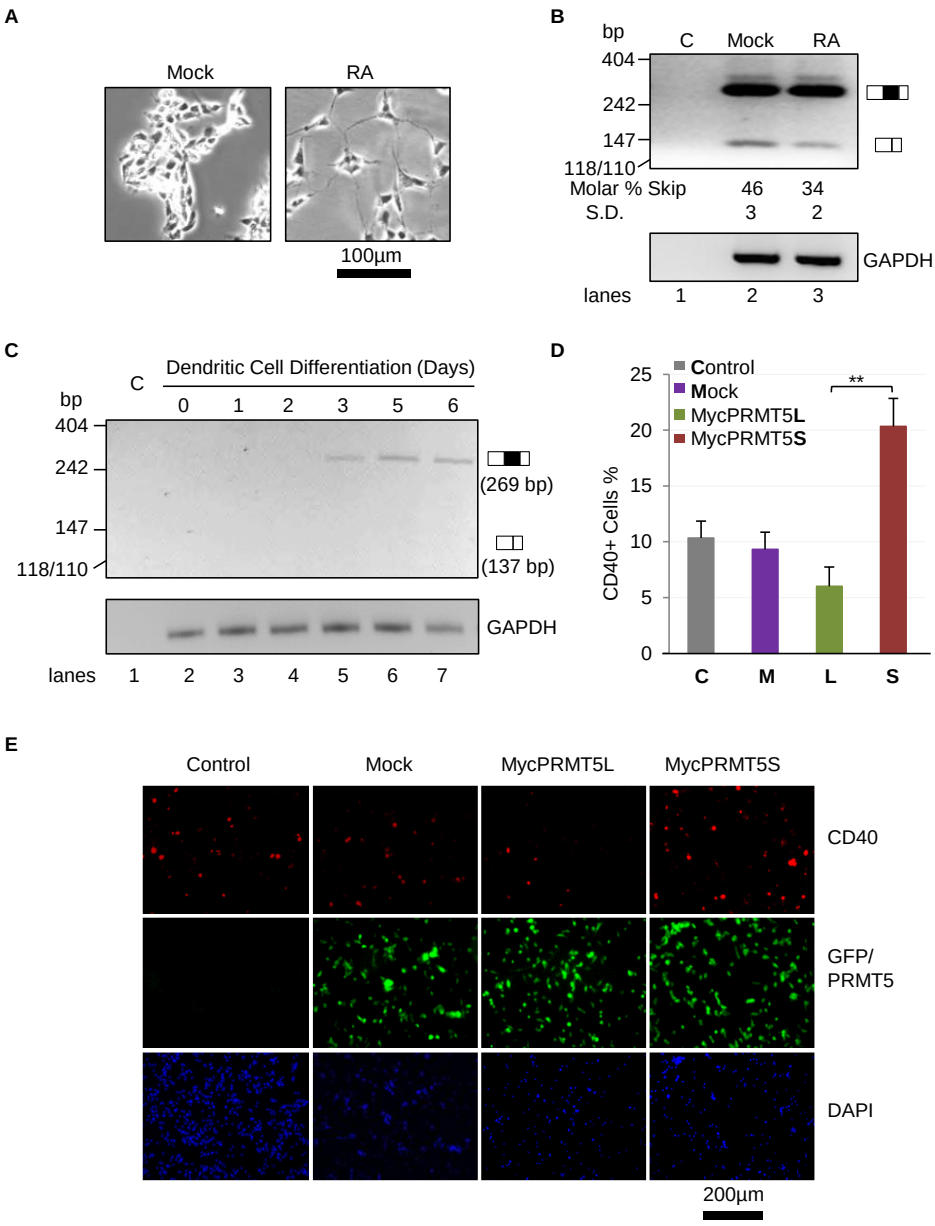


Figure 15. Differential expression and an opposite effect of PRMT5 isoforms on dendritic cell differentiation. A, Representative images of undifferentiated and differentiated LA-N-5 cells. Shown are Mock or 13-*cis*-retinoic acid (RA)-treated cells. **B,** Agarose gel electrophoresis images of semi-quantitative RT-PCR products of PRMT5 splice variants and GAPDH in undifferentiated or differentiated LA-N-5 cells (lanes 2, 3). C (lane 1): PCR control. **C,** Expression of PRMT5 in differentiating dendritic cells. Shown are agarose gel electrophoresis images of RT-PCR products of PRMT5 and GAPDH from precursors at day 0 (lane 2), or following days of differentiation into dendritic cells (lanes 3-7). C (lane 1): PCR control. **D,** Bar graph showing the percentages of CD40 positive control (non-transduced), mock, PRMT5L or S transduced dendritic cells. **: $p < 0.01$. **E,** Representative images of differentiated dendritic cells treated as in **D** and immunostained with anti-Myc and anti-CD40.

We then used this *in vitro* differentiation system to examine the effect of ectopic expression of either one of the MycPRMT5 isoforms on the differentiation/maturation of DCs. The precursor cells were transduced to express the irrelevant EGFP protein, PRMT5S or PRMT5L during the *in vitro* differentiation of DCs, which was defined by the expression of CD40. Non-transduced and mock-transduced cells showed about 10% CD40 positive cells while the PRMT5L-transduced cells showed only 6% CD40 positive cells (±1.7%, p = 0.01, n = 3 groups, 300 cells/group, same in the following analysis of CD40). In contrast, the PRMT5S-transduced cells showed about two-fold increase (±2.5%, p = 8.4E-05) of CD40 positive cells as compared to controls (Fig. 15D-E). Thus, ectopic expression of the PRMT5S variant in precursor cells promotes the differentiation of DCs, while PRMT5L inhibits it.

PRMT5 isoforms regulate genes involved in various cellular processes including apoptosis and differentiation

In our recently reported RNAseq data, we have showed that the two isoforms differentially regulate a group of genes including cell cycle-arresting genes preferentially regulated by PRMTS to promote cell cycle arrest whereas PRMT5L promotes mitosis [231]. To identify genes that are similarly regulated by both isoforms, we analyzed further the RNAseq data and identified 1636 genes in HEK293T cells expressing either isoforms with significant difference of the number of reads from the control sample ($p<0.05$). The number of reads of 1070 genes was increased whereas 566 genes decreased, suggesting that methylation by PRMT5 could either enhance or inhibit the expression of target genes.

We tested the expression of 25 genes by RT-PCR. Of the 22 genes with correct PCR products, all were confirmed to have the same direction of changes as predicted by the reads

from RNAseq analysis [231]. We used the minimal number of reads detectable by RT-PCR (42.4 on average, of 9 samples) as a cut-off value, which resulted in 1072 genes with at least 42 reads. We used these genes for functional clustering analysis using DAVID [274].

Table 3: Functional Clustering of genes commonly regulated by PRMT5L and PRMT5S.

Category	p-Value	Genes	# of Genes
Homeobox	.0003	POU6F1, HOXA13, PRRX1, ZEB2, PAX3, PDX1, HDX, HOXA3, LHX2, HLX, HOXA6, POU2F2, HOXA10, POU3F2, TLX3, BARHL2, ISL1, HOXD9, HOXB4, DLX1, MSX1, HOXB8, HOXB5, HOPX, HMX3	25
Phosphorylation	.0004	HMGCR, PDGFA, CSF1, TGFB3, TRIB3, PRDX2, ZEB2, BDKRB2, KIT, TRIB2, LATS2, SPRY4, IL11, LIF, SPRY2, DUSP19, NOD2, CDKN2B, BCL2, DUSP16, SPRED2, RAPGEF3, FAM129A, PPP1R14C, FRS2, MAP2K6, PPP1R14A, GHR, BCL10, SPHK1, PIM1, CDK5, CDC25A, PROK2, CDKN1A, CCND1, CNTF, CCND2, GADD45G, IL12A, DUSP8, GADD45A	42
Differentiation	.0009	S100A4, FGF19, ALDH1A2, RET, BCL2, CYP26A1, ZEB2, PAX3, ISL1, GDNF	10
Apoptosis	.001	PTGS2, HOXA13, ZMAT3, APH1B, TGFB3, SMNDC1, PCBP4, SQSTM1, AEN, **BCL2**, TICAM1, TNFRSF19, BCL3, INPP5D, RARB, PHLDA3, FOSL1, DEDD2, MAP2K6, KCNMA1, BCL10, AIFM2, KLF10, NR4A1, CDK5, DDIT3, **CASP10**, CDKN1A, EPHA7, NME3, BBC3, ENDOG, UBC, IL12A, RIPK2, NGFR, DCUN1D3, TP53INP1	38
Axonogenesis	.001	RAB3A, **NRXN3**, ISL1, CDK5, CXCL12, **SLIT2**, NUMBL, EPHA4, EPHA7, UNC5A, LHX2, BCL2, UBC, ROBO2, VCAN, SEMA3A, NGFR, UNC5C, SLITRK5, ETV4, GFRA3	21
Transcription	.001	IRAK2, PTHLH, NOD2, ID1, PIM1, PRDX2, CAT, PROX1, DDIT3	9

Notes: Genes with well-studied functions in their category are highlighted in bold.

These regulated genes cluster most significantly for six functions: homeobox, phosphorylation, differentiation, apoptosis, axonogenesis and transcription (Table 3). For instance, 38 genes including *BCL2* (B-cell lymphoma 2) and *CASP10* (Caspase 10, apoptosis-related cysteine peptidase) in apoptosis and 21 genes including *NRXN3* (Neurxin-3-alpha) and *SLIT2* (Slit homolog 2 protein) in axongenesis. Other functions related to the DC differentiation include innate immune response ($p = 0.01$) and semaphorins ($p = 0.01$). For example, the Semaphorin family of genes SEMA3A, D, E, G, and 6D are down-regulated by the PRMT5

isoforms. Interestingly, SEMA4D, which is required for the maturation of dendritic cells [281, 282], is up-regulated by 1.5 times (±0.1, n=3, p=0.02). These clusters are consistent with the diverse functions of PRMT5 in cell growth/mitosis and differentiation [150, 216-218, 221-224]. Therefore, the control of the expression of genes of different pathways by PRMT5 is likely part of a program of changes underlying its effect on cell properties. While the differentially regulated target genes by the two PRMT5 isoforms may contribute to their differential effects on cell mitosis and differentiation [231], the commonly regulated ones perhaps contribute to the control of common targets in cellular processes such as the dynamic changes of Golgi structure and differentiation.

DISCUSSION

PRMT5 is known to have diverse functions in both the cytoplasm and nucleus in cells but the source of this functional diversity has not been fully understood. Our data demonstrate that PRMT5 isoforms created by exon 3/4 alternative splicing are differentially expressed among different cells and during differentiation. They also exhibit distinct subcellular localization and control the compactness of the Golgi apparatus during cell cycle. Moreover, these two isoforms show opposite effects on prevalence of mitotic Golgi and dendritic cell differentiation, perhaps involving the control of the expression of a group of genes.

PRMT5 has complex roles in cells. For instance, it maintains embryonic stem cell pluripotency during development [222, 223], but on the other hand also controls the expression of critical genes in erythroid differentiation [283]. These seemingly opposite processes are thought to be carried out by the same PRMT5 protein. The discovery of the PRMT5S and its

distinct properties from PRMT5L makes it worthwhile to explore the presence and potential role of this novel variant in these cellular processes.

PRMT5S contains the intact methyltransferase domain, which is consistent with its capability to maintain methylation of histone 4, a PRMT5 target [231]; however, its skipped region corresponds to a part of N-terminal TIM barrel domain, which interacts with MEP50, a cofactor of PRMT5 [231, 284]. The interaction with MEP50 can mediate both substrate recognition and regulation of methyltransferase efficiency of PRMT5 [284-287]. Structural insights into PRMT5 particularly in complex with MEP50 indicate that the PRMT5S-skipped region is located in close proximity to the interface of these proteins [285, 286, 288]. Besides MEP50, the N-terminal region of PRMT5 is also involved in interactions with RioK1, pICln and Menin which can modulate PRMT5 activity and substrate specificity [284, 289, 290]. Therefore, altered interactions of PRMT5S with these proteins can contribute to the differential properties and functional consequences of this variant. Moreover, the N-terminal region of PRMT5 seems to harbour non-canonical nuclear exclusion signals [248], likely disrupted in PRMT5S resulting in a distinct diffusing localization pattern (Fig.s 12-14). Interestingly, PRMT5 exhibits subcellular localization-specific effects in several cell types [222, 228, 230, 248, 291-294]. Thus, the position of the region skipped in PRMT5S could potentially affect the protein-protein interaction, enzyme efficiency and subcellular localization of this variant leading to differential functional consequences. It would be interesting to explore if the PRMT5S properties could help explain at least some of the diverse effects of PRMT5 in cells observed previously.

The role of PRMT5 in maintaining the compactness of Golgi structures is supported by the increased area of Golgi marker Giantin at interphase and more diffusing distribution during

mitosis upon PRMT5 knock-down (Fig. 14B). This role is carried out by either of the PRMT5 isoforms with a similar effect (Fig. 14B). Besides their common effect on the compactness of the Golgi apparatus, the PRMT5 isoforms also exhibited opposite effects on the distribution of Golgi structures specific for the different phases of cell cycle. We observed the increase of interphase Golgi structures compacted near one end of the nucleus upon PRMT5 knockdown in cells expressing mainly PRMT5L (Fig.s 13-14). Interestingly in rescuing experiments, the two PRMT5 isoforms have opposite effects on the percentage of the interphase/mitotic Golgi structures (Fig. 14), though they both have protein methylation capability [231]. The different effects are consistent with the increase of the percentage of interphase cells by PRMT5S and mitotic cells by PRMT5L, likely through their differential control of a group of genes involved in cell cycle arrest in similar experiments [231].

The presence of PRMT5S all over the cells indicates that it might be a better candidate for the methylation of nuclear proteins such as histones and p53 [215, 217, 221], as well as cytoplasmic proteins outside of the Golgi apparatus [241, 295, 296].

In summary, we show here that the two PRMT5 variants have distinct expression, localization and effects on Golgi structure and function in cell differentiation. This adds to the repertoire of PRMT proteins by providing a variant candidate for the diverse and often complex effect of PRMT5 in cells.

CONTRIBUTIONS: I am the first author of this manuscript. Dr. Jiuyong Xie provided the conceptual basis of this study and I designed and carried out the experiments. Ms. Manli Zhang helped with BMDC culture and lentivirus transduction. Dr. Sam Kung provided the BMDC

system and gave advice on the dendritic cells differentiation experiment as well as contributed to writing this section and revised manuscript. Dr. David Litchfield provided U2OS cell line and contributed to revising the manuscript. Dr. Lisheng Wang provided H9 cells RNA and contributed to writing relevant text and revising manuscript. I prepared the draft of the manuscript and revised with my advisor Dr. Jiuyong Xie.

ACKNOWLEDGEMENTS: We thank Laszlo Gyenis for help in providing U2OS cells, Richard Gatti for MAL cells, Chongbo Zhao for help in dendritic cell culture, and Niaz Mahmood and Lei Lei for obtaining rat tissues and editing the manuscript. This work is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) to J.X. and partially supported by an operating grant from the Canadian Institutes of Health Research (CIHR) MOP-111224 to LW.

CHAPTER V

Thesis Discussion

Insertion of a regulatory element into a splice site: a novel evolutionary mechanism for the emergence of splice variants

The abundance of alternative exons is much greater in mammals as compared to lower vertebrates or other species [181-183], but the molecular basis of the emergence of alternative exons in mammals over the course of evolution remains largely unknown. Past studies suggest that the accumulation of mutations within regulatory elements and constitutive splice sites underlies evolutionary transition of constitutive exons to species-specific alternative exons in mammals [188-191].

Besides point mutations, our earlier study showed that *de novo* evolutionary insertions of G tract elements (REPA) particularly within a large group of mammalian splice sites can also regulate alternative splicing [49]. The finding of REPA suggests that the *de novo* appearance of regulatory elements also contributed to the creation of species-specific alternative exons at earlier stages of mammalian evolution (Fig. 6). The insertion is different from the point mutations found in splice sites and regulatory elements in the previous study suggesting transition of constitutive exons of lower vertebrates to alternative exons in mammals [190]. The evidence supporting the role of the point mutations is mainly based on computational analyses of genomic data and validation of a group of endogenous exons in several vertebrate species through RT-PCR [188-191]. The effect of point mutations on splicing regulation is further demonstrated in splicing reporter minigene assays with mutagenesis [190]. Based on their regulatory effect it has been proposed that the point mutations lead to suboptimal recognition of

exons [190, 191]. Considering the stepwise assembly of spliceosome [14], it is probable that splice site point mutations directly inhibit the interaction of constitutive factors of early complex, whereas mutations of regulatory elements beyond splice sites are more likely to involve regulatory factors. Although these studies provide valuable insights into the evolutionary dynamics that contributed to the emergence of alternative exons in higher species but the mechanistic understanding of such phenomenon has been lacking.

Consistent with past studies including our previous reports providing a mechanism of splicing repression by REPA CA repeats involving recruitment of hnRNP L [44-48, 81], our current study demonstrates that the evolutionarily emerged REPA G tracts recruit trans-acting regulatory factor hnRNP H/F in a similar way to inhibit U2AF65 binding to Py tract. The inhibition of U2AF65 leads to splicing repression (Fig. 7), apparently at an early stage of the spliceosome assembly. The specific recruitment of hnRNP H/F is also consistent with the binding specificity of the quasi RNA recognition motifs of this family of proteins for G tracts [82-84]. Thus, our study provides a mechanistic understanding of the role of evolved splicing regulatory elements in the emergence of species-specific alternative exons. This mechanism is demonstrated using PRMT5 as a model gene but REPA G tracts of several other tested human genes also showed silencing effect in splicing reporter assays and hnRNP H/F knockdown decreased the skipping of their downstream endogenous exons [49]. Therefore, our findings in this study suggest that the splicing repression through inhibition of U2AF65 binding to Py tract is likely a common mechanism underlying the splicing regulation of a large number of REPA G tract host genes in mammals. Interestingly, of the family of REPA G tract *trans*-acting factors hnRNP H1, H3 and GRSF1 seem to have emerged much earlier than REPA G tracts whereas hnRNP F and H2 were specifically found in mammals [193]. Therefore, differential expression

of these proteins over the course of evolution likely contributed to the regulation mechanism underlying REPA G tract mediated splicing repression.

The species in our comparisons for REPA G tract-mediated PRMT5 exon skipping include human, mouse, rat and zebra fish (Fig.s 6, 11). Although, we also showed the effects of REPA G tracts of other species such as platypus and opossum representing different lengths of G tracts in splicing reporter assays (Fig. 7), but the effect on endogenous exon splicing in these species remains unknown. Particularly, comparisons of endogenous exon splicing between species lacking REPA G tracts and with its first evolutionary appearance along with detection of hnRNP H/F can provide more precise insights into their role in transcriptomic diversity over the course of evolution.

Our investigations have been focused on REPA G tracts mainly in vertebrates. Thorough analyses of the patterns of evolutionary emergence of splicing regulatory elements in a wider range of species can provide global insights into the evolutionary emergence of alternative exons in eukaryotes. Moreover, the regulatory elements positioned beyond splice sites may also have evolutionarily emerged in a similar fashion and contributed to the transcriptomic diversity of mammals.

The differential levels of endogenous PRMT5 exon skipping among human tissues and cell lines are likely because of tissue-specific regulation, which could be due to the differential expression of *trans*-acting factors hnRNP H/F [297, 298]. Moreover, as shown in the responsiveness of splicing regulatory exonic G tracts of *Bcl-xL* to PKC inhibitor Ro-31-8220 [238], the REPA G tracts may also respond to tissue-specific cellular signals and exhibit differential levels of splicing repression.

The existence of REPA G tracts in a large group of genes as well as identification of similarly positioned potential regulatory elements (Fig. 5) [49] suggest that they likely contributed to overall higher occurrence of alternative splicing in mammals. Unlike the mere accumulation of mutations in splice sites or regulatory elements [188-191], this study underscores the role of evolved *de novo* regulatory elements in evolutionary emergence of alternative exons.

Functional impacts of PRMT5 splice variants

Mammalian species particularly humans exhibit highest occurrence of alternative splicing [3, 4, 181-183] and species-specific alternative exons of human and mouse have been predicted to alter the conserved regions of proteins more often than other events of alternative splicing [189]. However, this prediction of more frequent alterations of conserved domains of proteins by species-specific cassette exons is based on computational analyses between only two mammalian species. There has been no further analysis of other species particularly lower vertebrates with experimental data to provide a broader understanding of the role of evolutionarily emerged alternative exons in protein diversity. Therefore, the identification and functional characterization of protein isoforms emerged over the course of evolution has been an outstanding question.

The contribution of REPA G tract mediated splicing repression to the emergence of PRMT5S provides a unique example of an evolutionarily emerged protein isoform with distinct functional consequences. Consistent with the functional diversity exhibited by PRMT5 in the regulation of key cellular processes such as cell cycle and differentiation [213-217, 219-222, 224, 228, 283, 299-304], we showed that splice variant PRMT5S likely contributes to this

diversity as it showed cell cycle arresting and differentiation enhancing effects in opposite to the full length protein. Therefore, it is likely that the functional diversity of PRMT5 is achieved in many cases through PRMT5 alternative splicing.

Alternative splicing can produce protein isoforms with minute changes to completely opposite functional consequences [260]. Interestingly, splice variants of a PRMT family member exhibit differential localization, substrate specificity and methyltransferase activity [230], but PRMT5 splice variants have not been studied in the past for their differential properties. We showed here that PRMT5S exhibits distinct cellular localization (Fig. 12-13). Therefore alternative splicing attributes the diverse localization patterns to PRMT5 [222, 228, 248, 291-293].

The distinct subcellular localization patterns of PRMT5S suggest a link to its opposite functional consequences (Fig.s 9, 14-15), which is consistent with its localization-specific functions in several cell types [222, 228, 248, 291-294]. Interestingly, the differential subcellular localization of PRMT5 is associated with its preferential regulation of differentiation or proliferation [222, 291]. Moreover, cytoplasmic or nuclear PRMT5 proteins were found to be associated with high grade or carcinoid lung cancer tumors respectively [293, 294]. Similar correlation of cellular localization was consistent for germ cell tumors and melanoma tissue as compared to their respective normal tissues [291, 292, 294]. This trend was also observed for benign or premalignant prostate cancer tissues [228, 248, 294]. However, precise molecular mechanisms governing the opposite effects of differential localization patterns of PRMT5 isoforms remain to be explored.

The regulation of cell cycle and differentiation by PRMT5 often involves the direct methylation or transcriptional regulation of relevant genes [214, 215, 219-222, 224, 228, 283, 299-303]. Our data consistently showed that the opposite effects of PRMT5S on cell cycle and differentiation as compared to full length isoform are likely at least in part through preferential regulation of gene expression (Fig. 10). There could also be differential methylation target profiles of the PRMT5 isoforms because of distinct spatial distribution. However, the molecular basis of these opposite functional consequences requires further investigation, particularly structural differences of PRMT5S and identification of its potential distinct targets with their functions.

The peptide skipped in PRMT5S is in the N-terminal region of the protein (Fig. 8). This region of PRMT5 was suggested to have non-canonical nuclear export signal (NES) responsible for cytoplasmic localization [248]. Also, alternative splicing in the N-terminal region of PRMT1 creates splice variants with differential cellular localization [230]. Therefore, it is plausible that PRMT5S has distinct subcellular localization due to skipping of the peptide from the N terminal region. However, whether this skipping also accounts for substrate-specificity like PRMT1 remains to be verified. Moreover, the protein structure of PRMT5 from human and other species indicates that the skipped peptide of PRMT5S is near the interface with a cofactor MEP50, suggesting that its inclusion/exclusion from PRMT5 might regulate substrate interaction and specificity (Fig. 8) [285, 286, 288]. Moreover, the N-terminal region is also involved in interactions with other factors that can alter the substrate specificity and methyltransferase efficiency of PRMT5 [284, 289, 290]. Since the skipped part belongs to a domain critical for interaction with other proteins, this strongly suggests that the PRMT5S may interact differently with other proteins and is likely to have unique profiles of target proteins. Both the cellular

localization and potential differences in protein-protein interactions could provide a reasonable explanation for the opposite effects of the short splice variant on cells.

Although in our knockdown/rescue assays both isoforms showed similar effect on the methylation status of PRMT5 target histone 4 arginine 3 (Fig. 8), but the possibility of differential methylation of other targets still remains because of the potential interaction differences with the factors that can alter the substrate specificity and methyltransferase efficiency [284-286, 288-290].

Besides the distinct localization patterns the two PRMT5 isoforms are also differentially expressed in different tissues and cells lines (Fig. 11). The differential expression suggests that the ratio of isoforms may help achieve a fine balance of their apparently opposite functions in a spatial and temporal manner. Moreover, a large group of genes regulated by both isoforms significantly cluster in functional categories such as homeobox, differentiation and axonogenesis (Table 3), which indicates that the two isoforms may utilize these specific genes in different contexts to regulate differentiation and proliferation.

The colocalization of PRMT5 with Golgi apparatus and methylation of its component GM130 to maintain Golgi structure [225], prompted us to investigate the links between differential colocalization patterns of PRMT5 isoforms with Golgi apparatus during cell cycle and their effects on the maintenance of its structure. We consistently observed the colocalization of PRMT5L with Golgi marker Giantin, although the patterns are dynamic at different stages of cell cycle showing partial colocalization at certain stages (Fig. 14). However, detailed analysis of colocalization patterns requires further investigation by adding more Golgi markers such as GM130, P230 and ManII [225]. Using confocal microscopy to precisely determine the

localization of PRMT5L with respect to mitochondria (data not shown), our preliminary observations suggested that the apparent colocalization with an organelle may not overlap entirely but only in certain areas. Therefore, a detailed confocal analysis of localization of both PRMT5 isoforms particularly with respect to nucleus and Golgi can provide more precise location of proteins and help explain the functional differences. Since PRMT5S is diffusely distributed throughout the cell as compared to punctate distribution of PRMT5L around nucleus (Fig. 12-13), evolutionarily emerged alternative splicing may have enabled PRMT5 to move out of the organelles/areas around the nucleus. The ability of both rescuing PRMT5 isoforms to restore the slight increase in Golgi area in cells knocked down of PRMT5 (Fig. 14) suggests that both isoforms likely methylate GM130, which is a critical posttranslational modification for the maintenance of Golgi structure [225]. Further studies using antibodies capable of detecting *in vivo* methylation of GM130 can more specifically explain the role of this posttranslational modification in the maintenance of Golgi structure. However, regulation of the cell cycle-specific structural dynamics of Golgi [263-265, 276, 277] by the PRMT5 isoforms remains to be distinguished from the role of GM130 methylation. The two effects could be interlinked in a complex way to orchestrate the morphological changes of Golgi during mitosis.

PRMT5 accelerates cell cycle progression and promotes tumorigenesis [213, 220, 221, 224], and exhibits oncogene-like properties by repressing the activities of tumor suppressor genes [181, 215, 219-221, 299]. The elevated levels of PRMT5 in a number of cancers correlate its diverse functions with different types of cancers [305]. Moreover, PRMT5 may also play roles in infectious diseases [284, 306, 307]. It methylates the nuclear antigen protein of Epstein-Barr virus to stimulate gene expression, which suggests its role in latent infection [284, 306, 307]. The proteins p28 and p30 encoded by human T lymphotropic virus (HTLV) specifically

interact with host PRMT5 and viral gene expression of HTLV-2 is decreased upon reduction of host PRMT5 levels [308]. Furthermore, GATA4 in cardiomyocytes is also methylated by PRMT5, which inhibits the activity of GATA4 to promote transcription [284, 309]. Therefore, the implications of PRMT5 in a wide range of pathological conditions make it an important therapeutic target. With the availability of structure of PRMT5/MEP50, efforts are being made to target the activity of PRMT5 in cancers (patent #20140221345 and #WO2011079236A1) [285, 286, 294]. Interestingly, epizyme inhibitor EPZ0004777 that is directed against lysine methyltransferase also inhibited PRMT5 which suggests that its activity includes effect on PRMT5 [284, 310, 311]. Therefore, an important concern for such pharmacological treatments would be selective inhibition of PRMT5. Considering the opposite effects of PRMT5 isoforms the therapeutic approaches involving the functional characterization and modulation of ratio of isoforms in a particular pathological condition can be promising.

Together, the findings in this study provide a functional link between the evolutionary emergence of a splice variant and how it contributes to the protein diversity. This also suggests how the evolutionary emergence of splicing regulatory elements could have contributed to the higher abundance of alternative splicing, functional diversity and organismal complexity in mammals. Moreover, since the REPA $(G)_{5-8}$ and other similar elements are found in a large group of human genes, together they would likely have a broader impact on the evolutionary emergence of species-specific alternative exons in mammals and functional divergence. The localization differences of PRMT5 have been previously correlated with differential functional consequences. In this study we uncover the role of evolutionarily emerged splice variant of PRMT5 with distinct cellular localization in functional diversity of PRMT5 and provide a novel underlying mechanism.

CHAPTER VI

Future Directions

The identification of REPA G tracts and elucidation of their mechanism of splicing inhibition bring up the question of potential splicing regulatory role of other REPA identified in a large group of human genes [45, 47, 49, 193, 312]. The mechanism of action of REPA studied so far showed that they recruit *trans*-acting factors to inhibit early spliceosome assembly and splicing. Interestingly, similarly positioned elements can act as enhancers or repressors in a cell type specific manner [33]. Therefore, the study of other REPA for their potentially novel ways of splicing regulation could be an important direction for future research. Moreover, upstream signals are involved in splicing regulation by some REPA [193, 312]. Since splicing regulatory G tracts of *Bcl-xL* showed responsiveness to PKC inhibitor Ro-31-8220 [238], therefore the splicing repression by REPA G tracts may be modulated in a cell signal dependent manner. This provides an interesting future direction to explore the role of cell signalling in REPA G tracts mediated splicing regulation.

The splicing repression by REPA G tracts contributes to the emergence of PRMT5S which shows differential effects on cells. Particularly, PRMT5S shows distinct subcellular localization but the question of how it could contribute to the differential gene expression regulation as well as cell cycle arrest remains to be answered. Moreover, the functional analyses of PRMT5 isoforms in our study were based on transient expression and knockdown/rescue over a period of several days. However, the consequences of longer periods of PRMT5 knockdown and stable expression of rescuing isoforms remain to be investigated. The effects of longer

periods of knockdown/rescue on cell viability by assaying autophagy/apoptosis are interesting directions for future investigations.

PRMT5 is known to interact with other proteins to specifically methylate its substrates which regulate diverse cellular processes [168, 215, 284-286, 288-290]. Therefore, it would be interesting to investigate whether PRMT5S interacts with distinct proteins and how it could have differential effects on cells. Besides interaction with other proteins an important question is to determine the quantitative differences in methyltransferase activity of the two isoforms. Since the differential interactions with other factors can alter the methyltransferase activity [168, 215, 284-286, 288-290], therefore *in vitro* methylation assays would be useful to elucidate the enzymatic activities of the two isoforms. Besides the *in vitro* methylation assays with specific known targets, the identification of complete target profiles of the two isoforms will provide very important insights into their differential functional consequences.

PRMT5 shows localization-dependent differential regulation of differentiation particularly at early stages of embryonic development [222, 291]. Importantly, cytoplasmic PRMT5 specifically acts in the maintenance of pluripotency whereas in nucleus it correlates with differentiated cells [222]. Since the two isoforms exhibit differential localization with PRMT5S diffusely distributed throughout the cells as compared to punctate distribution of PRMT5L near nucleus, therefore an important question is to identify and elucidate the differential roles of either one of these isoforms during early development.

Besides the embryonic development the cytoplasmic or nuclear localization of PRMT5 has been correlated with different outcomes of several types of cancer cells [222, 228, 248, 291-293]. This provides an important future direction to investigate the link between expression of

PRMT5 isoforms and differential localization patterns in these cancer cells, and molecular mechanisms underlying different outcomes of cancers.

Because of the association of PRMT5 with a wide range of pathological conditions such as cancers, infectious and heart diseases [181, 213, 215, 219-221, 224, 284, 299, 305-309], it is being considered as a therapeutic target of pharmacological treatments [284, 294, 310, 311]. However, the splice variants of PRMT5 exhibit opposite functions and a balance of the two variants may help provide beneficial effects in the pathological conditions. Therefore, an interesting future direction could be to characterize the roles of either one of the splice variants in different pathological conditions and direct the therapeutic approaches towards achieving a balance of splice variants through modulation of alternative splicing.

Besides PRMT5, many other human genes also harbor similar REPA G tracts. A group of such genes showed similar regulation of splicing by these elements as observed for PRMT5 [49]. However, the consequences of this regulation on splicing outcome as well as on the functions of those genes remain largely unknown. This raises an important question of how the REPA G tract-mediated splicing repression contributes to the functional variations of the other host genes. Thus, an important direction for future research would be to study the protein isoforms created from the other host genes and their differential effects on the cells as well as underlying molecular mechanisms.

CHAPTER VII

Conclusion

REPA G tracts $(G)_{5-8}$ were identified in 130 human genes enriched in several functional clusters particularly 30 of them in cell growth and proliferation. The PRMT5 REPA G tract is evolved in mammals and is bound by hnRNP H/F to inhibit U2AF65 and splicing, contributing to the emergence of a shorter splice variant of PRMT5 (PRMT5S) in humans. Contrary to the full length PRMT5, it inhibits the cell cycle progression and specifically controls the expression of a group of cell cycle arresting genes.

This provides a novel example of how an evolutionarily emerged element regulates splicing and contributes to the proteomic diversity, likely common to the splicing regulation of a large number of human genes.

The PRMT5 splice variants created by the REPA G tract-mediated regulation are differentially expressed in various cell types and tissues, and exhibit distinct subcellular localization patterns. Moreover, the previously uncharacterized PRMT5S shows opposite effects on Golgi fragmentation during cell cycle as well as dendritic cell differentiation as compared to its full length counterpart. The expression of a large number of genes involved in various cellular functions is regulated by the PRMT5 splice variants.

The differential functional consequences of PRMT5 splice variants contribute to the understanding of the molecular basis of diverse roles of PRMT5 in cells. The counteracting effects and differential expression of PRMT5S likely determine the balance of complex PRMT5 effects on cells in a spatial and temporal manner.

Taken together, these studies contribute to the understanding of evolutionarily emerged RNA elements in the splicing regulation at 3′SS. This splicing regulation contributes to the functional diversity of PRMT5 and likely a large group of human genes.

References

1. **Maniatis T, Tasic B: Alternative pre-mRNA splicing and proteome expansion in metazoans. *Nature* 2002, 418(6894):236-243.**
2. **Nilsen TW, Graveley BR: Expansion of the eukaryotic proteome by alternative splicing. *Nature* 2010, 463(7280):457-463.**
3. **Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ: Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat Genet* 2008, 40(12):1413-1415.**
4. **Wang ET, Sandberg R, Luo S, Khrebtukova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB: Alternative isoform regulation in human tissue transcriptomes. *Nature* 2008, 456(7221):470-476.**
5. **Singh RK, Cooper TA: Pre-mRNA splicing in disease and therapeutics. *Trends in molecular medicine* 2012, 18(8):472-482.**
6. **Feng D, Xie J: Aberrant splicing in neurological diseases. *Wiley interdisciplinary reviews RNA* 2013, 4(6):631-649.**
7. **Umen JG, Guthrie C: The second catalytic step of pre-mRNA splicing. *RNA* 1995, 1(9):869-885.**
8. **Huang T, Vilardell J, Query CC: Pre-spliceosome formation in S.pombe requires a stable complex of SF1-U2AF(59)-U2AF(23). *EMBO J* 2002, 21(20):5516-5526.**
9. **Hollins C, Zorio DA, MacMorris M, Blumenthal T: U2AF binding selects for the high conservation of the C. elegans 3' splice site. *RNA* 2005, 11(3):248-253.**
10. **Blumenthal T, Steward K: RNA Processing and Gene Structure. 1997.**
11. **Kent WJ, Zahler AM: Conservation, regulation, synteny, and introns in a large-scale C. briggsae-C. elegans genomic alignment. *Genome Res* 2000, 10(8):1115-1125.**
12. **Reed R: The organization of 3' splice-site sequences in mammalian introns. *Genes Dev* 1989, 3(12B):2113-2123.**
13. **Smith CW, Porro EB, Patton JG, Nadal-Ginard B: Scanning from an independently specified branch point defines the 3' splice site of mammalian introns. *Nature* 1989, 342(6247):243-247.**
14. **Will CL, Luhrmann R: Spliceosome structure and function. *Cold Spring Harbor perspectives in biology* 2011, 3(7).**
15. **Berget SM: Exon recognition in vertebrate splicing. *J Biol Chem* 1995, 270(6):2411-2414.**
16. **Black DL: Mechanisms of alternative pre-messenger RNA splicing. *Annu Rev Biochem* 2003, 72:291-336.**
17. **Chen M, Manley JL: Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nat Rev Mol Cell Biol* 2009, 10(11):741-754.**
18. **Buratti E, Baralle FE: Influence of RNA secondary structure on the pre-mRNA splicing process. *Mol Cell Biol* 2004, 24(24):10505-10514.**
19. **Wu T, Fu XD: Genomic Functions of U2AF in Constitutive and Regulated Splicing. *RNA Biol* 2015:0.**
20. **Shao C, Yang B, Wu T, Huang J, Tang P, Zhou Y, Zhou J, Qiu J, Jiang L, Li H *et al*: Mechanisms for U2AF to define 3' splice sites and regulate alternative splicing in the human genome. *Nat Struct Mol Biol* 2014, 21(11):997-1005.**
21. **Saltzman AL, Pan Q, Blencowe BJ: Regulation of alternative splicing by the core spliceosomal machinery. *Genes Dev* 2011, 25(4):373-384.**

22. Rosel TD, Hung LH, Medenbach J, Donde K, Starke S, Benes V, Ratsch G, Bindereif A: RNA-Seq analysis in mutant zebrafish reveals role of U1C protein in alternative splicing regulation. *EMBO J* 2011, 30(10):1965-1976.
23. Sauliere J, Sureau A, Expert-Bezancon A, Marie J: The polypyrimidine tract binding protein (PTB) represses splicing of exon 6B from the beta-tropomyosin pre-mRNA by directly interfering with the binding of the U2AF65 subunit. *Mol Cell Biol* 2006, 26(23):8755-8769.
24. Gooding C, Roberts GC, Moreau G, Nadal-Ginard B, Smith CW: Smooth muscle-specific switching of alpha-tropomyosin mutually exclusive exon selection by specific inhibition of the strong default exon. *EMBO J* 1994, 13(16):3861-3872.
25. Perez I, Lin CH, McAfee JG, Patton JG: Mutation of PTB binding sites causes misregulation of alternative 3' splice site selection in vivo. *RNA* 1997, 3(7):764-778.
26. Smith CW, Nadal-Ginard B: Mutually exclusive splicing of alpha-tropomyosin exons enforced by an unusual lariat branch point location: implications for constitutive splicing. *Cell* 1989, 56(5):749-758.
27. Chan RC, Black DL: The polypyrimidine tract binding protein binds upstream of neural cell-specific c-src exon N1 to repress the splicing of the intron downstream. *Mol Cell Biol* 1997, 17(8):4667-4676.
28. Chou MY, Underwood JG, Nikolic J, Luu MH, Black DL: Multisite RNA binding and release of polypyrimidine tract binding protein during the regulation of c-src neural-specific splicing. *Mol Cell* 2000, 5(6):949-957.
29. Sharma S, Falick AM, Black DL: Polypyrimidine tract binding protein blocks the 5' splice site-dependent assembly of U2AF and the prespliceosomal E complex. *Mol Cell* 2005, 19(4):485-496.
30. Matlin AJ, Southby J, Gooding C, Smith CW: Repression of alpha-actinin SM exon splicing by assisted binding of PTB to the polypyrimidine tract. *RNA* 2007, 13(8):1214-1223.
31. Southby J, Gooding C, Smith CW: Polypyrimidine tract binding protein functions as a repressor to regulate alternative splicing of alpha-actinin mutally exclusive exons. *Mol Cell Biol* 1999, 19(4):2699-2711.
32. Zhang L, Liu W, Grabowski PJ: Coordinate repression of a trio of neuron-specific splicing events by the splicing regulator PTB. *RNA (New York)* 1999, 5(1):117-130.
33. Raj B, Irimia M, Braunschweig U, Sterne-Weiler T, O'Hanlon D, Lin ZY, Chen GI, Easton LE, Ule J, Gingras AC *et al*: A global regulatory mechanism for activating an exon network required for neurogenesis. *Mol Cell* 2014, 56(1):90-103.
34. Singh R, Banerjee H, Green MR: Differential recognition of the polypyrimidine-tract by the general splicing factor U2AF65 and the splicing repressor sex-lethal. *RNA* 2000, 6(6):901-911.
35. Valcarcel J, Singh R, Zamore PD, Green MR: The protein Sex-lethal antagonizes the splicing factor U2AF to regulate alternative splicing of transformer pre-mRNA. *Nature* 1993, 362(6416):171-175.
36. Lisbin MJ, Qiu J, White K: The neuron-specific RNA-binding protein ELAV regulates neuroglian alternative splicing in neurons and binds directly to its pre-mRNA. *Genes Dev* 2001, 15(19):2546-2561.
37. Anyanful A, Ono K, Johnsen RC, Ly H, Jensen V, Baillie DL, Ono S: The RNA-binding protein SUP-12 controls muscle-specific splicing of the ADF/cofilin pre-mRNA in C. elegans. *J Cell Biol* 2004, 167(4):639-647.
38. Ayala YM, Pagani F, Baralle FE: TDP43 depletion rescues aberrant CFTR exon 9 skipping. *FEBS Lett* 2006, 580(5):1339-1344.

39. Buratti E, Dork T, Zuccato E, Pagani F, Romano M, Baralle FE: Nuclear factor TDP-43 and SR proteins promote in vitro and in vivo CFTR exon 9 skipping. *EMBO J* 2001, 20(7):1774-1784.

40. Ho TH, Charlet BN, Poulos MG, Singh G, Swanson MS, Cooper TA: Muscleblind proteins regulate alternative splicing. *EMBO J* 2004, 23(15):3103-3112.

41. Warf MB, Diegel JV, von Hippel PH, Berglund JA: The protein factors MBNL1 and U2AF65 bind alternative RNA structures to regulate splicing. *Proc Natl Acad Sci U S A* 2009, 106(23):9203-9208.

42. Cao W, Razanau A, Feng D, Lobo VG, Xie J: Control of alternative splicing by forskolin through hnRNP K during neuronal differentiation. *Nucleic Acids Res* 2012, 40(16):8059-8071.

43. Zarnack K, Konig J, Tajnik M, Martincorena I, Eustermann S, Stevant I, Reyes A, Anders S, Luscombe NM, Ule J: Direct competition between hnRNP C and U2AF65 protects the transcriptome from the exonization of Alu elements. *Cell* 2013, 152(3):453-466.

44. Loh TJ, Cho S, Moon H, Jang HN, Williams DR, Jung DW, Kim IC, Ghigna C, Biamonti G, Zheng X *et al*: hnRNP L inhibits CD44 V exon splicing through interacting with its upstream intron. *Biochim Biophys Acta* 2015.

45. Liu G, Razanau A, Hai Y, Yu J, Sohail M, Lobo VG, Chu J, Kung SK, Xie J: A conserved serine of heterogeneous nuclear ribonucleoprotein L (hnRNP L) mediates depolarization-regulated alternative splicing of potassium channels. *J Biol Chem* 2012, 287(27):22709-22716.

46. Xie J, Jan C, Stoilov P, Park J, Black DL: A consensus CaMK IV-responsive RNA sequence mediates regulation of alternative exons in neurons. *RNA* 2005, 11(12):1825-1834.

47. Yu J, Hai Y, Liu G, Fang T, Kung SK, Xie J: The heterogeneous nuclear ribonucleoprotein L is an essential component in the Ca2+/calmodulin-dependent protein kinase IV-regulated alternative splicing through cytidine-adenosine repeats. *J Biol Chem* 2009, 284(3):1505-1513.

48. Heiner M, Hui J, Schreiner S, Hung LH, Bindereif A: HnRNP L-mediated regulation of mammalian alternative splicing by interference with splice site recognition. *RNA Biol* 2010, 7(1):56-64.

49. Sohail M, Cao W, Mahmood N, Myschyshyn M, Hong SP, Xie J: Evolutionarily emerged G tracts between the polypyrimidine tract and 3' AG are splicing silencers enriched in genes involved in cancer. *BMC Genomics* 2014, 15:1143.

50. Tange TO, Damgaard CK, Guth S, Valcarcel J, Kjems J: The hnRNP A1 protein regulates HIV-1 tat splicing via a novel intron silencer element. *EMBO J* 2001, 20(20):5748-5758.

51. Dembowski JA, Grabowski PJ: The CUGBP2 splicing factor regulates an ensemble of branchpoints from perimeter binding sites with implications for autoregulation. *PLoS Genet* 2009, 5(8):e1000595.

52. Zong FY, Fu X, Wei WJ, Luo YG, Heiner M, Cao LJ, Fang Z, Fang R, Lu D, Ji H *et al*: The RNA-binding protein QKI suppresses cancer-associated aberrant splicing. *PLoS Genet* 2014, 10(4):e1004289.

53. Corsini L, Bonnal S, Basquin J, Hothorn M, Scheffzek K, Valcarcel J, Sattler M: U2AF-homology motif interactions are required for alternative splicing regulation by SPF45. *Nat Struct Mol Biol* 2007, 14(7):620-629.

54. Taliaferro JM, Alvarez N, Green RE, Blanchette M, Rio DC: Evolution of a tissue-specific splicing network. *Genes Dev* 2011, 25(6):608-620.

55. Wang J, Bell LR: The Sex-lethal amino terminus mediates cooperative interactions in RNA binding and is essential for splicing regulation. *Genes Dev* 1994, 8(17):2072-2085.

56. Horabin JI, Schedl P: Sex-lethal autoregulation requires multiple cis-acting elements upstream and downstream of the male exon and appears to depend largely on controlling the use of the male exon 5' splice site. *Mol Cell Biol* 1993, 13(12):7734-7746.

57. Sakamoto H, Inoue K, Higuchi I, Ono Y, Shimura Y: Control of Drosophila Sex-lethal pre-mRNA splicing by its own female-specific product. *Nucleic Acids Res* 1992, 20(21):5533-5540.

58. Konig J, Zarnack K, Rot G, Curk T, Kayikci M, Zupan B, Turner DJ, Luscombe NM, Ule J: iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nat Struct Mol Biol* 2010, 17(7):909-915.

59. Gromak N, Matlin AJ, Cooper TA, Smith CW: Antagonistic regulation of alpha-actinin alternative splicing by CELF proteins and polypyrimidine tract binding protein. *RNA* 2003, 9(4):443-456.

60. Charlet BN, Logan P, Singh G, Cooper TA: Dynamic antagonism between ETR-3 and PTB regulates cell type-specific alternative splicing. *Mol Cell* 2002, 9(3):649-658.

61. Soares LM, Zanier K, Mackereth C, Sattler M, Valcarcel J: Intron removal requires proofreading of U2AF/3' splice site recognition by DEK. *Science* 2006, 312(5782):1961-1965.

62. Dowhan DH, Hong EP, Auboeuf D, Dennis AP, Wilson MM, Berget SM, O'Malley BW: Steroid hormone receptor coactivation and alternative RNA splicing by U2AF65-related proteins CAPERalpha and CAPERbeta. *Mol Cell* 2005, 17(3):429-439.

63. Huang G, Zhou Z, Wang H, Kleinerman ES: CAPER-alpha alternative splicing regulates the expression of vascular endothelial growth factor(1)(6)(5) in Ewing sarcoma cells. *Cancer* 2012, 118(8):2106-2116.

64. Tavanez JP, Madl T, Kooshapur H, Sattler M, Valcarcel J: hnRNP A1 proofreads 3' splice site recognition by U2AF. *Mol Cell* 2012, 45(3):314-329.

65. Wei WJ, Mu SR, Heiner M, Fu X, Cao LJ, Gong XF, Bindereif A, Hui J: YB-1 binds to CAUC motifs and stimulates exon inclusion by enhancing the recruitment of U2AF to weak polypyrimidine tracts. *Nucleic Acids Res* 2012, 40(17):8622-8636.

66. Tisserant A, Konig H: Signal-regulated Pre-mRNA occupancy by the general splicing factor U2AF. *PLoS One* 2008, 3(1):e1418.

67. Xue Y, Zhou Y, Wu T, Zhu T, Ji X, Kwon YS, Zhang C, Yeo G, Black DL, Sun H *et al*: Genome-wide analysis of PTB-RNA interactions reveals a strategy used by the general splicing repressor to modulate exon inclusion or skipping. *Mol Cell* 2009, 36(6):996-1006.

68. Olson S, Blanchette M, Park J, Savva Y, Yeo GW, Yeakley JM, Rio DC, Graveley BR: A regulator of Dscam mutually exclusive splicing fidelity. *Nat Struct Mol Biol* 2007, 14(12):1134-1140.

69. May GE, Olson S, McManus CJ, Graveley BR: Competing RNA secondary structures are required for mutually exclusive splicing of the Dscam exon 6 cluster. *RNA* 2011, 17(2):222-229.

70. Graveley BR: Mutually exclusive splicing of the insect Dscam pre-mRNA directed by competing intronic RNA secondary structures. *Cell* 2005, 123(1):65-73.

71. Wu JY, Maniatis T: Specific interactions between proteins implicated in splice site selection and regulated alternative splicing. *Cell* 1993, 75(6):1061-1070.

72. Fu XD, Maniatis T: The 35-kDa mammalian splicing factor SC35 mediates specific interactions between U1 and U2 small nuclear ribonucleoprotein particles at the 3' splice site. *Proc Natl Acad Sci U S A* 1992, 89(5):1725-1729.

73. Hertel KJ, Maniatis T: The function of multisite splicing enhancers. *Mol Cell* 1998, 1(3):449-455.

74. Graveley BR, Hertel KJ, Maniatis T: A systematic analysis of the factors that determine the strength of pre-mRNA splicing enhancers. *EMBO J* 1998, 17(22):6747-6756.

75. Nasim FU, Hutchison S, Cordeau M, Chabot B: High-affinity hnRNP A1 binding sites and duplex-forming inverted repeats have similar effects on 5' splice site selection in support of a common looping out and repression mechanism. *RNA* 2002, 8(8):1078-1089.

76. Chabot B, Blanchette M, Lapierre I, La Branche H: An intron element modulating 5' splice site selection in the hnRNP A1 pre-mRNA interacts with hnRNP A1. *Mol Cell Biol* 1997, 17(4):1776-1786.
77. Blanchette M, Chabot B: Modulation of exon skipping by high-affinity hnRNP A1-binding sites and by intron elements that repress splice site utilization. *EMBO J* 1999, 18(7):1939-1952.
78. Martinez-Contreras R, Fisette JF, Nasim FU, Madden R, Cordeau M, Chabot B: Intronic binding sites for hnRNP A/B and hnRNP F/H proteins stimulate pre-mRNA splicing. *PLoS Biol* 2006, 4(2):e21.
79. Mulligan GJ, Guo W, Wormsley S, Helfman DM: Polypyrimidine tract binding protein interacts with sequences involved in alternative splicing of beta-tropomyosin pre-mRNA. *J Biol Chem* 1992, 267(35):25480-25487.
80. Singh R, Valcarcel J, Green MR: Distinct binding specificities and functions of higher eukaryotic polypyrimidine tract-binding proteins. *Science* 1995, 268(5214):1173-1176.
81. Li H, Liu G, Yu J, Cao W, Lobo VG, Xie J: In vivo selection of kinase-responsive RNA elements controlling alternative splicing. *J Biol Chem* 2009, 284(24):16191-16201.
82. Caputi M, Zahler AM: Determination of the RNA binding specificity of the heterogeneous nuclear ribonucleoprotein (hnRNP) H/H'/F/2H9 family. *J Biol Chem* 2001, 276(47):43850-43859.
83. Dominguez C, Allain FH: NMR structure of the three quasi RNA recognition motifs (qRRMs) of human hnRNP F and interaction studies with Bcl-x G-tract RNA: a novel mode of RNA recognition. *Nucleic Acids Res* 2006, 34(13):3634-3645.
84. Garneau D, Revil T, Fisette JF, Chabot B: Heterogeneous nuclear ribonucleoprotein F/H proteins modulate the alternative splicing of the apoptotic mediator Bcl-x. *J Biol Chem* 2005, 280(24):22641-22650.
85. Chou MY, Rooke N, Turck CW, Black DL: hnRNP H is a component of a splicing enhancer complex that activates a c-src alternative exon in neuronal cells. *Mol Cell Biol* 1999, 19(1):69-77.
86. Schaub MC, Lopez SR, Caputi M: Members of the heterogeneous nuclear ribonucleoprotein H family activate splicing of an HIV-1 splicing substrate by promoting formation of ATP-dependent spliceosomal complexes. *J Biol Chem* 2007, 282(18):13617-13626.
87. Wang E, Dimova N, Cambi F: PLP/DM20 ratio is regulated by hnRNPH and F and a novel G-rich enhancer in oligodendrocytes. *Nucleic Acids Res* 2007, 35(12):4164-4178.
88. Zamore PD, Green MR: Identification, purification, and biochemical characterization of U2 small nuclear ribonucleoprotein auxiliary factor. *Proc Natl Acad Sci U S A* 1989, 86(23):9243-9247.
89. Moore MJ: Intron recognition comes of AGe. *Nature structural biology* 2000, 7(1):14-16.
90. Merendino L, Guth S, Bilbao D, Martinez C, Valcarcel J: Inhibition of msl-2 splicing by Sex-lethal reveals interaction between U2AF35 and the 3' splice site AG. *Nature* 1999, 402(6763):838-841.
91. Zorio DA, Blumenthal T: Both subunits of U2AF recognize the 3' splice site in Caenorhabditis elegans. *Nature* 1999, 402(6763):835-838.
92. Wu S, Romfo CM, Nilsen TW, Green MR: Functional recognition of the 3' splice site AG by the splicing factor U2AF35. *Nature* 1999, 402(6763):832-835.
93. Smith CW, Chu TT, Nadal-Ginard B: Scanning and competition between AGs are involved in 3' splice site selection in mammalian introns. *Mol Cell Biol* 1993, 13(8):4939-4952.
94. Kanopka A, Muhlemann O, Akusjarvi G: Inhibition by SR proteins of splicing of a regulated adenovirus pre-mRNA. *Nature* 1996, 381(6582):535-538.

95. Moriyoshi K, Masu M, Ishii T, Shigemoto R, Mizuno N, Nakanishi S: Molecular cloning and characterization of the rat NMDA receptor. *Nature* 1991, 354(6348):31-37.
96. Ladd AN, Charlet N, Cooper TA: The CELF family of RNA binding proteins is implicated in cell-specific and developmentally regulated alternative splicing. *Mol Cell Biol* 2001, 21(4):1285-1296.
97. Faustino NA, Cooper TA: Identification of putative new splicing targets for ETR-3 using sequences identified by systematic evolution of ligands by exponential enrichment. *Mol Cell Biol* 2005, 25(3):879-887.
98. Suzuki H, Takeuchi M, Sugiyama A, Alam AK, Vu LT, Sekiyama Y, Dam HC, Ohki SY, Tsukahara T: Alternative splicing produces structural and functional changes in CUGBP2. *BMC Biochem* 2012, 13:6.
99. Philips AV, Timchenko LT, Cooper TA: Disruption of splicing regulated by a CUG-binding protein in myotonic dystrophy. *Science* 1998, 280(5364):737-741.
100. Timchenko LT, Miller JW, Timchenko NA, DeVore DR, Datar KV, Lin L, Roberts R, Caskey CT, Swanson MS: Identification of a (CUG)n triplet repeat RNA-binding protein and its expression in myotonic dystrophy. *Nucleic Acids Res* 1996, 24(22):4407-4414.
101. Kim-Ha J, Kerr K, Macdonald PM: Translational regulation of oskar mRNA by bruno, an ovarian RNA-binding protein, is essential. *Cell* 1995, 81(3):403-412.
102. Suzuki H, Jin Y, Otani H, Yasuda K, Inoue K: Regulation of alternative splicing of alpha-actinin transcript by Bruno-like proteins. *Genes Cells* 2002, 7(2):133-141.
103. Paillard L, Omilli F, Legagneux V, Bassez T, Maniey D, Osborne HB: EDEN and EDEN-BP, a cis element and an associated factor that mediate sequence-specific mRNA deadenylation in Xenopus embryos. *EMBO J* 1998, 17(1):278-287.
104. Timchenko NA, Welm AL, Lu X, Timchenko LT: CUG repeat binding protein (CUGBP1) interacts with the 5' region of C/EBPbeta mRNA and regulates translation of C/EBPbeta isoforms. *Nucleic Acids Res* 1999, 27(22):4517-4525.
105. Tsuda K, Kuwasako K, Takahashi M, Someya T, Inoue M, Terada T, Kobayashi N, Shirouzu M, Kigawa T, Tanaka A *et al*: Structural basis for the sequence-specific RNA-recognition mechanism of human CUG-BP1 RRM3. *Nucleic Acids Res* 2009, 37(15):5151-5166.
106. Guo M, Jan LY, Jan YN: Control of daughter cell fates during asymmetric division: interaction of Numb and Notch. *Neuron* 1996, 17(1):27-41.
107. Pece S, Serresi M, Santolini E, Capra M, Hulleman E, Galimberti V, Zurrida S, Maisonneuve P, Viale G, Di Fiore PP: Loss of negative regulation by Numb over Notch is relevant to human breast carcinogenesis. *J Cell Biol* 2004, 167(2):215-221.
108. Spana EP, Doe CQ: Numb antagonizes Notch signaling to specify sibling neuron cell fates. *Neuron* 1996, 17(1):21-26.
109. Feng Y, Bankston A: The star family member QKI and cell signaling. *Adv Exp Med Biol* 2010, 693:25-36.
110. Galarneau A, Richard S: Target RNA motif and target mRNAs of the Quaking STAR protein. *Nat Struct Mol Biol* 2005, 12(8):691-698.
111. Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, Rothballer A, Ascano M, Jr., Jungkamp AC, Munschauer M *et al*: Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* 2010, 141(1):129-141.
112. Corioni M, Antih N, Tanackovic G, Zavolan M, Kramer A: Analysis of in situ pre-mRNA targets of human splicing factor SF1 reveals a function in alternative splicing. *Nucleic Acids Res* 2011, 39(5):1868-1879.

113. **Lallena MJ, Chalmers KJ, Llamazares S, Lamond AI, Valcarcel J: Splicing regulation at the second catalytic step by Sex-lethal involves 3' splice site recognition by SPF45. *Cell* 2002, 109(3):285-296.**

114. **Sharma S, Kohlstaedt LA, Damianov A, Rio DC, Black DL: Polypyrimidine tract binding protein controls the transition from exon definition to an intron defined spliceosome. *Nat Struct Mol Biol* 2008, 15(2):183-191.**

115. **Solnick D: Alternative splicing caused by RNA secondary structure. *Cell* 1985, 43(3 Pt 2):667-676.**

116. **Clouet d'Orval B, d'Aubenton Carafa Y, Sirand-Pugnet P, Gallego M, Brody E, Marie J: RNA secondary structure repression of a muscle-specific exon in HeLa cell nuclear extracts. *Science* 1991, 252(5014):1823-1828.**

117. **Chebli K, Gattoni R, Schmitt P, Hildwein G, Stevenin J: The 216-nucleotide intron of the E1A pre-mRNA contains a hairpin structure that permits utilization of unusually distant branch acceptors. *Mol Cell Biol* 1989, 9(11):4852-4861.**

118. **Eperon LP, Graham IR, Griffiths AD, Eperon IC: Effects of RNA secondary structure on alternative splicing of pre-mRNA: is folding limited to a region behind the transcribing RNA polymerase? *Cell* 1988, 54(3):393-401.**

119. **Shepard PJ, Hertel KJ: Conserved RNA secondary structures promote alternative splicing. *RNA* 2008, 14(8):1463-1469.**

120. **Zhang J, Kuo CC, Chen L: GC content around splice sites affects splicing through pre-mRNA secondary structures. *BMC Genomics* 2011, 12:90.**

121. **Meyer M, Plass M, Perez-Valle J, Eyras E, Vilardell J: Deciphering 3'ss selection in the yeast genome reveals an RNA thermosensor that mediates alternative splicing. *Mol Cell* 2011, 43(6):1033-1039.**

122. **Hiller M, Zhang Z, Backofen R, Stamm S: Pre-mRNA secondary structures influence exon recognition. *PLoS Genet* 2007, 3(11):e204.**

123. **Monie TP, Hernandez H, Robinson CV, Simpson P, Matthews S, Curry S: The polypyrimidine tract binding protein is a monomer. *RNA* 2005, 11(12):1803-1808.**

124. **Oberstrass FC, Auweter SD, Erat M, Hargous Y, Henning A, Wenter P, Reymond L, Amir-Ahmady B, Pitsch S, Black DL *et al*: Structure of PTB bound to RNA: specific binding and implications for splicing regulation. *Science* 2005, 309(5743):2054-2057.**

125. **Kim JH, Hahm B, Kim YK, Choi M, Jang SK: Protein-protein interaction among hnRNPs shuttling between nucleus and cytoplasm. *J Mol Biol* 2000, 298(3):395-405.**

126. **Rahman MA, Masuda A, Ohe K, Ito M, Hutchinson DO, Mayeda A, Engel AG, Ohno K: HnRNP L and hnRNP LL antagonistically modulate PTB-mediated splicing suppression of CHRNA1 pre-mRNA. *Sci Rep* 2013, 3:2931.**

127. **Mollet I, Barbosa-Morais NL, Andrade J, Carmo-Fonseca M: Diversity of human U2AF splicing factors. *FEBS J* 2006, 273(21):4807-4816.**

128. **Rooke N, Markovtsov V, Cagavi E, Black DL: Roles for SR proteins and hnRNP A1 in the regulation of c-src exon N1. *Mol Cell Biol* 2003, 23(6):1874-1884.**

129. **Gahura O, Hammann C, Valentova A, Puta F, Folk P: Secondary structure is required for 3' splice site recognition in yeast. *Nucleic Acids Res* 2011, 39(22):9759-9767.**

130. **Deshler JO, Rossi JJ: Unexpected point mutations activate cryptic 3' splice sites by perturbing a natural secondary structure within a yeast intron. *Genes Dev* 1991, 5(7):1252-1263.**

131. **Charpentier B, Rosbash M: Intramolecular structure in yeast introns aids the early steps of in vitro spliceosome assembly. *RNA* 1996, 2(6):509-522.**

132. Luco RF, Pan Q, Tominaga K, Blencowe BJ, Pereira-Smith OM, Misteli T: Regulation of alternative splicing by histone modifications. *Science* 2010, 327(5968):996-1000.
133. Choudhary C, Kumar C, Gnad F, Nielsen ML, Rehman M, Walther TC, Olsen JV, Mann M: Lysine acetylation targets protein complexes and co-regulates major cellular functions. *Science* 2009, 325(5942):834-840.
134. Ip JY, Schmidt D, Pan Q, Ramani AK, Fraser AG, Odom DT, Blencowe BJ: Global impact of RNA polymerase II elongation inhibition on alternative splicing regulation. *Genome Res* 2011, 21(3):390-401.
135. de la Mata M, Alonso CR, Kadener S, Fededa JP, Blaustein M, Pelisch F, Cramer P, Bentley D, Kornblihtt AR: A slow RNA polymerase II affects alternative splicing in vivo. *Mol Cell* 2003, 12(2):525-532.
136. Roberts GC, Gooding C, Mak HY, Proudfoot NJ, Smith CW: Co-transcriptional commitment to alternative splice site selection. *Nucleic Acids Res* 1998, 26(24):5568-5572.
137. Fong N, Kim H, Zhou Y, Ji X, Qiu J, Saldi T, Diener K, Jones K, Fu XD, Bentley DL: Pre-mRNA splicing is facilitated by an optimal RNA polymerase II elongation rate. *Genes Dev* 2014, 28(23):2663-2676.
138. Chen W, Luo L, Zhang L: The organization of nucleosomes around splice sites. *Nucleic Acids Res* 2010, 38(9):2788-2798.
139. Huang H, Yu S, Liu H, Sun X: Nucleosome organization in sequences of alternative events in human genome. *Biosystems* 2012, 109(2):214-219.
140. Reisman DN, Strobeck MW, Betz BL, Sciariotta J, Funkhouser W, Jr., Murchardt C, Yaniv M, Sherman LS, Knudsen ES, Weissman BE: Concomitant down-regulation of BRM and BRG1 in human tumor cell lines: differential effects on RB-mediated growth arrest vs CD44 expression. *Oncogene* 2002, 21(8):1196-1207.
141. Kadam S, Emerson BM: Transcriptional specificity of human SWI/SNF BRG1 and BRM chromatin remodeling complexes. *Mol Cell* 2003, 11(2):377-389.
142. Corey LL, Weirich CS, Benjamin IJ, Kingston RE: Localized recruitment of a chromatin-remodeling activity by an activator in vivo drives transcriptional elongation. *Genes Dev* 2003, 17(11):1392-1401.
143. Batsche E, Yaniv M, Muchardt C: The human SWI/SNF subunit Brm is a regulator of alternative splicing. *Nat Struct Mol Biol* 2006, 13(1):22-29.
144. Matter N, Herrlich P, Konig H: Signal-dependent regulation of splicing via phosphorylation of Sam68. *Nature* 2002, 420(6916):691-695.
145. Webby CJ, Wolf A, Gromak N, Dreger M, Kramer H, Kessler B, Nielsen ML, Schmitz C, Butler DS, Yates JR, 3rd *et al*: Jmjd6 catalyses lysyl-hydroxylation of U2AF65, a protein associated with RNA splicing. *Science* 2009, 325(5936):90-93.
146. Zhang Y, Madl T, Bagdiul I, Kern T, Kang HS, Zou P, Mausbacher N, Sieber SA, Kramer A, Sattler M: Structure, phosphorylation and U2AF65 binding of the N-terminal domain of splicing factor 1 during 3'-splice site recognition. *Nucleic Acids Res* 2013, 41(2):1343-1354.
147. Wang X, Bruderer S, Rafi Z, Xue J, Milburn PJ, Kramer A, Robinson PJ: Phosphorylation of splicing factor SF1 on Ser20 by cGMP-dependent protein kinase regulates spliceosome assembly. *EMBO J* 1999, 18(16):4549-4559.
148. Manceau V, Swenson M, Le Caer JP, Sobel A, Kielkopf CL, Maucuer A: Major phosphorylation of SF1 on adjacent Ser-Pro motifs enhances interaction with U2AF65. *FEBS J* 2006, 273(3):577-587.

149. Rain JC, Rafi Z, Rhani Z, Legrain P, Kramer A: Conservation of functional domains involved in RNA binding and protein-protein interactions in human and Saccharomyces cerevisiae pre-mRNA splicing factor SF1. *RNA* 1998, 4(5):551-565.
150. Bezzi M, Teo SX, Muller J, Mok WC, Sahu SK, Vardy LA, Bonday ZQ, Guccione E: Regulation of constitutive and alternative splicing by PRMT5 reveals a role for Mdm4 pre-mRNA in sensing defects in the spliceosomal machinery. *Genes Dev* 2013, 27(17):1903-1916.
151. Edmond V, Moysan E, Khochbin S, Matthias P, Brambilla C, Brambilla E, Gazzeri S, Eymin B: Acetylation and phosphorylation of SRSF2 control cell fate decision in response to cisplatin. *EMBO J* 2011, 30(3):510-523.
152. Kuhn AN, van Santen MA, Schwienhorst A, Urlaub H, Luhrmann R: Stalling of spliceosome assembly at distinct stages by small-molecule inhibitors of protein acetylation and deacetylation. *RNA* 2009, 15(1):153-175.
153. Seyedarabi A, Sullivan JA, Sasakawa C, Pickersgill RW: A disulfide driven domain swap switches off the activity of Shigella IpaH9.8 E3 ligase. *FEBS Lett* 2010, 584(19):4163-4168.
154. Song EJ, Werner SL, Neubauer J, Stegmeier F, Aspden J, Rio D, Harper JW, Elledge SJ, Kirschner MW, Rape M: The Prp19 complex and the Usp4Sart3 deubiquitinating enzyme control reversible ubiquitination at the spliceosome. *Genes Dev* 2010, 24(13):1434-1447.
155. Kaida D, Motoyoshi H, Tashiro E, Nojima T, Hagiwara M, Ishigami K, Watanabe H, Kitahara T, Yoshida T, Nakajima H *et al*: Spliceostatin A targets SF3b and inhibits both splicing and nuclear retention of pre-mRNA. *Nat Chem Biol* 2007, 3(9):576-583.
156. Habelhah H, Shah K, Huang L, Ostareck-Lederer A, Burlingame AL, Shokat KM, Hentze MW, Ronai Z: ERK phosphorylation drives cytoplasmic accumulation of hnRNP-K and inhibition of mRNA translation. *Nature cell biology* 2001, 3(3):325-330.
157. Razanau A, Xie J: Emerging mechanisms and consequences of calcium regulation of alternative splicing in neurons and endocrine cells. *Cell Mol Life Sci* 2013, 70(23):4527-4536.
158. van der Houven van Oordt W, Diaz-Meco MT, Lozano J, Krainer AR, Moscat J, Caceres JF: The MKK(3/6)-p38-signaling cascade alters the subcellular distribution of hnRNP A1 and modulates alternative splicing regulation. *J Cell Biol* 2000, 149(2):307-316.
159. Allemand E, Guil S, Myers M, Moscat J, Caceres JF, Krainer AR: Regulation of heterogenous nuclear ribonucleoprotein A1 transport by phosphorylation in cells stressed by osmotic shock. *Proc Natl Acad Sci U S A* 2005, 102(10):3605-3610.
160. Xie J, Lee JA, Kress TL, Mowry KL, Black DL: Protein kinase A phosphorylation modulates transport of the polypyrimidine tract-binding protein. *Proc Natl Acad Sci U S A* 2003, 100(15):8776-8781.
161. Ma S, Liu G, Sun Y, Xie J: Relocalization of the polypyrimidine tract-binding protein during PKA-induced neurite growth. *Biochim Biophys Acta* 2007, 1773(6):912-923.
162. Caceres JF, Screaton GR, Krainer AR: A specific subset of SR proteins shuttles continuously between the nucleus and the cytoplasm. *Genes Dev* 1998, 12(1):55-66.
163. Tacke R, Chen Y, Manley JL: Sequence-specific RNA binding by an SR protein requires RS domain phosphorylation: creation of an SRp40-specific splicing enhancer. *Proc Natl Acad Sci U S A* 1997, 94(4):1148-1153.
164. Cao W, Jamison SF, Garcia-Blanco MA: Both phosphorylation and dephosphorylation of ASF/SF2 are required for pre-mRNA splicing in vitro. *RNA* 1997, 3(12):1456-1467.
165. Xiao SH, Manley JL: Phosphorylation-dephosphorylation differentially affects activities of splicing factor ASF/SF2. *EMBO J* 1998, 17(21):6359-6367.
166. Ma CT, Ghosh G, Fu XD, Adams JA: Mechanism of dephosphorylation of the SR protein ASF/SF2 by protein phosphatase 1. *J Mol Biol* 2010, 403(3):386-404.

167. Dagher SF, Fu XD: Evidence for a role of Sky1p-mediated phosphorylation in 3' splice site recognition involving both Prp8 and Prp17/Slu4. *RNA* 2001, 7(9):1284-1297.

168. Boisvert FM, Cote J, Boulanger MC, Richard S: A proteomic analysis of arginine-methylated protein complexes. *Mol Cell Proteomics* 2003, 2(12):1319-1330.

169. Koumbadinga G, Mahmood, N., Lei, L., Kan, YC, Cao, WG, Lobo, VG, Yao, XJ, Zhang, SZ and Xie, J. : Increased stability of heterogeneous ribonucleoproteins by a deacetylase inhibitor. *BBA-Gene Regulatory Mechanisms* 2015, in press.

170. Okuda J, Toyotome T, Kataoka N, Ohno M, Abe H, Shimura Y, Seyedarabi A, Pickersgill R, Sasakawa C: Shigella effector IpaH9.8 binds to a splicing factor U2AF(35) to modulate host immune responses. *Biochem Biophys Res Commun* 2005, 333(2):531-539.

171. Bellare P, Kutach AK, Rines AK, Guthrie C, Sontheimer EJ: Ubiquitin binding by a variant Jab1/MPN domain in the essential pre-mRNA splicing factor Prp8p. *RNA* 2006, 12(2):292-302.

172. Bellare P, Small EC, Huang X, Wohlschlegel JA, Staley JP, Sontheimer EJ: A role for ubiquitin in the spliceosome assembly pathway. *Nat Struct Mol Biol* 2008, 15(5):444-451.

173. Kotake Y, Sagane K, Owa T, Mimori-Kiyosue Y, Shimizu H, Uesugi M, Ishihama Y, Iwata M, Mizui Y: Splicing factor SF3b as a target of the antitumor natural product pladienolide. *Nat Chem Biol* 2007, 3(9):570-575.

174. Nakajima H, Sato B, Fujita T, Takase S, Terano H, Okuhara M: New antitumor substances, FR901463, FR901464 and FR901465. I. Taxonomy, fermentation, isolation, physico-chemical properties and biological activities. *J Antibiot (Tokyo)* 1996, 49(12):1196-1203.

175. Gozani O, Potashkin J, Reed R: A potential role for U2AF-SAP 155 interactions in recruiting U2 snRNP to the branch site. *Mol Cell Biol* 1998, 18(8):4752-4760.

176. Konarska MM, Sharp PA: Interactions between small nuclear ribonucleoprotein particles in formation of spliceosomes. *Cell* 1987, 49(6):763-774.

177. Pikielny CW, Rymond BC, Rosbash M: Electrophoresis of ribonucleoproteins reveals an ordered assembly pathway of yeast splicing complexes. *Nature* 1986, 324(6095):341-345.

178. Rutz B, Seraphin B: Transient interaction of BBP/ScSF1 and Mud2 with the splicing machinery affects the kinetics of spliceosome assembly. *RNA* 1999, 5(6):819-831.

179. Corrionero A, Minana B, Valcarcel J: Reduced fidelity of branch point recognition and alternative splicing induced by the anti-tumor drug spliceostatin A. *Genes Dev* 2011, 25(5):445-459.

180. Roybal GA, Jurica MS: Spliceostatin A inhibits spliceosome assembly subsequent to prespliceosome formation. *Nucleic Acids Res* 2010, 38(19):6664-6672.

181. Alekseyenko AV, Kim N, Lee CJ: Global analysis of exon creation versus loss and the role of alternative splicing in 17 vertebrate genomes. *RNA* 2007, 13(5):661-670.

182. Barbosa-Morais NL, Irimia M, Pan Q, Xiong HY, Gueroussov S, Lee LJ, Slobodeniuc V, Kutter C, Watt S, Colak R *et al*: The evolutionary landscape of alternative splicing in vertebrate species. *Science* 2012, 338(6114):1587-1593.

183. Merkin J, Russell C, Chen P, Burge CB: Evolutionary dynamics of gene and isoform regulation in Mammalian tissues. *Science* 2012, 338(6114):1593-1599.

184. Kondrashov FA, Koonin EV: Origin of alternative splicing by tandem exon duplication. *Hum Mol Genet* 2001, 10(23):2661-2669.

185. Krull M, Brosius J, Schmitz J: Alu-SINE exonization: en route to protein-coding function. *Mol Biol Evol* 2005, 22(8):1702-1711.

186. Sorek R: The birth of new exons: mechanisms and evolutionary consequences. *RNA* 2007, 13(10):1603-1608.

187. Sorek R, Ast G, Graur D: Alu-containing exons are alternatively spliced. *Genome Res* 2002, 12(7):1060-1067.
188. Keren H, Lev-Maor G, Ast G: Alternative splicing and evolution: diversification, exon definition and function. *Nat Rev Genet* 2010, 11(5):345-355.
189. Pan Q, Bakowski MA, Morris Q, Zhang W, Frey BJ, Hughes TR, Blencowe BJ: Alternative splicing of conserved exons is frequently species-specific in human and mouse. *Trends Genet* 2005, 21(2):73-77.
190. Lev-Maor G, Goren A, Sela N, Kim E, Keren H, Doron-Faigenboim A, Leibman-Barak S, Pupko T, Ast G: The "alternative" choice of constitutive exons throughout evolution. *PLoS Genet* 2007, 3(11):e203.
191. Koren E, Lev-Maor G, Ast G: The emergence of alternative 3' and 5' splice site exons from constitutive exons. *PLoS Comput Biol* 2007, 3(5):e95.
192. Yeo G, Hoon S, Venkatesh B, Burge CB: Variation in sequence and organization of splicing regulatory elements in vertebrate genes. *Proc Natl Acad Sci U S A* 2004, 101(44):15700-15705.
193. Xie J: Differential evolution of signal-responsive RNA elements and upstream factors that control alternative splicing. *Cell Mol Life Sci* 2014, 71(22):4347-4360.
194. Liu G, Lei L, Yu J, Kung S, Xie J: Refinement of the spectra of exon usage by combined effects of extracellular stimulus and intracellular factors. *Biochim Biophys Acta* 2014, 1839(7):537-545.
195. Schor IE, Rascovan N, Pelisch F, Allo M, Kornblihtt AR: Neuronal cell depolarization induces intragenic chromatin modifications affecting NCAM alternative splicing. *Proc Natl Acad Sci U S A* 2009, 106(11):4325-4330.
196. Sanchez SE, Petrillo E, Beckwith EJ, Zhang X, Rugnone ML, Hernando CE, Cuevas JC, Godoy Herz MA, Depetris-Chauvin A, Simpson CG *et al*: A methyl transferase links the circadian clock to the regulation of alternative splicing. *Nature* 2010, 468(7320):112-116.
197. Cech TR, Steitz JA: The noncoding RNA revolution-trashing old rules to forge new ones. *Cell* 2014, 157(1):77-94.
198. Kishore S, Stamm S: The snoRNA HBII-52 regulates alternative splicing of the serotonin receptor 2C. *Science* 2006, 311(5758):230-232.
199. Beltran M, Puig I, Pena C, Garcia JM, Alvarez AB, Pena R, Bonilla F, de Herreros AG: A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial-mesenchymal transition. *Genes Dev* 2008, 22(6):756-769.
200. Hastings ML, Milcarek C, Martincic K, Peterson ML, Munroe SH: Expression of the thyroid hormone receptor gene, erbAalpha, in B lymphocytes: alternative mRNA processing is independent of differentiation but correlates with antisense RNA levels. *Nucleic Acids Res* 1997, 25(21):4296-4300.
201. Morrissy AS, Griffith M, Marra MA: Extensive relationship between antisense transcription and alternative splicing in the human genome. *Genome Res* 2011, 21(8):1203-1212.
202. Yan MD, Hong CC, Lai GM, Cheng AL, Lin YW, Chuang SE: Identification and characterization of a novel gene Saf transcribed from the opposite strand of Fas. *Hum Mol Genet* 2005, 14(11):1465-1474.
203. Kralovicova J, Knut M, Cross NC, Vorechovsky I: Identification of U2AF(35)-dependent exons by RNA-Seq reveals a link between 3' splice-site organization and activity of U2AF-related proteins. *Nucleic Acids Res* 2015, 43(7):3747-3763.
204. Shepard J, Reick M, Olson S, Graveley BR: Characterization of U2AF(6), a splicing factor related to U2AF(35). *Mol Cell Biol* 2002, 22(1):221-230.

205. Saltzman AL, Kim YK, Pan Q, Fagnani MM, Maquat LE, Blencowe BJ: Regulation of multiple core spliceosomal proteins by alternative splicing-coupled nonsense-mediated mRNA decay. *Mol Cell Biol* 2008, 28(13):4320-4330.

206. Wollerton MC, Gooding C, Robinson F, Brown EC, Jackson RJ, Smith CW: Differential alternative splicing activity of isoforms of polypyrimidine tract binding protein (PTB). *RNA* 2001, 7(6):819-832.

207. Pink RC, Wicks K, Caley DP, Punch EK, Jacobs L, Carter DR: Pseudogenes: pseudo-functional or key regulators in health and disease? *RNA* 2011, 17(5):792-798.

208. Yeo GW, Van Nostrand EL, Liang TY: Discovery and analysis of evolutionarily conserved intronic splicing regulatory elements. *PLoS Genet* 2007, 3(5):e85.

209. Yoshida K, Sanada M, Shiraishi Y, Nowak D, Nagata Y, Yamamoto R, Sato Y, Sato-Otsubo A, Kon A, Nagasaki M *et al*: Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature* 2011, 478(7367):64-69.

210. Patnaik MM, Lasho TL, Finke CM, Hanson CA, Hodnefield JM, Knudson RA, Ketterling RP, Pardanani A, Tefferi A: Spliceosome mutations involving SRSF2, SF3B1, and U2AF35 in chronic myelomonocytic leukemia: prevalence, clinical correlates, and prognostic relevance. *Am J Hematol* 2013, 88(3):201-206.

211. Cazzola M, Della Porta MG, Malcovati L: The genetic basis of myelodysplasia and its clinical relevance. *Blood* 2013, 122(25):4021-4034.

212. Papaemmanuil E, Gerstung M, Malcovati L, Tauro S, Gundem G, Van Loo P, Yoon CJ, Ellis P, Wedge DC, Pellagatti A *et al*: Clinical and biological implications of driver mutations in myelodysplastic syndromes. *Blood* 2013, 122(22):3616-3627; quiz 3699.

213. Bao X, Zhao S, Liu T, Liu Y, Yang X: Overexpression of PRMT5 promotes tumor cell growth and is associated with poor disease prognosis in epithelial ovarian cancer. *J Histochem Cytochem* 2013, 61(3):206-217.

214. Gkountela S, Li Z, Chin CJ, Lee SA, Clark AT: PRMT5 is required for human embryonic stem cell proliferation but not pluripotency. *Stem Cell Rev* 2014, 10(2):230-239.

215. Jansson M, Durant ST, Cho EC, Sheahan S, Edelmann M, Kessler B, La Thangue NB: Arginine methylation regulates the p53 response. *Nat Cell Biol* 2008, 10(12):1431-1439.

216. Kwak YT, Guo J, Prajapati S, Park KJ, Surabhi RM, Miller B, Gehrig P, Gaynor RB: Methylation of SPT5 regulates its interaction with RNA polymerase II and transcriptional elongation properties. *Mol Cell* 2003, 11(4):1055-1066.

217. Majumder S, Alinari L, Roy S, Miller T, Datta J, Sif S, Baiocchi R, Jacob ST: Methylation of histone H3 and H4 by PRMT5 regulates ribosomal RNA gene transcription. *J Cell Biochem* 2010, 109(3):553-563.

218. Nagamatsu G, Kosaka T, Kawasumi M, Kinoshita T, Takubo K, Akiyama H, Sudo T, Kobayashi T, Oya M, Suda T: A germ cell-specific gene, Prmt5, works in somatic cell reprogramming. *J Biol Chem* 2011, 286(12):10641-10648.

219. Pal S, Vishwanath SN, Erdjument-Bromage H, Tempst P, Sif S: Human SWI/SNF-associated PRMT5 methylates histone H3 arginine 8 and negatively regulates expression of ST7 and NM23 tumor suppressor genes. *Mol Cell Biol* 2004, 24(21):9630-9645.

220. Powers MA, Fay MM, Factor RE, Welm AL, Ullman KS: Protein arginine methyltransferase 5 accelerates tumor growth by arginine methylation of the tumor suppressor programmed cell death 4. *Cancer Res* 2011, 71(16):5579-5587.

221. Scoumanne A, Zhang J, Chen X: PRMT5 is required for cell-cycle progression and p53 tumor suppressor function. *Nucleic Acids Res* 2009, 37(15):4965-4976.

222. **Tee WW, Pardo M, Theunissen TW, Yu L, Choudhary JS, Hajkova P, Surani MA: Prmt5 is essential for early mouse development and acts in the cytoplasm to maintain ES cell pluripotency. *Genes Dev* 2010, 24(24):2772-2777.**

223. **Torres-Padilla ME, Parfitt DE, Kouzarides T, Zernicka-Goetz M: Histone arginine methylation regulates pluripotency in the early mouse embryo. *Nature* 2007, 445(7124):214-218.**

224. **Wei TY, Juan CC, Hisa JY, Su LJ, Lee YC, Chou HY, Chen JM, Wu YC, Chiu SC, Hsu CP *et al*: Protein arginine methyltransferase 5 is a potential oncoprotein that upregulates G1 cyclins/cyclin-dependent kinases and the phosphoinositide 3-kinase/AKT signaling cascade. *Cancer Sci* 2012, 103(9):1640-1650.**

225. **Zhou Z, Sun X, Zou Z, Sun L, Zhang T, Guo S, Wen Y, Liu L, Wang Y, Qin J *et al*: PRMT5 regulates Golgi apparatus structure through methylation of the golgin GM130. *Cell Res* 2010, 20(9):1023-1033.**

226. **Ancelin K, Lange UC, Hajkova P, Schneider R, Bannister AJ, Kouzarides T, Surani MA: Blimp1 associates with Prmt5 and directs histone arginine methylation in mouse germ cells. *Nat Cell Biol* 2006, 8(6):623-630.**

227. **Liu F, Zhao X, Perna F, Wang L, Koppikar P, Abdel-Wahab O, Harr MW, Levine RL, Xu H, Tefferi A *et al*: JAK2V617F-mediated phosphorylation of PRMT5 downregulates its methyltransferase activity and promotes myeloproliferation. *Cancer Cell* 2011, 19(2):283-294.**

228. **Cho EC, Zheng S, Munro S, Liu G, Carr SM, Moehlenbrink J, Lu YC, Stimson L, Khan O, Konietzny R *et al*: Arginine methylation controls growth regulation by E2F-1. *EMBO J* 2012, 31(7):1785-1797.**

229. **Xie J, Black DL: A CaMK IV responsive RNA element mediates depolarization-induced alternative splicing of ion channels. *Nature* 2001, 410(6831):936-939.**

230. **Goulet I, Gauvin G, Boisvenue S, Cote J: Alternative splicing yields protein arginine methyltransferase 1 isoforms with distinct activity, substrate specificity, and subcellular localization. *J Biol Chem* 2007, 282(45):33009-33021.**

231. **Sohail M, Xie J: Evolutionary emergence of a novel splice variant with an opposite effect on the cell cycle. *Mol Cell Biol* 2015, 35(12):2203-2214.**

232. **Black DL: Protein diversity from alternative splicing: a challenge for bioinformatics and post-genome biology. *Cell* 2000, 103(3):367-370.**

233. **Feng D, Xie J: Aberrant splicing in neurological diseases. *Wiley Interdiscip Rev RNA* 2013, 4:631-649.**

234. **Chen CD, Kobayashi R, Helfman DM: Binding of hnRNP H to an exonic splicing silencer is involved in the regulation of alternative splicing of the rat beta-tropomyosin gene. *Genes Dev* 1999, 13(5):593-606.**

235. **Honore B, Rasmussen HH, Vorum H, Dejgaard K, Liu X, Gromov P, Madsen P, Gesser B, Tommerup N, Celis JE: Heterogeneous nuclear ribonucleoproteins H, H', and F are members of a ubiquitously expressed subfamily of related but distinct proteins encoded by genes mapping to different chromosomes. *J Biol Chem* 1995, 270(48):28780-28789.**

236. **Mahe D, Mahl P, Gattoni R, Fischer N, Mattei MG, Stevenin J, Fuchs JP: Cloning of human 2H9 heterogeneous nuclear ribonucleoproteins. Relation with splicing and early heat shock-induced splicing arrest. *J Biol Chem* 1997, 272(3):1827-1836.**

237. **McCullough AJ, Berget SM: G triplets located throughout a class of small vertebrate introns enforce intron borders and regulate splice site selection. *Mol Cell Biol* 1997, 17(8):4562-4571.**

238. **Hai Y, Cao W, Liu G, Hong SP, Elela SA, Klinck R, Chu J, Xie J: A G-tract element in apoptotic agents-induced alternative splicing. *Nucleic Acids Res* 2008, 36(10):3320-3331.**

239. Han K, Yeo G, An P, Burge CB, Grabowski PJ: A combinatorial code for splicing silencing: UAGG and GGGG motifs. *PLoS biology* 2005, 3(5):e158.
240. Sironi M, Menozzi G, Riva L, Cagliani R, Comi GP, Bresolin N, Giorda R, Pozzoli U: Silencer elements as possible inhibitors of pseudoexon splicing. *Nucleic Acids Res* 2004, 32(5):1783-1791.
241. Meister G, Eggert C, Buhler D, Brahms H, Kambach C, Fischer U: Methylation of Sm proteins by a complex containing PRMT5 and the putative U snRNP assembly factor pICln. *Curr Biol* 2001, 11(24):1990-1994.
242. Kung SK: Introduction of shRNAs into primary NK cells with lentivirus. *Methods Mol Biol* 2010, 612:233-247.
243. An DS, Kung SK, Bonifacino A, Wersto RP, Metzger ME, Agricola BA, Mao SH, Chen IS, Donahue RE: Lentivirus vector-mediated hematopoietic stem cell gene transfer of common gamma-chain cytokine receptor in rhesus macaques. *J Virol* 2001, 75(8):3547-3555.
244. Black DL: Activation of c-src neuron-specific splicing by an unusual RNA element in vivo and in vitro. *Cell* 1992, 69(5):795-807.
245. Anders S, Huber W: Differential expression analysis for sequence count data. *Genome Biol* 2010, 11(10):R106.
246. Robinson MD, McCarthy DJ, Smyth GK: edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010, 26(1):139-140.
247. Lee JA, Xing Y, Nguyen D, Xie J, Lee CJ, Black DL: Depolarization and CaM kinase IV modulate NMDA receptor splicing through two essential RNA elements. *PLoS biology* 2007, 5(2):e40.
248. Gu Z, Li Y, Lee P, Liu T, Wan C, Wang Z: Protein arginine methyltransferase 5 functions in opposite ways in the cytoplasm and nucleus of prostate cancer cells. *PLoS One* 2012, 7(8):e44033.
249. Meloche S, Pouyssegur J: The ERK1/2 mitogen-activated protein kinase pathway as a master regulator of the G1- to S-phase transition. *Oncogene* 2007, 26(22):3227-3239.
250. Tong X, Xie D, O'Kelly J, Miller CW, Muller-Tidow C, Koeffler HP: Cyr61, a member of CCN family, is a tumor suppressor in non-small cell lung cancer. *J Biol Chem* 2001, 276(50):47709-47714.
251. Vairapandi M, Balliet AG, Hoffman B, Liebermann DA: GADD45b and GADD45g are cdc2/cyclinB1 kinase inhibitors with a role in S and G2/M cell cycle checkpoints induced by genotoxic stress. *J Cell Physiol* 2002, 192(3):327-338.
252. Noguchi K, Fukazawa H, Murakami Y, Uehara Y: Nek11, a new member of the NIMA family of kinases, involved in DNA replication and genotoxic stress responses. *J Biol Chem* 2002, 277(42):39655-39665.
253. Church DL, Guan KL, Lambie EJ: Three genes of the MAP kinase cascade, mek-2, mpk-1/sur-1 and let-60 ras, are required for meiotic cell cycle progression in Caenorhabditis elegans. *Development* 1995, 121(8):2525-2535.
254. Jin S, Tong T, Fan W, Fan F, Antinore MJ, Zhu X, Mazzacurati L, Li X, Petrik KL, Rajasekaran B *et al*: GADD45-induced cell cycle G2-M arrest associates with altered subcellular distribution of cyclin B1 and is independent of p38 kinase activity. *Oncogene* 2002, 21(57):8696-8704.
255. Kubota S, Hattori T, Shimo T, Nakanishi T, Takigawa M: Novel intracellular effects of human connective tissue growth factor expressed in Cos-7 cells. *FEBS Lett* 2000, 474(1):58-62.
256. Melixetian M, Klein DK, Sorensen CS, Helin K: NEK11 regulates CDC25A degradation and the IR-induced G2/M checkpoint. *Nat Cell Biol* 2009, 11(10):1247-1253.
257. Yang L, Besschetnova TY, Brooks CR, Shah JV, Bonventre JV: Epithelial cell cycle arrest in G2/M mediates kidney fibrosis after injury. *Nat Med* 2010, 16(5):535-543, 531p following 143.

258. Singh AK, Swarnalatha M, Kumar V: c-ETS1 facilitates G1/S-phase transition by up-regulating cyclin E and CDK2 genes and cooperates with hepatitis B virus X protein for their deregulation. *J Biol Chem* 2011, 286(25):21961-21970.

259. Lacroix M, El Messaoudi S, Rodier G, Le Cam A, Sardet C, Fabbrizio E: The histone-binding protein COPR5 is required for nuclear functions of the protein arginine methyltransferase PRMT5. *EMBO Rep* 2008, 9(5):452-458.

260. Graveley BR: Alternative splicing: increasing diversity in the proteomic world. *Trends Genet* 2001, 17(2):100-107.

261. Yeo G, Holste D, Kreiman G, Burge CB: Variation in alternative splicing across human tissues. *Genome Biol* 2004, 5(10):R74.

262. Grabowski PJ: Genetic evidence for a Nova regulator of alternative splicing in the brain. *Neuron* 2000, 25(2):254-256.

263. Cabrera-Poch N, Pepperkok R, Shima DT: Inheritance of the mammalian Golgi apparatus during the cell cycle. *Biochim Biophys Acta* 1998, 1404(1-2):139-151.

264. Stanley H, Botas J, Malhotra V: The mechanism of Golgi segregation during mitosis is cell type-specific. *Proc Natl Acad Sci U S A* 1997, 94(26):14467-14470.

265. Corda D, Barretta ML, Cervigni RI, Colanzi A: Golgi complex fragmentation in G2/M transition: An organelle-based cell-cycle checkpoint. *IUBMB Life* 2012, 64(8):661-670.

266. Sohail M, Cao W, Mahmood N, Myschyshyn M, Hong SP, Xie J: Evolutionarily emerged G tracts between the polypyrimidine tract and 3' AG are splicing silencers enriched in genes involved in cancer. *BMC Genomics* 2014, 15(1):1143.

267. Pontinen T, Melin A, Varadi G, Khanmoradi K, Chewaproug D, Kung SC, Zaki R, Ortiz J: Cutaneous metastasis of pancreatic adenocarcinoma after kidney transplant: a case report and review of the literature. *Exp Clin Transplant* 2010, 8(4):273-276.

268. Li L, Wang S, Jezierski A, Moalim-Nour L, Mohib K, Parks RJ, Retta SF, Wang L: A unique interplay between Rap1 and E-cadherin in the endocytic pathway regulates self-renewal of human embryonic stem cells. *Stem Cells* 2010, 28(2):247-257.

269. Mitui M, Nahas SA, Du LT, Yang Z, Lai CH, Nakamura K, Arroyo S, Scott S, Purayidom A, Concannon P *et al*: Functional and computational assessment of missense variants in the ataxia-telangiectasia mutated (ATM) gene: mutations with increased cancer risk. *Human mutation* 2009, 30(1):12-21.

270. Hill DP, Robertson KA: Differentiation of LA-N-5 neuroblastoma cells into cholinergic neurons: methods for differentiation, immunohistochemistry and reporter gene introduction. *Brain Res Brain Res Protoc* 1998, 2(3):183-190.

271. Pence JC, Shorter NA: In vitro differentiation of human neuroblastoma cells caused by vasoactive intestinal peptide. *Cancer Res* 1990, 50(16):5177-5183.

272. Thiele CJ, Cohen PS, Israel MA: Regulation of c-myb expression in human neuroblastoma cells during retinoic acid-induced differentiation. *Mol Cell Biol* 1988, 8(4):1677-1683.

273. Zhang L, Procuik M, Fang T, Kung SK: Functional analysis of the quantitative expression of a costimulatory molecule on dendritic cells using lentiviral vector-mediated RNA interference. *J Immunol Methods* 2009, 344(2):87-97.

274. Huang da W, Sherman BT, Lempicki RA: Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2009, 4(1):44-57.

275. Huang da W, Sherman BT, Lempicki RA: Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* 2009, 37(1):1-13.

276. Colanzi A, Hidalgo Carcedo C, Persico A, Cericola C, Turacchio G, Bonazzi M, Luini A, Corda D: The Golgi mitotic checkpoint is controlled by BARS-dependent fission of the Golgi ribbon into separate stacks in G2. *EMBO J* 2007, 26(10):2465-2476.

277. Tang D, Wang Y: Cell cycle regulation of Golgi membrane dynamics. *Trends Cell Biol* 2013, 23(6):296-304.

278. Su X, Qian C, Zhang Q, Hou J, Gu Y, Han Y, Chen Y, Jiang M, Cao X: miRNomes of haematopoietic stem cells and dendritic cells identify miR-30b as a regulator of Notch1. *Nat Commun* 2013, 4:2903.

279. King KY, Goodell MA: Inflammatory modulation of HSCs: viewing the HSC as a foundation for the immune response. *Nat Rev Immunol* 2011, 11(10):685-692.

280. Auffray C, Sieweke MH, Geissmann F: Blood monocytes: development, heterogeneity, and relationship with dendritic cells. *Annu Rev Immunol* 2009, 27:669-692.

281. Kumanogoh A, Suzuki K, Ch'ng E, Watanabe C, Marukawa S, Takegahara N, Ishida I, Sato T, Habu S, Yoshida K *et al*: Requirement for the lymphocyte semaphorin, CD100, in the induction of antigen-specific T cells and the maturation of dendritic cells. *J Immunol* 2002, 169(3):1175-1181.

282. Takamatsu H, Takegahara N, Nakagawa Y, Tomura M, Taniguchi M, Friedel RH, Rayburn H, Tessier-Lavigne M, Yoshida Y, Okuno T *et al*: Semaphorins guide the entry of dendritic cells into the lymphatics by activating myosin II. *Nat Immunol* 2010, 11(7):594-600.

283. Zhao Q, Rank G, Tan YT, Li H, Moritz RL, Simpson RJ, Cerruti L, Curtis DJ, Patel DJ, Allis CD *et al*: PRMT5-mediated methylation of histone H4R3 recruits DNMT3A, coupling histone and DNA methylation in gene silencing. *Nat Struct Mol Biol* 2009, 16(3):304-311.

284. Stopa N, Krebs JE, Shechter D: The PRMT5 arginine methyltransferase: many roles in development, cancer and beyond. *Cell Mol Life Sci* 2015, 72(11):2041-2059.

285. Ho MC, Wilczek C, Bonanno JB, Xing L, Seznec J, Matsui T, Carter LG, Onikubo T, Kumar PR, Chan MK *et al*: Structure of the arginine methyltransferase PRMT5-MEP50 reveals a mechanism for substrate specificity. *PLoS One* 2013, 8(2):e57008.

286. Antonysamy S, Bonday Z, Campbell RM, Doyle B, Druzina Z, Gheyi T, Han B, Jungheim LN, Qian Y, Rauch C *et al*: Crystal structure of the human PRMT5:MEP50 complex. *Proc Natl Acad Sci U S A* 2012, 109(44):17960-17965.

287. Furuno K, Masatsugu T, Sonoda M, Sasazuki T, Yamamoto K: Association of Polycomb group SUZ12 with WD-repeat protein MEP50 that binds to histone H2A selectively in vitro. *Biochem Biophys Res Commun* 2006, 345(3):1051-1058.

288. Sun L, Wang M, Lv Z, Yang N, Liu Y, Bao S, Gong W, Xu RM: Structural insights into protein arginine symmetric dimethylation by PRMT5. *Proc Natl Acad Sci U S A* 2011, 108(51):20538-20543.

289. Guderian G, Peter C, Wiesner J, Sickmann A, Schulze-Osthoff K, Fischer U, Grimmler M: RioK1, a new interactor of protein arginine methyltransferase 5 (PRMT5), competes with pICln for binding and modulates PRMT5 complex composition and substrate specificity. *J Biol Chem* 2011, 286(3):1976-1986.

290. Gurung B, Feng Z, Iwamoto DV, Thiel A, Jin G, Fan CM, Ng JM, Curran T, Hua X: Menin epigenetically represses Hedgehog signaling in MEN1 tumor syndrome. *Cancer Res* 2013, 73(8):2650-2658.

291. Eckert D, Biermann K, Nettersheim D, Gillis AJ, Steger K, Jack HM, Muller AM, Looijenga LH, Schorle H: Expression of BLIMP1/PRMT5 and concurrent histone H2A/H4 arginine 3 dimethylation in fetal germ cells, CIS/IGCNU and germ cell tumors. *BMC Dev Biol* 2008, 8:106.

292. **Nicholas C, Yang J, Peters SB, Bill MA, Baiocchi RA, Yan F, Sif S, Tae S, Gaudio E, Wu X *et al*: PRMT5 is upregulated in malignant and metastatic melanoma and regulates expression of MITF and p27(Kip1.). *PLoS One* 2013, 8(9):e74710.**

293. **Shilo K, Wu X, Sharma S, Welliver M, Duan W, Villalona-Calero M, Fukuoka J, Sif S, Baiocchi R, Hitchcock CL *et al*: Cellular localization of protein arginine methyltransferase-5 correlates with grade of lung tumors. *Diagn Pathol* 2013, 8:201.**

294. **Koh CM, Bezzi M, Guccione E: The Where and the How of PRMT5. *Current Molecular Biology Reports* 2015, 1(1):19-28.**

295. **Bruns AF, Grothe C, Claus P: Fibroblast growth factor 2 (FGF-2) is a novel substrate for arginine methylation by PRMT5. *Biol Chem* 2009, 390(1):59-65.**

296. **Massenet S, Pellizzoni L, Paushkin S, Mattaj IW, Dreyfuss G: The SMN complex is associated with snRNPs throughout their cytoplasmic assembly pathway. *Mol Cell Biol* 2002, 22(18):6533-6541.**

297. **Kamma H, Portman DS, Dreyfuss G: Cell type-specific expression of hnRNP proteins. *Exp Cell Res* 1995, 221(1):187-196.**

298. **Honore B, Baandrup U, Vorum H: Heterogeneous nuclear ribonucleoproteins F and H/H' show differential expression in normal and selected cancer tissues. *Exp Cell Res* 2004, 294(1):199-209.**

299. **Wang L, Pal S, Sif S: Protein arginine methyltransferase 5 suppresses the transcription of the RB family of tumor suppressors in leukemia and lymphoma cells. *Mol Cell Biol* 2008, 28(20):6262-6277.**

300. **Chittka A, Nitarska J, Grazini U, Richardson WD: Transcription factor positive regulatory domain 4 (PRDM4) recruits protein arginine methyltransferase 5 (PRMT5) to mediate histone arginine methylation and control neural stem cell proliferation and differentiation. *J Biol Chem* 2012, 287(51):42995-43006.**

301. **Dacwag CS, Bedford MT, Sif S, Imbalzano AN: Distinct protein arginine methyltransferases promote ATP-dependent chromatin remodeling function at different stages of skeletal muscle differentiation. *Mol Cell Biol* 2009, 29(7):1909-1921.**

302. **Dacwag CS, Ohkawa Y, Pal S, Sif S, Imbalzano AN: The protein arginine methyltransferase Prmt5 is required for myogenesis because it facilitates ATP-dependent chromatin remodeling. *Mol Cell Biol* 2007, 27(1):384-394.**

303. **Kanade SR, Eckert RL: Protein arginine methyltransferase 5 (PRMT5) signaling suppresses protein kinase Cdelta- and p38delta-dependent signaling and keratinocyte differentiation. *J Biol Chem* 2012, 287(10):7313-7323.**

304. **Paul C, Sardet C, Fabbrizio E: The histone- and PRMT5-associated protein COPR5 is required for myogenic differentiation. *Cell Death Differ* 2012, 19(5):900-908.**

305. **Yang Y, Bedford MT: Protein arginine methyltransferases and cancer. *Nat Rev Cancer* 2013, 13(1):37-50.**

306. **Liu CD, Cheng CP, Fang JS, Chen LC, Zhao B, Kieff E, Peng CW: Modulation of Epstein-Barr virus nuclear antigen 2-dependent transcription by protein arginine methyltransferase 5. *Biochem Biophys Res Commun* 2013, 430(3):1097-1102.**

307. **Shire K, Kapoor P, Jiang K, Hing MN, Sivachandran N, Nguyen T, Frappier L: Regulation of the EBNA1 Epstein-Barr virus protein by serine phosphorylation and arginine methylation. *J Virol* 2006, 80(11):5261-5272.**

308. **Doueiri R, Anupam R, Kvaratskhelia M, Green KB, Lairmore MD, Green PL: Comparative host protein interactions with HTLV-1 p30 and HTLV-2 p28: insights into difference in pathobiology of human retroviruses. *Retrovirology* 2012, 9:64.**

309. Chen M, Yi B, Sun J: Inhibition of cardiomyocyte hypertrophy by protein arginine methyltransferase 5. *J Biol Chem* 2014, 289(35):24325-24335.

310. Daigle SR, Olhava EJ, Therkelsen CA, Majer CR, Sneeringer CJ, Song J, Johnston LD, Scott MP, Smith JJ, Xiao Y *et al*: Selective killing of mixed lineage leukemia cells by a potent small-molecule DOT1L inhibitor. *Cancer Cell* 2011, 20(1):53-65.

311. Yu W, Chory EJ, Wernimont AK, Tempel W, Scopton A, Federation A, Marineau JJ, Qi J, Barsyte-Lovejoy D, Yi J *et al*: Catalytic site remodelling of the DOT1L methyltransferase by selective inhibitors. *Nat Commun* 2012, 3:1288.

312. Xie J: Control of alternative pre-mRNA splicing by Ca(++) signals. *Biochim Biophys Acta* 2008, 1779(8):438-452.

List of Publications

- **Sohail M.**, Zhang M., Litchfield D., Wang L., Kung S., Xie J. Differential Expression, Distinct Localization and Opposite Effect on Cell Differentiation of a Novel Splice Variant of Human PRMT5. BBA-Molecular Cell Research **(Submitted after revision)**.
- **Sohail M**, Xie J. Evolutionary emergence of a novel splice variant with an opposite effect on the cell cycle. *Mol Cell Biol* 2015, 35:2203-2214.
- **Sohail M**, Cao W, Mahmood N, Myschyshyn M, Hong SP, Xie J. Evolutionarily emerged G tracts between the polypyrimidine tract and 3′ AG are splicing silencers enriched in genes involved in cancer. *BMC Genomics* 2014, 15:1143.
- Liu G, Razanau A, Hai Y, Yu J, **Sohail M**, Lobo VG, Chu J, Kung SK, Xie J. A conserved serine of heterogeneous nuclear ribonucleoprotein L (hnRNP L) mediates depolarization-regulated alternative splicing of potassium channels. *J Biol Chem* 2012, 287:22709-22716.
- Cao W, **Sohail M**, Liu G, Koumbadinga GA, Lobo VG, Xie J. Differential effects of PKA-controlled CaMKK2 variants on neuronal differentiation. *RNA Biol* 2011, 8:1061-1072.