Studies on the regulation of RNA binding protein with serine-rich domain 1 and its implication in cervical cancer cells

A thesis submitted for the partial fulfillment of the requirement for the degree of

Doctor of Philosophy

by

BHAGYASHREE DEKA

Under the supervision of

Dr. KUSUM K. SINGH

Department of Biosciences and Bioengineering

Indian Institute of Technology Guwahati

Assam, 781039

July 2022

Studies on the regulation of RNA binding protein with serine-rich domain 1 and its implication in cervical cancer cells

A thesis submitted for the partial fulfillment of the requirement for the degree of

Doctor of Philosophy

by

BHAGYASHREE DEKA

Under the supervision of

Dr. KUSUM K. SINGH

Department of Biosciences and Bioengineering

Indian Institute of Technology Guwahati

Assam, 781039

July 2022

Dedicated to My Parents

Indian Institute of Technology Guwahati

Department of Biosciences and Bioengineering

Statement

I do hereby declare that the research findings in this thesis entitled "Studies on the regulation of RNA binding protein with serine-rich domain 1 and its implication in cervical cancer cells" is the result of work carried out in the Department of Biosciences and Bioengineering, Indian Institute of Technology Guwahati, under the supervision of Dr. Kusum K. Singh.

In keeping with the general practice of reporting scientific observations, due acknowledgement has been made wherever the work described is based on the findings of other investigators.

Bhagyashree Deka

Date: 28th July 2022

Bhagyashree Deka

Indian Institute of Technology Guwahati

Department of Biosciences and Bioengineering

Certificate

It is certified that the work described in this thesis entitled "Studies on the regulation of RNA binding protein with serine-rich domain 1 and its implication in cervical cancer cells" by Bhagyashree Deka for the award of the degree of Doctor of Philosophy is an authentic record of the results obtained from the research work carried out under my supervision in the Department of Biosciences and Bioengineering, IIT Guwahati. The work embodied in this thesis has not been submitted elsewhere for a degree.

Dr. Kusum K. Singh
Thesis Supervisor
Department of Biosciences and Bioengineering
Indian Institute of Technology Guwahati
Assam 781039 India

ACKNOWLEDGEMENTS

Firstly, I would like to acknowledge my thesis supervisor, Dr. Kusum K. Singh, for giving me the chance to work in her laboratory and pursue my dream. I am fortunate to receive her continued support, guidance and scientific insights during every step of my Ph.D. journey. I am grateful to her for giving me the freedom to work independently and helping me grow as a scientist and teaching me the skills to be an independent researcher.

I would also extend my gratitude to the members of my doctoral committee, Dr. Anil M. Limaye (Chairperson), Prof. Sachin Kumar and Dr. Akshai Kumar A.S., for their unbiased opinions and insightful recommendations that have been the driving force required for the completion of my thesis.

I would like to thank the Heads of the Department, Prof. Kannan Pakshirajan, Prof. Latha Rangan and Prof. Rakhi Chaturvedi as well as other faculty members for their support and for lending the facilities of their laboratories.

I would like to express my sincere thanks and gratitude to Dr. Sachin Singh Gautam for providing a robust computational facility to carry out RNA Sequencing analysis.

I would also like to extend my acknowledgement towards lab members of Molecular endocrinology lab, Viral immunology lab, Cancer lab and MAB lab for helping me establish basic protocols in the new laboratory during the initial years of my Ph.D. as a first student.

I want to thank all members of the RBP laboratory who have always given me a friendly, supportive and edifying environment in the lab. I would like to acknowledge the help I got from Pratap, Sweta, Ayushi and Priyanka for the work performed in chapter III and IV. I would like to thank Pratap, Sweta, Ayushi, Priyanka and Sayan for their help in preparing the figures. I also appreciate the support from Ajay in gel and buffer

preparations. I would also like to thank Pratap for his help in formatting the thesis. I would always cherish the memories we made together as RBP family.

I extend my thanks to the Department of Biosciences and Bioengineering for providing the research facilities to accomplish my Ph.D. thesis. I want to thanks all the staff members of the Department of Biosciences and Bioengineering for providing me with the logistical support essential to perform my research. Further, I would like to acknowledge IIT Guwahati, Ministry of Human Resources Development, India and the Department of Biotechnology, Government of India, for financial assistance and funding for my Ph.D. project.

I feel fortunate to have Renu, Rachayeeta, Christy, Nayan and Manash as friends who were constantly with me during the ups and downs of my Ph.D. journey. I will never forget the happy time and friendship we shared over the years. I also consider myself lucky to receive much care, love, support and knowledge from Renu.

Finally, I would like to thank my late Father, Mother, Dipika, Gaurav and Bikram for being the pillars of strength during my Ph.D. It is their innumerable sacrifices that have enabled me to reach this stage in life. This thesis is dedicated to the loving memory of my Father for encouraging me for graduate school, for boosting me up during the failures of my Ph.D. journey and for believing in me.

I also express my gratitude to all others whom I may have missed.

Thank you,

Bhagyashree Deka

Table of contents

List of figures

List of table

Abbreviations

5-FU	Fluorouracil
aa	Aminoacid
ActD	Actinomycin D
AML	Acute myeloid leukemia
ANAPC5	Anaphase-promoting complex subunit 5
ANAPC7	Anaphase-promoting complex subunit 7
APC/C	Anaphase-promoting complex/cyclosome
APP	Amyloid precursor protein
ARE	Adenylate uridylate (AU-rich) elements
ARE-BP	ARE-binding proteins
AS	Alternative splicing
ASAP	Apoptosis and splicing associated protein
ATP	Adenosine triphosphate
AURKB	Aurora B kinase
Bcl-2	B cell leukemia-2
bp	Base pair
BP	Branch point
CaCl2	Calcium chloride
CARE	CA-rich elements
CASP1	Caspase 1
CASP4	Caspase 4
CBC	Cap-binding complex
CESC	Cervical Squamous Cell Carcinoma and Endocervical Adenocarcinoma
CLL	Chronic lymphocytic leukemia
C-terminal	Carboxy-terminal
CTSV	Cathepsin V

CURE	CU-rich element
DCP2	Decapping protein 2
dH_2O	Deionized water
DICE	Differentiation control element
DMEM	Dulbecco's Modified Eagle Medium
DMPK	Dystrophia myotonica protein kinase
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
Dox	Doxorubicin
DTT	Dithiothreitol
E. Coli	Escherichia coli
ECM	Extracellular matrix
EDTA	Ethylenediaminetetraacetic acid
EIF4A3	Eukaryotic translation initiation factor 4A3
EJC	Exon junction complex
ELISA	Enzyme-linked immunosorbent assay
ESE	Exonic splicing enhancers
ESS	Exonic splicing silencers
EST	Expressed sequence tags
EtBr	Ethidium bromide
EtOH	Ethanol
FBS	Fetal bovine serum
FRT	Flp Recombination Target
FXR1	Fragile-X mental retardation protein 1
GO	Gene Ontology
GRE	GU-rich elements
HCL	Hydrochloric acid
HDF	Human dermal fibroblast

HIV	Human immunodeficiency virus
hnRNP L	Heterogeneous nuclear ribonucleoprotein L
HNRNPC	Heterogeneous nuclear ribonucleoprotein C
hnRNPs	Heterogeneous nuclear ribonucleoproteins
HPV	Human papillomavirus
HuR	Hu proteins
IBP160	Intron-binding protein of 160 kDa
IF	Immunofluorescence
IgG	Immunoglobulin G
IgM	immunoglobulin μ-chain
ILF2	Interleukin enhancer-binding factor 2
ILF3	Interleukin enhancer-binding factor 3
IMP13	Karyopherin receptor importin 13
IP	Immunoprecipitation
IPTG	Isopropylthio-β-galactoside
ISE	Intronic splicing enhancers
ISS	Intronic splicing silencers
ITR	Inverted terminal repeat
JNK	c-Jun NH(2)-terminal kinases
kDa	Kilo Dalton
KH	K-homology domain
LB	Luria Bertani
Luc	Luciferase
M	Molar
MDM2	Murine double minute 2
MDS	Myelodysplastic syndromes
miRNA	MicroRNA
MMP	Matrix Metalloproteinase
MRE	miRNA response elements

mRNA	Messenger RNA
MSH2	MutS homolog 2
MSH6	MutS homolog 6
NaCl	Sodium chloride
NMD	Nonsense-mediated mRNA decay
NP-40	Nonidet P-40
N-terminal	Amino terminal
OD	Optical density
OXA	Oxandrolone
PAGE	Polyacrylamide gel electrophoresis
PAR-CLIP	Photoactivatable ribonucleoside cross-linking and immunoprecipitation
PAZ	Piwi/Argonaute/Zwille
PBS	Phosphate buffered saline
PCR	Polymerase Chain Reaction
PMSF	Phenylmethylsulfonylfluoride
PPT	Polypyrimidine-tract
Pre-miRNA	Precursor miRNA
Pri-miRNA	Primary miRNA
PSAP	PININ, SAP18 and RNPS1
PTBP	Polypyrimidine-tract binding protein
PTBP1	Polypyrimidine tract-binding protein 1
PTC	Premature termination codons
qRT-PCR	Quantitative Reverse Transcription PCR
RBD	RNA-binding domains
RBP	RNA-binding protein
RBPs	RNA-binding proteins
RGG	Arg-Gly-Gly
Rho-GDI	Rho GDP-dissociation inhibitor

RIP	RNA Immunoprecipitation
RISC	RNA-induced silencing complexes
RNA	Ribonucleic acid
RNPs	Ribonucleoprotein complexes
RNPS1	RNA-binding protein with serine-rich domain 1
RRM	RNA recognition motif
RS Domain	Arg/Ser-rich domain
RSB	RNPS1-SAP18 binding
RT	Room temperature
RT-PCR	Reverse Transcription PCR
S100A1	S100 Calcium Binding Protein A1
SAP	SAF-A/B, Acinus and PIAS
SD	Standard Deviation
SDS	Sodium Dodecyl Sulfate
siRNA	Small interfering RNA
snRNP	small nuclear ribonucleoprotein particles
SR	Serine/arginine-rich protein
SRm160	Serine/Arginine Repetitive Matrix 1
SS	Splice site
STAT3	Signal transducer and activator of transcription 3
TBS	Tris-buffered saline
TCGA	The Cancer Genome Atlas
TGFβ-RII	Transforming growth factor-β receptor II
TIA-1	T-cell-restricted intracellular antigen 1
TIMP	Tissue inhibitors of matrix metalloproteinases
UBL	Ubiquitin-like domain
UCEC	Uterine corpus endometrial carcinoma
UTR	Untranslated region
WDR1	WD-repeat domain 1

Wnt	Wingless
WT	Wild-type
XRN1	Exoribonuclease 1
µg	Microgram
µL	Microlitre
µm	Micrometer
µM	Micromolar

Chapter 1: Introduction and Literature Review

The review of literature embodied in this chapter is partially published.

Deka B, Singh KK. Multifaceted Regulation of Gene Expression by the Apoptosis- and Splicing-Associated Protein Complex and Its Components. International Journal of Biological Sciences. 2017; 13(5):545-560.

1.1 Pre-mRNA splicing

Most of the protein-coding genes in eukaryotes are composed of exon and intron sequences. During the process of splicing, introns are spliced out from the pre-mRNA, and mature functional mRNA is produced from the ligation of exons. Splicing is mediated through cis-elements (splice-sites) localized within the transcript and trans-acting elements, mainly the spliceosomal machinery [1]. Introns are composed of three conserved cis-elements, the 5' splice site (5'SS), branch point (BP) adenosine, and 3' splice site (3'SS). The spliceosome is composed of small nuclear ribonucleoprotein complexes termed U1, U2, U4, U5, and U6 snRNPs and more than 100 proteins [2].

The spliceosome assembly begins with the interaction of U1 snRNP to the 5'SS by base-pairing the 5'end of U1 snRNA with the 5'SS, yielding the early spliceosome complex (E-complex) (Figure 1.1) [3]. Besides U1 snRNA, U1 snRNP consists of three U1-specific proteins (U1-70k, U1A, and U1C) and the Sm proteins [4]. The DEAD-box helicases Prp5 and Sub2 then recruit the U2 snRNP to the BP site, eventually generating the prespliceosome A complex [5, 6]. The marking of an intron by the U1 and U2 snRNPs is believed to be the critical step for splice-site selection and commits an intron to be spliced out.

Subsequently, the U4/U6.U5 tri-snRNP associates with the prespliceosome to form the pre-B complex (pre-catalytic spliceosome) (Figure 1.1). The 5'end of U2 snRNA pairs with the 3'end of U6 snRNA, forming U6/U2 snRNA pairing within the pre-B complex [7]. During the transition of the pre-B complex to the B complex, Prp28, a component of tri-snRNP, unpairs the U1/5' SS duplex and facilitates the transfer of the 5' SS to the U5 and U6 snRNAs [8]. The transfer of the 5' SS is followed by the release of Prp28 and U1 snRNP [9, 10]. The RNA helicase Brr2 then releases U4 snRNA from U6 snRNA to form the active site of the spliceosome in the catalytically active B^{act} complex [11, 12]. The BP adenosine is then docked into the active site; the branching reaction generates the cleaved 5' exon and the lariat-intron intermediate, yielding the C complex [13]. The same active site is used for the second catalytic step of splicing. The DEAH-box ATPase

Prp16 mediates remodeling of the spliceosome for the second step of splicing, exon ligation [14]. The 5' and 3' exons are ligated by the nucleophilic attack of the 5' exon at the 3' SS and the ligated exons are then released from the active site, hence generating the mature mRNA (Figure 1.1) [15].

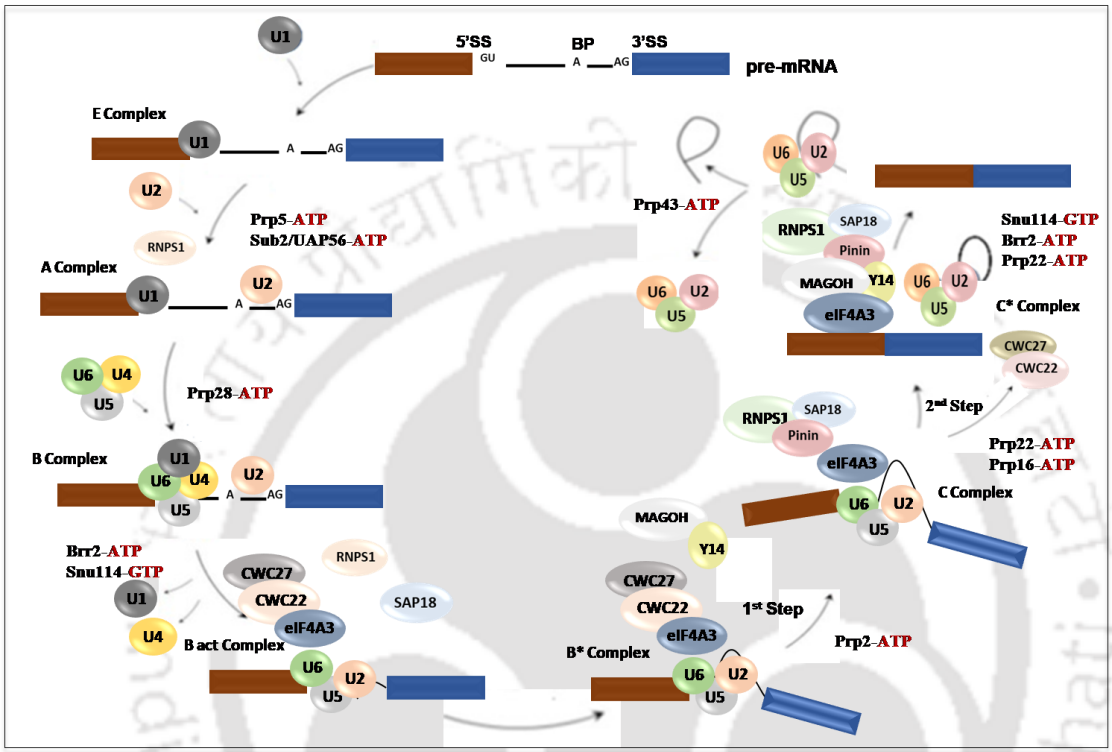


Figure 1.1: Schematic representation of the assembly of spliceosomal machinery onto pre-mRNA. Exons are indicated by brown and blue boxes, while thin black lines show intron and intron lariat. U1 snRNP binds to the 5' splice site of the pre-mRNA and U2 snRNP binds at the branch point (BP) sequence to form the A complex. Next, the U4/U6.U5 tri-snRNP associates with the A complex to form the pre-B complex. Removal of U1 and U4 snRNAs lead to the formation of catalytically active B^{act} complex. After completion of the first splicing step, the C complex is formed. Finally, the C* complex catalyses the release of the intron lariat and facilitate ligation of the exons. Proteomic and structural studies of spliceosomal complexes have demonstrated presence of EJC core and RNPS1 on the spliceosome.

1.1.1 Alternative splicing

Splicing is a tightly controlled process. It is highly coordinated in a tissue-specific and temporal manner to shape the proteome composition of a cell. It is believed that alternative splicing majorly determines the complexity of the proteome. In 1978, Walter Gilbert first proposed the concept of alternative splicing that can generate many functional proteins from a single pre-mRNA by the differential usage of splice-sites [16].

For instance, the *Drosophila melanogaster* gene *Down syndrome cell adhesion molecule* (*Dscam*) can produce 38,016 distinct mRNA isoforms [17]. Further, transcriptome analysis suggests that more than 90% of human genes undergo alternative splicing [18].

Analyses of microarray data and Expressed sequence tags (ESTs) have revealed seven types of alternative splicing (AS) (Figure 1.2). Cassette-type alternative exon is the most widespread splicing event (~30%) in mammals. Alternative selections of 5' or 3' splice sites together account for one-quarter of the AS events (~25%) [19]. Another class of AS events, such as intron retention, is associated with weaker splice sites and short intron length [20].

Alternative splicing is regulated by the binding of regulatory trans-acting proteins to the cis-acting sequences within the pre-mRNA. These cis-acting sequences are grouped as an enhancer (ESEs/ISEs) or suppressor (ESSs/ISSs) elements based on their function [1]. Exonic splicing enhancers (ESEs) and intronic splicing enhancers (ISE) are bound by positive regulators, such as SR proteins. In contrast, exonic splicing silencers (ESSs) and intronic splicing silencers are bound by negative regulators, such as heterogeneous nuclear ribonucleoproteins (hnRNPs).

SR proteins are members of a large family of RNA binding proteins that are conserved across vertebrates and invertebrates. SR proteins promote exon inclusion by recruiting the U1 snRNP or the U2AF auxiliary splicing factor to the 5' and 3' splice sites, respectively [21, 22]. Hence, SR proteins are classically considered to be splicing enhancers. SR proteins contain the signature Arg/Ser-rich domain (called the RS domain) and the RNA recognition motif (RRM) domain [23]. The RS domain at the C-terminus facilitates interactions with other proteins, particularly components of spliceosomal machinery [24]. Studies suggest that the RS domain can interact with the pre-mRNA via the BP and the 5' SS [25, 26]. SRSF1 (SF2/ASF) was the first SR protein discovered as a factor required to reconstitute splicing in splicing-deficient HeLa cytoplasmic extract (S100). It was also identified as a factor in HEK 293 cells which could alter the selection

of 5' SS of an SV40 (simian virus 40) early pre-mRNA [27, 28]. Most SR protein family members function in both constitutive and alternative splicing.

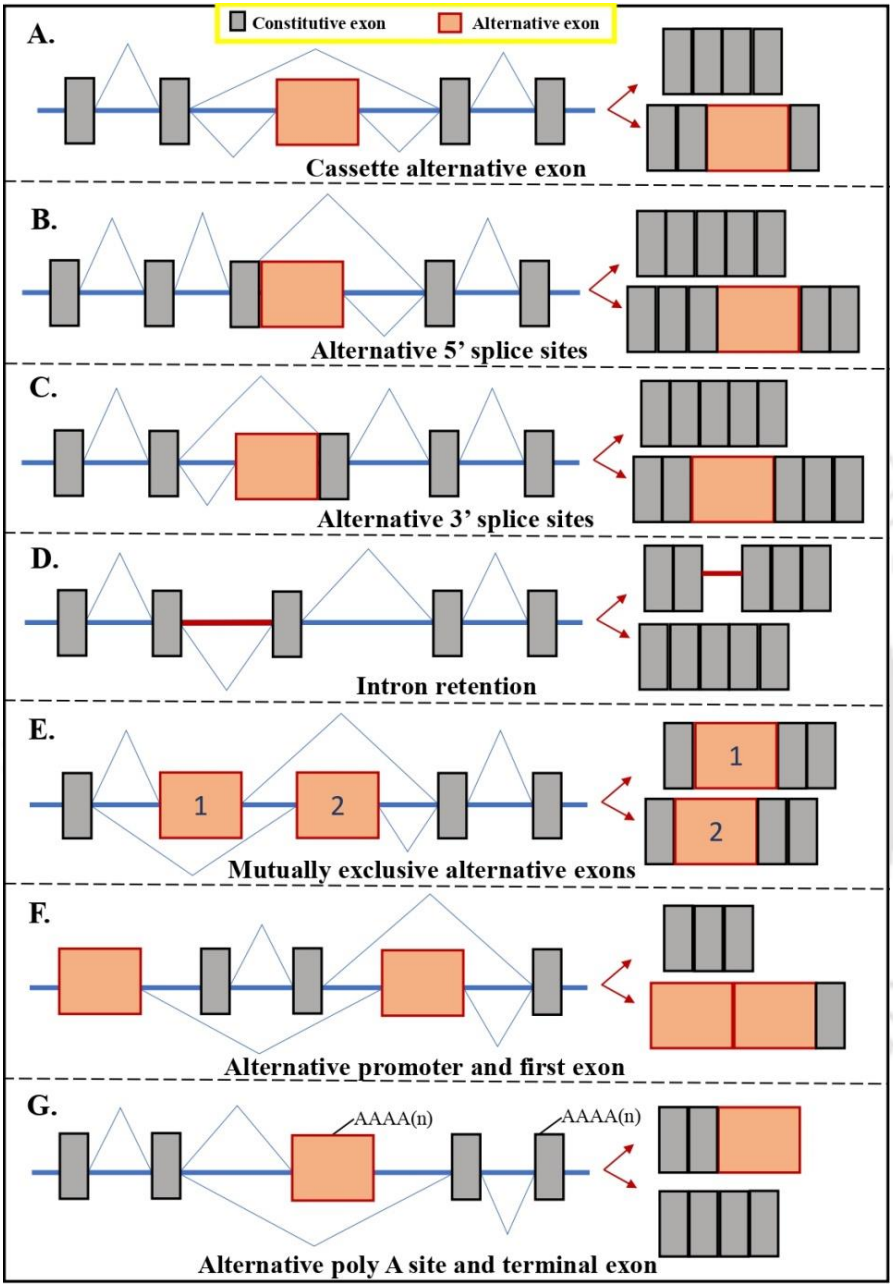


Figure 1.2: Schematic representation of the types of alternative splicing. Alternative splicing can take place through a number of different processes to generate functionally distinct transcripts. (A) Cassette alternative exon, also known as exon skipping (B) Alternative 5' splice sites (C) Alternative 3' splice sites (D) Intron retention (E) Mutually exclusive alternative exons (F) Alternative promoter and first exon (G) Alternative poly A site and terminal exon. Exons and final transcripts are depicted as boxes, while introns are illustrated by lines. Constitutively expressed exons are depicted in grey and alternative exons are depicted in orange. Folded lines are used to connect spliced ends. Intron retained is depicted with red intervening intron in the final transcript (recreated from Blencowe, 2006).

1.1.2 Assembly of mRNPs

In addition to the altering of mRNA molecules, splicing also facilitates the assembly of a number of RNA-binding proteins (RBPs) to form ribonucleoprotein complexes (mRNPs). These depositions of RBPs dictate the fate of spliced mRNAs by regulating and integrating the post-transcriptional steps of gene expression, including export, localization, translation and mRNA decay. The composition of the mRNP complex on a spliced transcript is subjected to extensive remodeling and consequently dynamic over the life of the mRNA. Remodeling may occur by ATP hydrolysis or by diffusion-mediated replacement reactions.

mRNAs destined for translation undergo remodeling at the 5' and 3' end of mRNAs after transport from the nucleus to the cytosol. In the nucleus, the 5' -terminal 7-methylguanosine (m7 G) cap of an mRNA is bound by the cap-binding complex (CBC), whereas in the cytosol, CBC is replaced by eIF4E (a subunit of the translation initiation factor complex eIF4F) [29, 30]. eIF4E aids in the recruitment of small ribosomal subunit and promotes translation initiation. Similarly, polyadenylation factors such as poly(A) binding protein PABPN1 in the nucleus is replaced by PABPC1 in the cytosol, thereby enabling circularization of the mRNA and translation of the mRNA (Figure 1.3) [31].

Interestingly, remodeling of mRNPs may also occur by covalent modifications of the mRNA, such as deadenylation, readenylation, uridylation, editing, and/or base modifications [32-34]. Covalent modifications alter the binding of RBPs and, eventually, the composition of mRNPs. For instance, deadenylation leads to the dissociation of poly(A) binding proteins and, as a result, influences the translation efficiency of the mRNA [35]. Likewise, methylation of adenosines to N6-methyladenosine (m6 A) is another RNA modification that alters the mRNP composition and stability of the mRNA [36].

Among a large number of RBPs implicated in the RNA life cycle, a central component of post-splicing mRNPs is the exon junction complex (EJC). EJC acts as a 'molecular mark' of

exon-exon junctions on spliced mRNA and carries the molecular memory of introns forward to the cytoplasm.

1.2 Exon Junction Complex (EJC)

Characterization of the spliced mRNA led to the discovery that during splicing, the spliceosome deposits a multiprotein complex termed EJC onto mRNAs at a distance of 24 nt upstream of the exon-exon junction in a weak sequence-specific manner [37]. This protein complex remains associated with mRNA till the first round of translation, when the ribosome dissociates it from the spliced mRNA (Figure 1.3) [38].

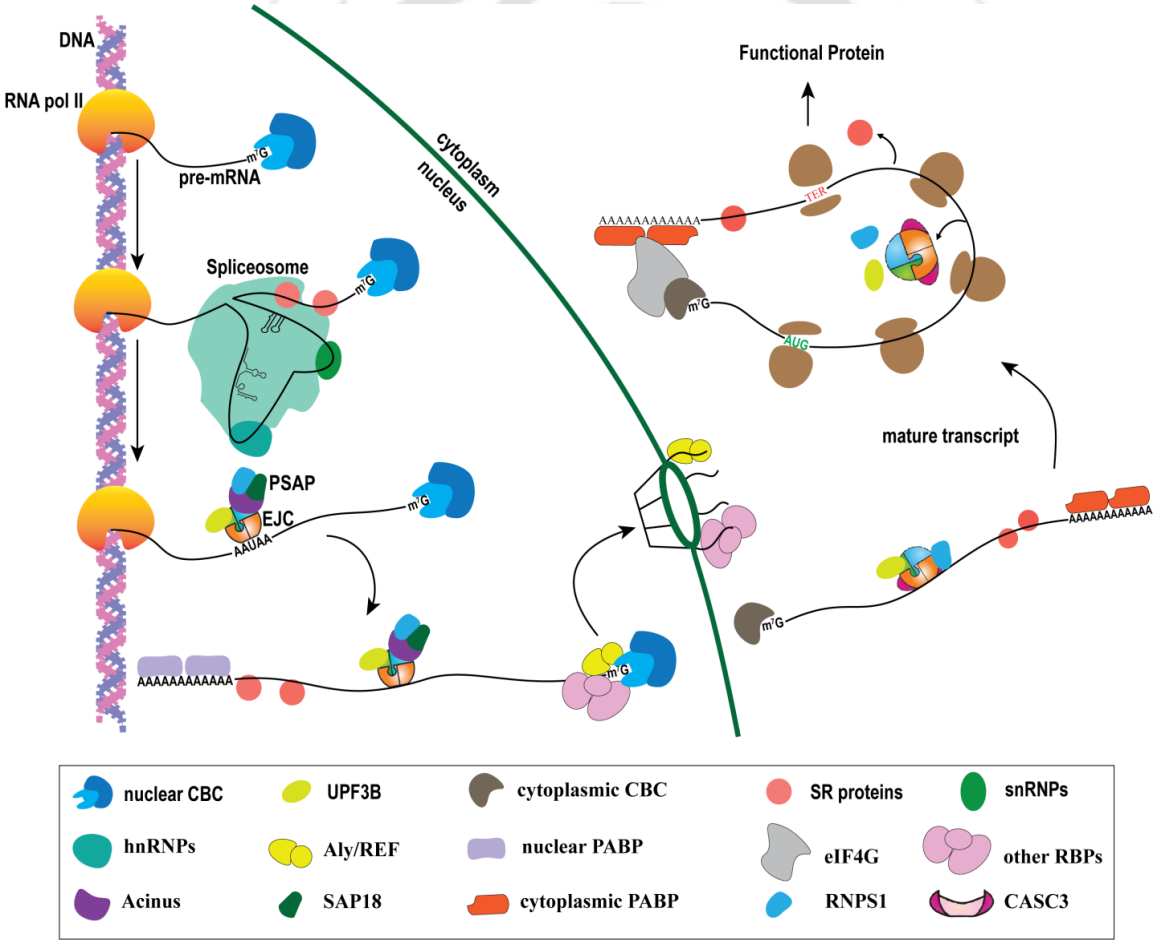


Figure 1.3: Schematic representation of life cycle of EJC and RNPS1. During the process of splicing, spliceosomal complex deposits EJC onto mRNA. The core EJC consists of eIF4A3, MAGOH and Y14. The peripheral EJC proteins are the SR proteins, RNPS1, Acinus, Pinin, SAP18, Aly/REF export factors, and UPF proteins. The core EJC along with RNPS1 remain associated with mRNA till the first round of translation, when the ribosome dissociates it from the spliced mRNA.

The subunits that form the core EJC are eukaryotic translation initiation factor 4A3 (eIF4A3), Y14, and MAGOH [39, 40]. During the course of the journey of mRNA from the nucleus to the cytoplasm, the EJC core acts as a binding platform for many different RBPs. This dynamic composition of EJC couples pre-mRNA splicing with other post-splicing processes (Figure 1.3).

Intriguingly, spliceosomal B and C complexes contain the EJC core proteins, eIF4A3, MAGOH and Y14, indicating that these proteins are present in the spliceosome prior to exon ligation (Figure 1.1) [41-43]. eIF4A3 forms a stable complex with CWC22, a core splicing factor, enabling recruitment of eIF4A3 to the spliceosomal complex [44, 45]. It is believed that MAGOH–Y14 and eIF4A3–CWC22 are recruited to spliceosome independently [46]. CWC22 maintains eIF4A3 in an open conformation; as a result, the binding of eIF4A3 to RNA and MAGOH-Y14 is restricted [46, 47]. Subsequently, eIF4A3 is released from CWC22 to allow eIF4A3 to bind RNA. However, it remains to be determined how eIF4A3 and MAGOH-Y14 are recruited onto spliced mRNA from the C complex. In this regard, RNA helicase IBP160 (intron-binding protein of 160 kDa) is probably involved in the recruitment of EJC proteins to the spliced mRNA. In the C complex, IBP160 interacts with all three EJC core proteins prior to EJC deposition [48]. Nevertheless, the role of IBP160 in the recruitment of the core EJC proteins onto the spliced mRNA is not yet understood.

Structural and reconstitution studies reveal how the core EJC associates with spliced RNA. eIF4A3 is a DEAD-box RNA helicase and comprises of two globular domains (the N-terminal RecA1 domain and the C-terminal RecA2 domain) that are linked by a flexible linker [49]. The two RecA-like domains of eIF4A3 are involved in RNA binding and ATP hydrolysis. In the presence of ATP, the two domains attain a closed conformation and bind to the sugar–phosphate backbones of RNA. MAGOH-Y14 heterodimer then interacts with the two domains of eIF4A3 and locks the core EJC on the RNA by restraining the conformational change upon ATP hydrolysis. The MAGOH-Y14

heterodimer is able to interact with only the ATP and RNA-bound conformation of eIF4A3 [50, 51].

During the export of the mRNA from the nucleus to the cytoplasm, EJC remains bound to the mRNA (Figure 1.3). In the cytoplasm, EJC is removed from the mRNA by the ribosome during the pioneer round of translation, suggesting that EJC remains associated with mRNA until all the regulatory steps of RNA metabolism are completed. The translation-dependent disassembly of EJC is carried out by the ribosome-associated factor PYM (partner of Y14 and MAGOH) [52]. PYM forms a trimeric complex with MAGOH-Y14 via its N-terminal domain, whereas the C-terminal domain of PYM interacts with the 40S subunit of the ribosome. The association of MAGOH-Y14 to either PYM or eIF4A3 is mutually exclusive and leads to steric hindrance; consequently, PYM triggers EJC dissociation by destabilizing MAGOH-Y14 association with eIF4AIII [52, 53]. Notably, it is shown that ribosomes can dissociate EJCs in the absence of PYM. Hence it is suggested that the function of PYM might be to sequester dissociated MAGOH and Y14 to prevent EJC reassembly in the cytoplasm [52]. The heterodimer is imported back to the nucleus by IMP13 (karyopherin receptor importin 13). PYM is dissociated from MAGOH-Y14 following the binding of MAGOH-Y14 heterodimer with IMP13 [54, 55]. Nonetheless, the molecular mechanism by which eIF4A3 is imported back to the nucleus is unknown.

Studies have shown a varied role of EJC at different stages of the mRNA life cycle. To this end, the core EJC dynamically interacts with many peripheral proteins and these peripheral proteins impart functional roles to EJC (Figure 1.3) [56]. Some notable peripheral EJC proteins are the SR proteins, RNPS1, Acinus, Pinin, SAP18, Aly/REF export factors, and UPF proteins which define the wide functions of EJC from splicing in the nucleus to other functions in the cytoplasm. In the cytoplasm, EJC enhances translational efficiency and triggers nonsense-mediated mRNA decay (NMD), a surveillance pathway that mediates the degradation of mRNAs containing premature stop codons. Thus, the EJC ensures that the transcripts are correctly processed and only the error-free transcripts are favored for translation.

1.3 RNA-binding protein with serine-rich domain 1 (RNPS1)

One of the important proteins present in the spliceosome before exon ligation and associates with core EJC is RNPS1 (Figure 1.1). RNPS1 remains associated with the EJC on the spliced mRNA and also interacts with CBP80, a subunit of the nuclear 7-methyl-G cap-binding complex in the cytoplasm. However, RNPS1 does not interact with eIF4E in the cytoplasm, suggesting that RNPS1 remains associated with the spliced mRNA in the cytoplasm until the first round of translation and, as a result, is involved in mRNA biology both in the nucleus and cytoplasm (Figure 1.3) [38]. The details of RNPS1 structure and function will be discussed in the following chapters.

1.3.1 Structure of RNPS1

The amino acid sequence of RNPS1 is related to the serine/arginine-rich (SR) protein family (Figure 1.4A). RNPS1 consists of an N-terminal serine-rich domain (S domain), a middle RRM domain and a carboxy-terminal arginine/serine/proline-rich (RS/P) domain. The RRM domain consists of conserved RNP-1 and RNP-2 sub-motifs and a conserved decapeptide box [57]. In contrast, RNPS1 mediates protein-protein interaction with its RRM domain which is generally used by the SR proteins to bind nucleic acids. The RRM domain of RNPS1 is also critical for the assembly of two different ternary complexes, named ASAP (apoptosis and splicing associated protein) and PSAP (contains PININ, SAP18 and RNPS1) [58]. Interestingly, RNPS1 associates with the core EJC as a member of fully assembled ASAP or PSAP complex. In this regard, it has been shown that mutant RNPS1 (termed 176 amino acid), which is unable to form PSAP/ASAP complex, also fails to interact with eIF4A3 and spliced mRNA [59].

Acinus, one of the subunits of the ASAP complex, is a nuclear factor and exists in three isoforms: Acinus-L, Acinus-S and Acinus-S' (Figure 1.4B). Acinus-L is the longest isoform (1341 amino acids), while Acinus-S and Acinus-S' contain 583 and 568 amino acids, respectively. Acinus isoforms share a common central RRM domain, RSB (RNPS1-SAP18

binding) motif and a C-terminal domain [58, 60]. The isoforms differ only in their amino-terminal domain. Compared to Acinus-S and Acinus-S', Acinus-L contains an additional

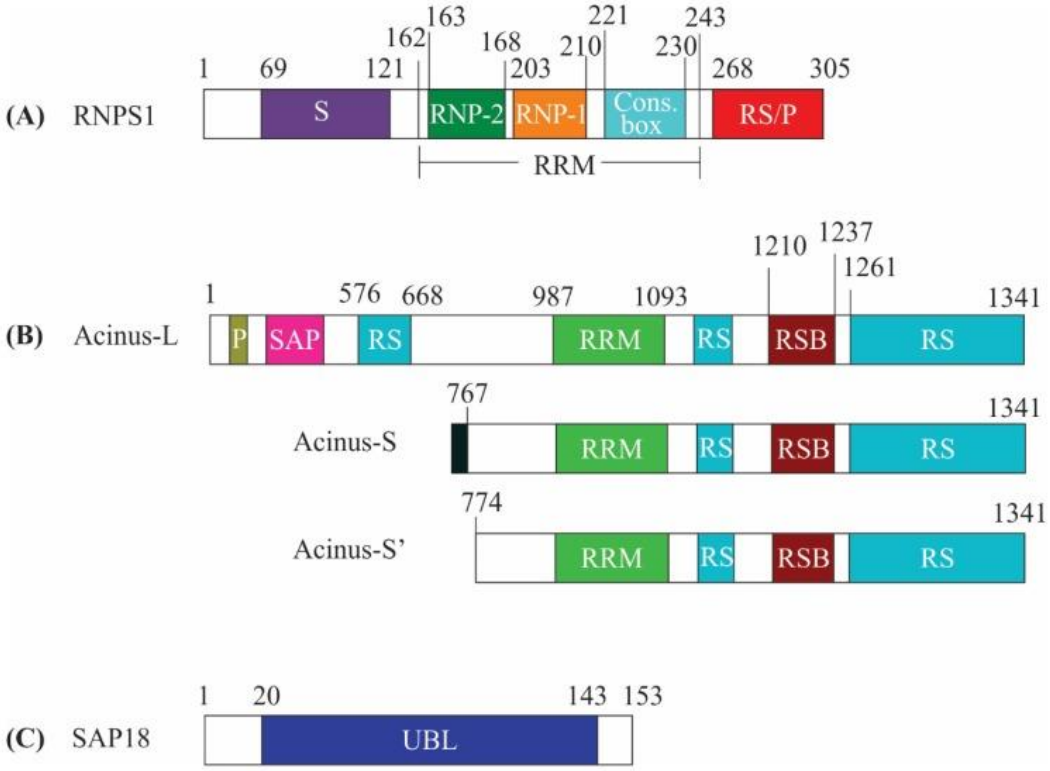


Figure 1.4: Schematic representation of the domain structure of apoptosis-and splicing-associated protein (ASAP) complex components. (A) RNA-binding protein with serine-rich domain 1 (RNPS1) (B) The three isoforms of apoptotic chromatin condensation inducer in the nucleus (Acinus-L, S and S') are indicated. (C) Sin3-associated protein of 18 kDa (SAP18).

SAP [after SAF-A/B, Acinus and PIAS (STAT inhibitors)] motif and an RS domain in its amino-terminal domain [61]. Nonetheless, all the isoforms of Acinus can form the ASAP complex and hence, have no functional differences [39, 62].

The structure of SAP18, another subunit of the ASAP complex, contains a conserved ubiquitin-like (UBL) domain (Figure 1.4C). The UBL domain comprises of antiparallel five stranded β sheets curled around a central α helix [63]. Proteins containing the UBL domain generally serve as a scaffold for the assembly of protein complexes [64]. To this end, SAP18 associates with RNPS1:Acinus heterodimer to form the ASAP complex but does not interact with RNPS1 and Acinus in isolation [39].

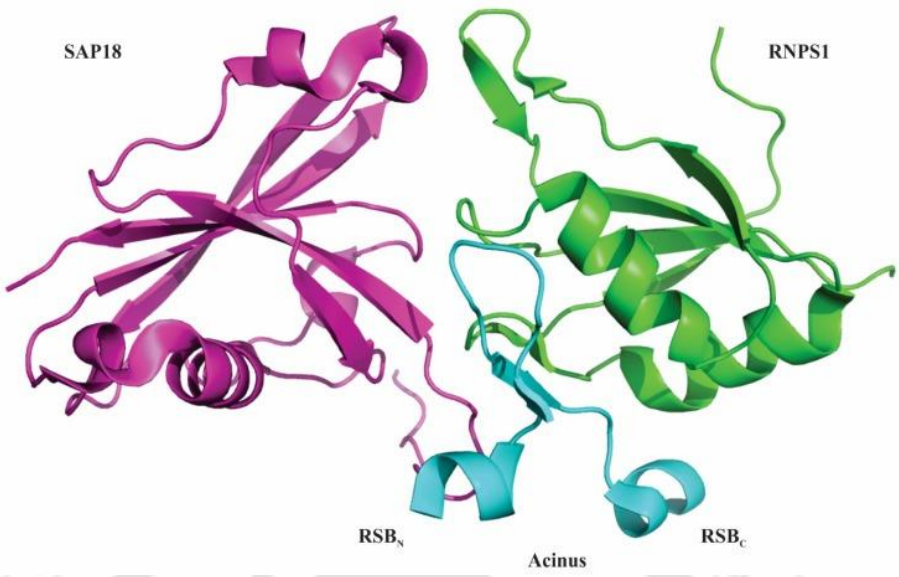


Figure 1.5: Structure of ASAP complex. RNA-binding protein with serine-rich domain 1 (RNPS1) is represented in green, apoptotic chromatin condensation inducer in the nucleus (Acinus) is in blue and Sin3-associated protein of 18 kDa (SAP18) is in pink. The interacting surface of ASAP complex comprises of the RNA recognition motif (RRM) of RNPS1, the ubiquitin-like (UBL) domain of SAP18 and RNPS1-SAP18 binding (RSB) motif of Acinus. Atomic coordinates of ASAP complex with PDB accession code 4A8X were modelled with PyMol (http://www.pymol.org/) according to Murachelli et al. 2012.

In 2012, Murachelli et al. solved the crystal structure of a minimal ASAP complex (Figure 1.5). They utilized 159-244 residues of human RNPS1, 656-683 residues of *D.melanogaster* Acinus and 14-143 residues of mouse SAP18. The crystal structure revealed that the RRM domain of RNPS1, the RSB motif of Acinus and the UBL domain of SAP18 constitute the ASAP complex. The N-terminal helical segment of the RSB motif of Acinus interacts with SAP18 and the carboxy-terminal helical segment of the Acinus RSB motif interacts with RNPS1 [58].

Intriguingly, RNPS1 and SAP18 also associate with Pinin (PNN), a splicing coactivator, to generate an alternative protein complex known as PSAP. Similar to the RSB motif of Acinus, Pinin also utilizes the RSB motif to interact with RNPS1 and SAP18. A competition experiment between Acinus and Pinin revealed that the incubation of PSAP complex with an excess amount of Acinus did not yield ASAP complex. This observation suggests that the RSB motif of Pinin might have a higher binding affinity for RNPS1 and

SAP18 than Acinus. Furthermore, it is envisioned that the ASAP and PSAP are probably recruited separately to the core EJC [58].

1.3.2 Functional roles of RNPS1 in mRNA biology

As an important component of the post-splicing mRNP, RNPS1 regulates many stages of the mRNA life cycle. In the nucleus, RNPS1 regulates splicing activity, while in the cytoplasm, RNPS1 is involved in augmenting translation efficiency and mRNA quality control.

(a) RNPS1 regulates splicing activity

RNPS1 was originally discovered as an activator of constitutive splicing. Initial *in vitro* splicing assay showed the splicing stimulatory role of RNPS1 on various model mRNAs. For instance, in the presence of limiting amounts of splicing factor SRSF1, RNPS1 promotes constitutive splicing of *β-globin* transcript. Likewise, RNPS1 also stimulates the constitutive splicing of pre-mRNAs such as *immunoglobulin μ-chain* (*IgM*), *HIV-tat*, *Drosophila ftz* and *δ-crystalline* [57]. In line with this observation, a loss of function study in HeLa cells revealed the role of RNPS1 in efficient splicing of *AURKB* (Aurora B kinase) and *MDM2* (murine double minute 2) pre-mRNA [65]. Studies have also confirmed the role of RNPS1 in enhancing the formation of spliceosomal A complex and its presence in spliceosome B and C complexes [41, 66]. Of note, RNPS1 promotes splicing of pre-mRNA with GU-AG intron (spliced by major spliceosome) but has no effect on pre-mRNA with AT-AC intron (spliced by minor spliceosome), suggesting that RNPS1 is primarily involved in the major splicing pathway [57].

Studies in *Drosophila* indicated the importance of RNPS1 in splicing out weak and long neighbouring introns, suggesting its function in exon definition. Exon definition is a splicing process where splice sites across small exons are recognized rather than long introns [67]. For instance, depletion of core EJC (eIF4A3, Mago and Y14) or RNPS1 results in the skipping of exons in heterochromatin genes having large intron (greater than 250 bp) [68, 69]. In this regard, several theories have been proposed to explain the

role of RNPS1 and EJC in exon definition. One of the theories states that EJC probably recruits SR proteins and aids in exon definition. Another probable reason is that EJC masks the pseudo splice-sites present within pre-mRNA, preventing exon skipping. The third possible explanation is that EJC facilitates the association of the spliceosomal complex with splice sites [68].

Similarly, RNPS1 also suppresses transposable elements by controlling the splicing activity of piwi transcripts in the *Drosophila* germline and surrounding somatic follicle cells [70, 71]. Piwi is a part of the piRNA pathway that represses transposable elements, mainly in gonads [72]. The *piwi* pre-mRNA has a weak fourth intron (large intron) and depletion of core EJC (Mago and Y14), Acinus, or RNPS1 causes retention of the fourth intron, consequently decreasing piwi protein level. Additionally, the splicing of this weak fourth intron depends on the splicing of neighboring strong flanking introns. Hence, it has been proposed that splicing of strong neighboring intron causes deposition of EJC together with RNPS1 at exon-exon junction in the close vicinity of a weak intron, eventually facilitating spliceosome assembly at weak intron (Figure 1.6) [70, 71].

Furthermore, in *Drosophila,* RNPS1 also contributes to the regulation of Wingless (Wg/Wnt) signaling. Wnt signalling is important for cell proliferation, cell polarity, pattern formation, cell migration and stem cell maintenance. Depletion of core EJC or RNPS1 deregulated the splicing of Wnt signaling determinant, *dlg1* transcript, resulting in reduced expression of Dlg1. Dlg1 safeguards scaffold protein of Wnt signaling, Dsh, from lysosomal degradation. Consequently, loss of RNPS1 decreases the expression of Dlg1 and Dsh proteins, resulting in decreased Wnt signaling [73].

Interestingly, the transcriptome-wide analysis revealed that knockdown of RNPS1 in human cell lines leads to the mis-splicing of a large number of mRNAs due to the de-repression of cryptic splice sites [59]. Cryptic splice sites are pseudo splice sites localized on the pre-mRNA that are masked and repressed by the deposition of RNPS1 and EJC core under normal conditions. It has been observed that RNPS1 bound to an upstream

EJC represses the use of downstream pseudo 5' SS, thus acting as a protection sphere on the bound transcript (Figure 1.7). Nevertheless, the effect of EJC-bound RNPS1 is limited to the region adjacent to the EJC binding site. Remarkably, depletion of SAP18 and Pinin,

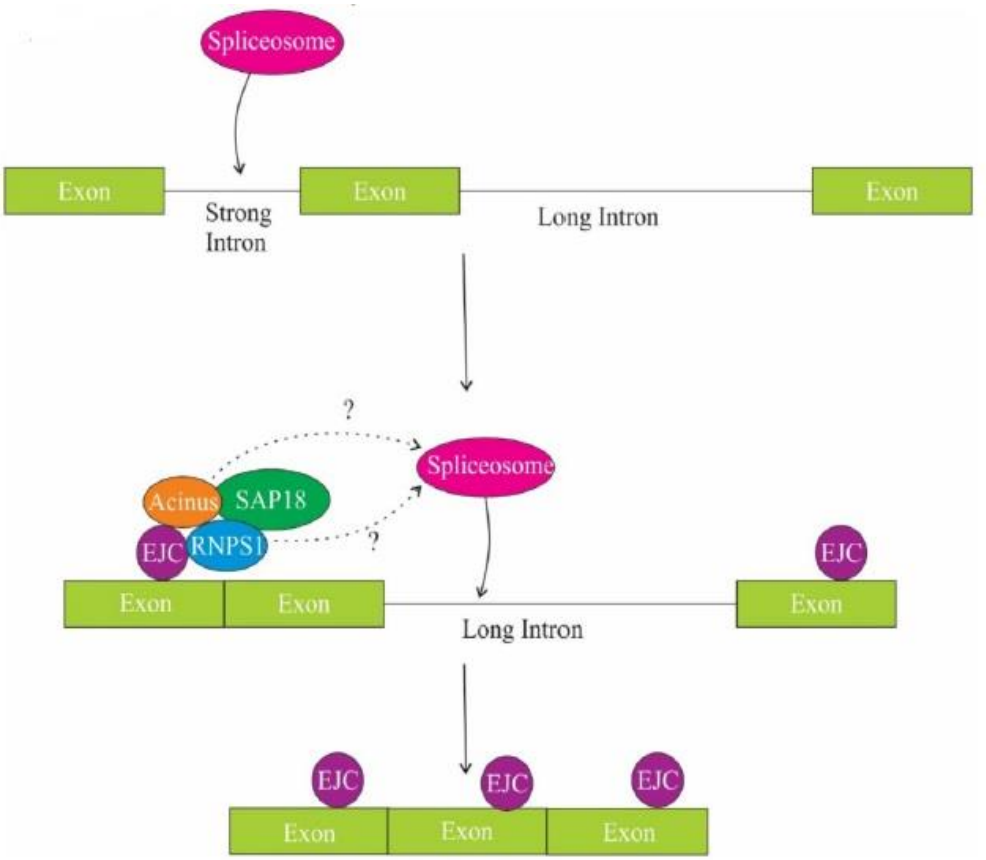


Figure 1.6: A proposed mechanism of splicing regulation by RNPS1. One of the mechanisms of regulation of splicing by RNPS1 is splicing of long sub-optimal introns. Splicing of strong adjacent introns deposit EJC at exon junctions and EJC in turn recruits RNPS1 in the close proximity of weak intron, thus facilitating spliceosome assembly and splicing of the long weak intron.

but not Acinus, causes de-suppression of cryptic splice sites, indicating PSAP-bound EJC is involved in cryptic splice site repression rather than ASAP-bound EJC. Furthermore, it was revealed that RNPS1 is the effector molecule in PSAP-bound EJC mediated 5' splice-site regulation, whereas EJC and other PSAP proteins are only needed for proper deposition of RNPS1 to the mRNA. To this end, tethering of Pinin and SAP18 to reporter mRNA in RNPS1 knockdown cells did not rescue splicing defects, while tethering PSAP-incompatible RNPS1 (176) could rescue splicing defects [59].

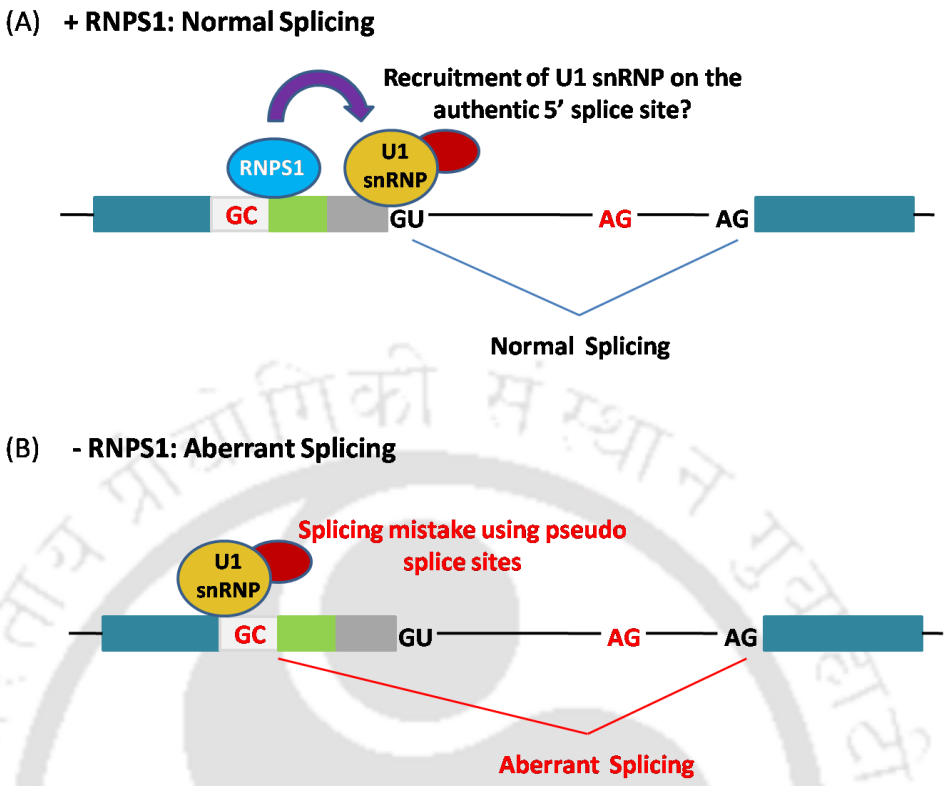


Figure 1.7: A proposed mechanism by which RNPS1 maintains fidelity of pre-mRNA splicing [75]. (A) Deposition of RNPS1 masks pseudo splice sites or cryptic splice sites present on the pre-mRNA and prevent loss of exonic sequences. (B) Loss of RNPS1 exposes the pseudo splice sites on pre-mRNA which are then recognized by the spliceosomal complex, leading to aberrant splicing (recreated from Fukumura et al 2018).

Another parallel study also reported a similar observation that putative splice sites are frequently generated by newly formed exon junctions, causing recursive splicing and loss of exonic sequences. In this regard, core EJC proteins, RNPS1 or Pinin suppress the splicing of reconstituted splice sites [74]. Altogether, this indicates a critical role of RNPS1 in maintaining splicing fidelity and, as a result, safeguarding transcriptome integrity.

Interestingly, a genome-wide CRISPRCas9-based screen in a human cell line revealed RNPS1 as a significant alternative splicing regulator of neuronal microexons [76]. Microexons are set of very small exons of size 3-30 nucleotides (1-10 amino acids). Knockdown of RNPS1 results in enhanced skipping of the neuronal microexons. A fascinating observation is that these microexons are mis-spliced in the brains of

approximately one-third of analyzed autistic subjects [77]. Furthermore, these individual microexons have been found to play vital functions during neurodevelopment and other brain functions [78]. Microexons most often encode for amino-acid residues that are surface accessible and therefore modulate protein-protein interactions. In fact, many reports have demonstrated that the inclusions of neuronal microexons are critical for promoting protein-protein interactions in neuronal cells [79, 80]. Additionally, it has been reported that most microexons exhibit high inclusion at later stages of neuronal differentiation in genes associated with axonogenesis and the development and function of synapses [81]. This implies that the differential regulation of splicing of microexons is imperative for the proper development and function of the nervous system. Because microexons are heavily dependent on splicing regulators like RNPS1 for their regulation, this raises the question of whether the expression of RNPS1 is altered to cause changes in alternative splicing of microexons. Collectively, the reports suggest RNPS1 as a key regulator of both constitutive and alternative splicing.

(b) RNPS1 enhances translation efficiency and 3'end processing of pre-mRNA

Apart from the functions of RNPS1 in splicing, RNPS1 also functions in translation. Several reports provide convincing evidence that spliced transcripts produce a high quantity of protein per mRNA molecule compared to their identical transcripts lacking introns [82]. Studies confirmed that this enhancement in translational efficiency is directly correlated to an increased association of spliced mRNAs with polysome. Remarkably, the tethering of RNPS1 and other EJC proteins to an intron-less reporter enhanced both translation yield and polysome association, indicating an important role of RNPS1 in aiding polysome assembly [83].

The mRNA processing at the 5' and 3' end is of key importance in augmenting translational efficiency, as mRNA stability majorly depends on the presence of cap at the 5' end and the length of poly (A) present at the 3' end. The export of mRNA from the nucleus to the cytoplasm is also influenced by polyadenylation. In this regard, tethering of RNPS1 or SRm160 (Serine/Arginine Repetitive Matrix 1) to the intronless *CAT* and β-

globin transcript increased the 3' end processing and mRNA stability, resulting in increased expression of the intronless reporter [84]. One of the probable underlying reasons is the recruitment of the 3' end processing machinery, as SRm160 is known to interact with CPSF-160 [85]. CPSF-160 plays a role in 3' end processing by binding to the AAUAAA signal sequence in pre-mRNA and associating with poly (A) polymerase and other processing factors to activate 3' end cleavage and polyadenylation [86]. Together this suggests that the interaction of RNPS1 with SRm160 probably influences the 3' end processing of mRNAs; as a result, RNPS1 enhances the expression of spliced mRNAs.

(c) Function of RNPS1 in nonsense-mediated mRNA decay

Nonsense-mediated mRNA decay (NMD) is a process that degrades faulty transcripts with premature termination codons (PTCs) that could generate harmful truncated proteins. Besides maintaining mRNA quality control, NMD also maintains homeostasis by controlling the expression of normal genes. NMD is generally activated when a PTC is localized more than 50 nt upstream of the exon-exon junction in a transcript [87, 88]. NMD is generally classified into two types, EJC-enhanced NMD and EJC-independent NMD. In EJC-enhanced NMD, EJC bound downstream of the PTC serves as a signal for the assembly of NMD machinery on the faulty mRNA [89].

Tethering assay showed that RNPS1 tethered to a PTC-containing reporter mRNA triggers NMD and, as a result, promotes degradation of the mRNA [90, 91]. This data suggest that RNPS1 has a functional role in NMD. In a bid to comprehensively characterize the role of RNPS1 in NMD, several transcriptome-wide analyses were performed. However, the global analysis revealed that RNPS1 affects the NMD efficiency of only a subset of transcripts [92].

Collectively, the studies indicate that RNPS1 functions as an essential global regulator of splicing, while it has a minor role in mRNA quality control.

1.4 Post-transcriptional gene regulation

The crucial role of RNPS1 in RNA biology raises a fundamental question, whether the expression of RNPS1 is regulated to orchestrate differential splicing activity observed during cellular differentiation and organism development. In this regard, the regulation of *RNPS1* gene expression and its effects on the cellular program is still unexplored.

The sequencing of the human genome and several other higher eukaryotic genomes has shown that a very small proportion of genetic material (~1.5%) codes for protein [93]. Indeed, most regions of the genomic DNA participate in the regulation of gene expression, which either takes place at the transcriptional level or the post-transcriptional level. The transcriptional mode of gene regulation controls the transcription of a gene, whereas the post-transcriptional mode of gene regulation determines the outcome of the transcribed mRNA, including stability, subcellular localization and translation efficiency.

In eukaryotes, the processed mRNA molecule has a tripartite structure comprising a 5' untranslated region (5' UTR), a coding region and a 3'untranslated region (3' UTR). The post-transcriptional control of gene expression is mainly mediated by the untranslated region. The essential role of UTR in gene regulation is reinforced by the fact that mutations in the UTR can cause serious pathology [94, 95]. For example, myotonic dystrophy type 1 is caused by the amplification of CTG repeats in the 3'UTR of the *DMPK* (dystrophia myotonica protein kinase) gene [96].

In comparison to 5'UTR, 3'UTR contains a wealth of regulatory elements to modulate the fate of mRNA. The 5'UTR or the coding region has to accommodate the translational machinery and is therefore under rigid structural constraints. Hence, evolution has taken advantage of the greater accessibility of 3'UTR to harbor regulatory factors [97].

Regulation by 3'UTRs is mediated mainly by two leading players, RNA-binding proteins and microRNAs (miRNAs). Compared to DNA-binding protein, whose interaction with

the cis-element relies on the primary structure of DNA, whereas the regulation at the RNA level depends on both the primary and secondary structure of RNA [98].

1.4.1 RNA-binding proteins (RBPs) in post-transcriptional gene regulation

RBPs (RNA-binding proteins) are trans-acting elements that regulate mRNA stability, transport, and translation regulation by binding to the regulatory regions (cis-acting elements) within the 5' or 3' UTR of mRNA. RBPs use one or more RNA-binding domains (RBDs) to bind to RNA targets. Some well-studied RNA-binding domains are: RNA recognition motif (RRM), DEAD/DEAH box, zinc finger (ZnF), K-homology (KH) domain, RGG (Arg-Gly-Gly) box, Pumilio/FBF (PUF or Pum-HD) domain, Piwi/Argonaute/Zwille (PAZ) domain, cold-shock domain, Sm domain and the double-stranded RNA-binding domain (dsRBD) [99]. Bioinformatics investigations using these motifs indicated that eukaryotic genomes encode a high number of RBPs. In yeast, 5–8% of genes encode proteins that probably function as RBPs, while around 2% of the genome in *C. elegans* and *D. melanogaster* likely encodes RBPs [100]. However, the number of RBPs is expected to be substantially larger, as there are likely to be more RNA-binding domains yet to be found. This finding raises the question, why do eukaryotes require such a large number of RBPs? One hypothesis is that when eukaryotes acquired complex post-transcriptional mechanisms to fine-tune gene expression, the number of RBPs required to function in these processes expanded at the same time [101].

The RBPs bind to specific cis-acting sequences to mediate post-transcriptional gene regulation. The following are a few examples of well-studied cis-acting elements and RBPs:

AU-rich elements: One of the most prevalent regulatory elements in the 3' UTR of mRNA is the adenylate uridylate (AU-rich) elements (AREs), which affect mRNA stability, translation progress, and alternative pre-mRNA processing. AREs are present generally in short-lived mRNAs. The pentamer "AUUUA" motif is the basic motif of AREs, which appears in various repetitions. Using chimeric genes, researchers demonstrated that

increasing the number of pentameric motifs leads to mRNA instability. A family of proteins known as ARE-binding proteins (ARE-BPs) recognises AREs. Some well-known ARE-BPs are nuclear factor NF90, the family of Hu proteins (HuR) and T-cell-restricted intracellular antigen 1 (TIA-1) [98].

CU-rich elements and differentiation control elements: The CU-rich element (CURE) and the differentiation control element (DICE) were once thought to be the same regulatory element; however, they are now recognized as two separate CU-rich elements. The polypyrimidine-tract binding protein (PTBP) is one of the well-studied CURE binding proteins. PTB has the ability to bind to distant places on the same RNA molecule, thereby introducing RNA loops that are necessary for ribosome recruitment to internal ribosomal entry sites. Polyadenylation, cap-dependent translation, and mRNA stability can all be influenced by PTBP [102].

GU-rich elements: The pentamer motif "GUUUG" is the basic sequence motif of GU-rich elements (GRE). GREs are made up of 2 to 5 overlapping pentamers, but unlike AREs, the efficacy of mRNA decay does not appear to be affected by the number of motifs. In addition, GREs have a functional resemblance to ARE. GREs have a role in post-transcriptional processes such as deadenylation and mRNA degradation. GU-rich elements interact with proteins from the CELF family, such as CUGBP [103].

CA-rich elements: Another most common dinucleotide repeat in the human genome are CA-rich elements (CAREs). They can be found in both coding and non-coding regions and have functions in mRNA stability. Heterogeneous nuclear ribonucleoprotein L (hnRNP L) has been identified as one of the important CARE interacting proteins. The binding of hnRNP L changes the vulnerability of mRNAs to endo- or exo-nuclease destruction [104].

Despite the fact that RBPs have been studied widely for several decades, there remains much to learn about their roles and interactions in gene regulation. Several novel RBDs have recently been found and are being characterized until date. Even well-studied RBPs are being discovered to have new roles. We have only begun to understand the dynamic

interplay of the RBPs and their corresponding RNA targets in response to the cellular state and external stimuli.

1.4.2 MicroRNAs (miRNAs) in post-transcriptional gene regulation

MicroRNAs are a class of short non-coding RNAs that post-transcriptionally regulates gene expression. MicroRNA exerts gene regulation through the formation of the RISC complex (RNA-induced silencing complexes) composed minimally of AGO protein and a miRNA molecule. The RISC complex then serves as a platform to recruit proteins that decide the fate of target mRNAs of the RISC complex [105].

An individual miRNA can regulate the expression of more than one target mRNA, and the expression of a single mRNA may be controlled by many miRNAs. The specific binding sites of miRNA are termed as seed sequence (also known as miRNA response element), which is mostly present within the 3' UTRs of target mRNAs (Figure 1.8). MicroRNAs bind these seed sequences by base pairing at positions 2-7 from their 5' end [106]. In a few mRNAs, seed sequences are also present in the 5' UTR and coding region. However, the regulatory effect of the miRNA seed sequence in the coding region is weaker in comparison to the 3'UTR seed sequence [107]. If miRNA seed sequence is present in the 5' UTR, it generally causes translation inhibition. Nevertheless, transcriptome-wide mapping of miRNA binding sites (AGO HITS-CLIP) reported that many miRNA-target interactions also follow non-canonical seed pairing rules [108].

MicroRNAs regulate gene expression by promoting accelerated target mRNA degradation or repressing translation (Figure 1.8). Transcriptome-wide analysis together with ribosome-profiling in cultured mammalian cells, revealed that the target mRNA degradation is the most prevalent mechanism of miRNA-mediated regulation (66-90%) [109]. Evidence suggests that the degradation of target mRNA is catalyzed by the 5' to 3' mRNA decay pathway. The poly(A) tail is first removed from the target mRNA by the CCR4–NOT deadenylase complex and PolyA Nuclease proteins PAN2–PAN3 [110]. Deadenylated transcripts are then decapped by decapping protein 2 (DCP2). Lastly,

decapped and deadenylated transcripts are degraded by the 5'-to-3'exoribonuclease 1 (XRN1) [111].

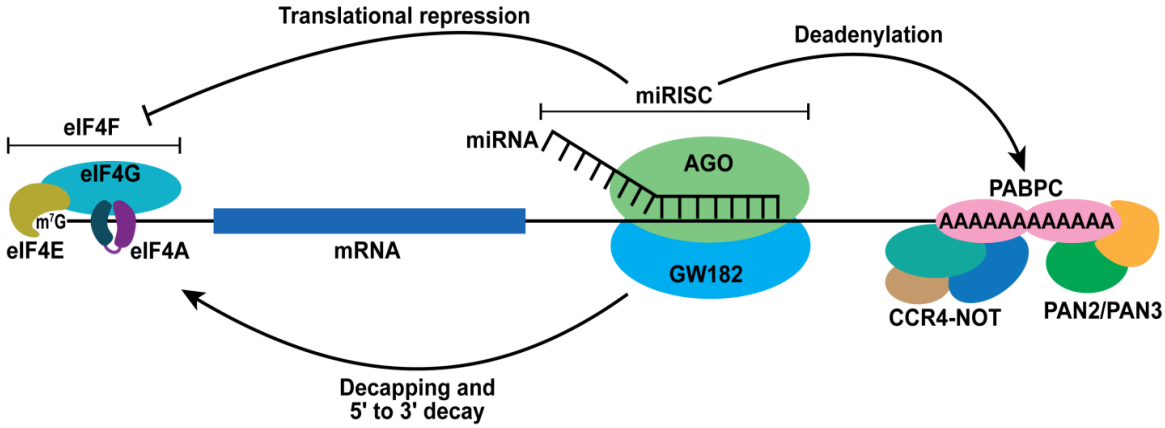


Figure 1.8: Proposed mechanisms of miRNA mediated regulation of gene expression. The miRNA-RISC complex binds by partial complementarity to seed region in the 3'UTR of the target mRNA. Subsequently, the RISC complex recruits proteins involved in deadenylation of target mRNA along with decapping enzymes. Besides mRNA degradation, miRNA also sometimes inhibits translation by interacting with the translation initiation factors (recreated from Jonas and Izaurralde 2015).

Evidence suggests that translation repression accounts for 6-26% of the miRNA-mediated regulation of gene expression [109]. Several studies report that RISCs target translation initiation factors, eIF4A RNA helicases (eIF4A1 and eIF4A2), and as a result, hinder the process of ribosome scanning [112, 113]. Hence, it is perceived that miRNAs inhibit the translation of the target mRNAs at the initiation step (Figure 1.8).

Intriguingly, some miRNAs are also involved in translation upregulation. One such condition is the consequence of the G0 stage on miRNA-dependent gene regulation. It has been observed that GW182, an important component of the RISC complex, is downregulated in the G0 state and immature oocytes; as a result, miRNA-mediated negative regulation is abrogated. In the absence of GW182, another protein FXR1 (Fragile-X mental retardation protein 1), interacts with AGO2, thus modulating the function of the RISC complex from negative to positive regulator [114]. It has been shown that FXR1-AGO2-RISC bound to 3' UTR of a target mRNA can interact with the translation initiation factor, eIF4E. This interaction supports the formation of a closed-loop even in the absence of a poly(A) tail and, hence, could activate translation directly [115].

(a) miRNAs are regulators of splicing factors

Splicing-factor levels are frequently altered in human malignancies, but only a small percentage of these tumors have copy number changes. Remarkably, recent studies demonstrated that miRNAs can control the expression of splicing factors. For instance, miRNAs miR-30a-5p, miR-181a-5p and miR-216b-5p negatively regulate the expression of SR protein, SRSF7, in renal cancer cells [116]. Altered expression of miRNAs is frequently reported in cancers and their function as promoters or repressors of tumorigenesis is becoming increasingly evident. In line with this, silencing of SRSF7 decreased the proliferation rate of the renal cancer cell line, Caki-2 and miR-216b could recapitulate the effects of SRSF7 silencing. Further, miR-216b-5p mediated downregulation of SRSF7 also shifted the splicing pattern of a downstream target gene of SRSF7, *SPP1* [116]. Hence, miRNAs can also act as indirect regulators of alternative splicing.

Similarly, miRNAs are also involved in the regulation of splicing factors, SRSF1, SRSF2 and hnRNP A1 in renal cancer cells. Additionally, SRSF1 and SRSF2 can regulate the expression of their own regulatory miRNA; as a result, SR proteins and their miRNAs are involved in a regulatory feedback loop [117]. Furthermore, the regulation of hnRNP A1 via miRNAs was also demonstrated in colon cancer and acute myeloid leukemia (AML). In AML, downregulation of miR-451 elicits upregulation of its target, hnRNP A1, thereby contributing to the oncogenic activity of hnRNP A1 [118]. Altogether, these results imply the importance of miRNA-mediated regulation of splicing factors.

1.5 Deregulation of splicing factors in cancer

Alternative splicing plays a key role in producing cell type and temporal specific splice isoforms, thereby defining the proteome diversity during differentiation and development. Consequently, aberrant splicing appears to be a critical cause of various diseases and disorders, such as cancer [119, 120]. Accumulating evidences suggest that cancer-specific splicing events are involved in the development of cancer hallmarks,

including proliferation, migration, invasion and drug resistance. It is believed that changes in the splicing machinery confer benefits to the tumor cells either through the generation of oncogenic protein isoforms or by altering the balance in cellular isoform ratios (Figure 1.9). Cancer cells exploit the transcriptional plasticity offered by alternative splicing to generate isoforms that promote cell proliferation, migration, or invasion [121]. Accordingly, RNA splicing regulators are considered a new class of oncoproteins or tumor suppressors that contribute to cancer development. Changes in the components of splicing machinery are more relevant in cancer because they alter a plethora of downstream splicing targets, while mutation altering splicing of a single gene generally perturbs only one isoform expression.

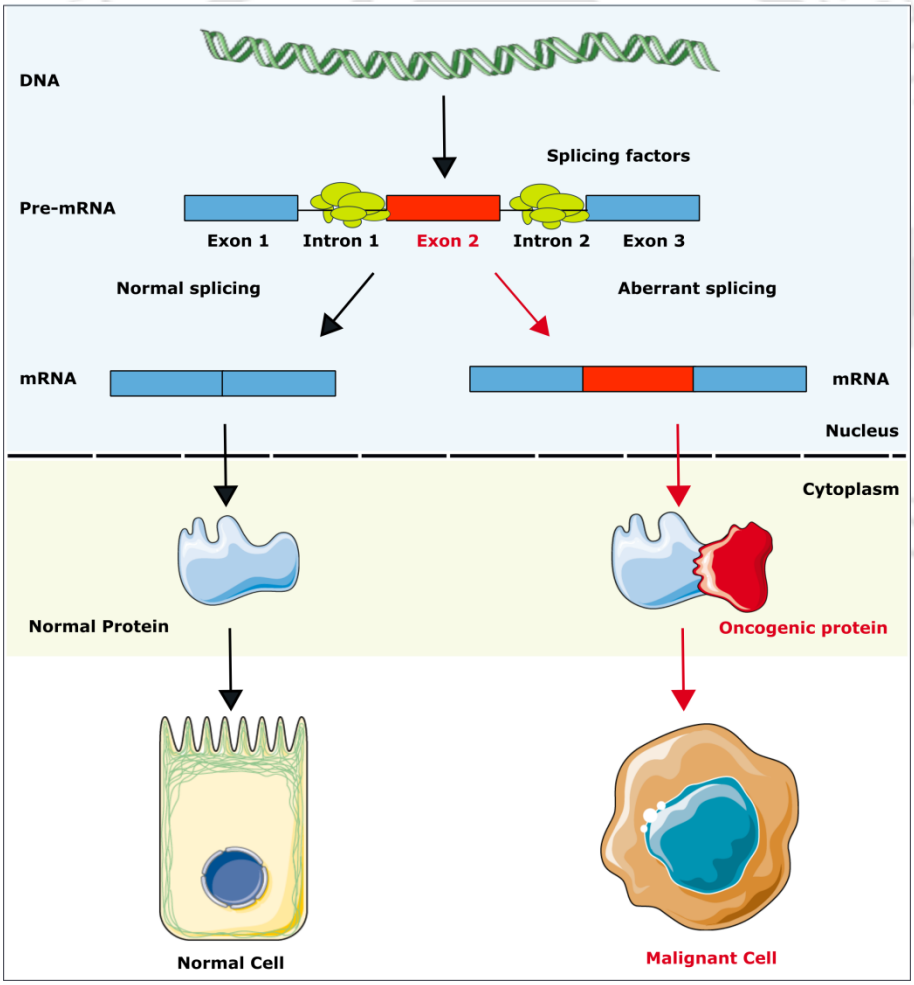


Figure 1.9: The cellular splicing machinery is altered during oncogenic transformation of the cell. In cancerous cell, abnormal splicing of critical mRNAs gives rise to predominant production of protein isoforms with oncogenic properties (recreated from Pajares et al. 2007).

The 5′ splice site is recognized by the U1 snRNP through base-pairing interactions between U1 snRNA and the 5′ end of the intron. Mutations in U1 snRNA cause alterations in 5′ splice site usage, which influence cancer driver genes across many malignancies and lead to poor prognosis in patients with chronic lymphocytic leukaemia [123]. Likewise, mutations in the splicing factor, SRSF2, are often observed in hematological malignancies, including 2% of acute myeloid leukemia (AML), 10% of myelodysplastic syndromes (MDS), and 31–47% of chronic myelomonocytic leukemia [124]. MDS patients with SRSF2 mutations have a lower overall survival rate and a higher rate of progression to AML [125]. Intriguingly, missense mutations in SRSF2 are present around proline 95 amino acid, near the RRM domain, which imparts RNA-binding specificity [126].

Besides mutation, a recent transcriptome-wide analysis revealed that the expressions of splicing factors are frequently deregulated in multiple types of cancer. It was found that several SRSF and hnRNP proteins are overexpressed in tumors and have oncogenic properties [127]. One of the well-known splicing factors, SRSF1, has been reported to be an oncogene. The expression of SRSF1 is frequently upregulated in lung, breast and colon cancer [128-130]. Furthermore, SRSF1 and MYC work in synergistic manner, and their coexpression is linked to a high-grade tumor and shorter survival in lung and breast cancer patients [128, 129]. Moreover, upregulation of SRSF1 is linked to cisplatin and topotecan resistance in lung cancer patients [131]. SRSF1 regulates the alternative splicing of genes affecting several hallmarks of cancer. In this regard, SRSF1 is involved in the upregulation of the oncogenic splice variant of cyclin D1b [132]. SRSF1 also regulates the expression of pro-angiogenic and anti-angiogenic VEGF isoforms [133]. Furthermore, SRSF1 promotes alternative splicing from pro-apoptotic to antiapoptotic variants of caspase 9, Bim and BIN1 [128]. Interestingly, one of the regulators controlling the expression of splicing factor, SRSF1, in cancer is miRNA, particularly miR10b-5p and miR-203a-3p [117].

Similarly, SRSF3 is also upregulated in breast, lung, ovarian, colon, liver, bladder and oral cancer [134, 135]. SRSF3 functions as a proto-oncogene since its overexpression can transform human fibroblasts in vitro, whereas its downregulation causes cancer cell lines to stop growing [134]. SRSF3 controls alternative splicing of genes involved in cancer, such as PKM2 (which affects cell metabolism) and TP53 (which causes cellular senescence) [136, 137].

More recently, it has been reported that the expression of RNPS1 is higher in uterine corpus endometrial carcinoma (UCEC) tumors than in normal tissues. It has been shown that RNPS1 promotes proliferation and poor prognosis of UCEC. The study reports that RNPS1 might regulate the expression of mismatch repair genes such as *MSH2* (MutS homolog 2) and *MSH6* (MutS homolog 6), thereby contributing to UCEC [138]. Nevertheless, a comprehensive characterization of the functional role of RNPS1 in cancer still remains to be examined. Additionally, it remains unclear whether RNPS1 contributes to the generation of cancer-specific splice isoforms. Moreover, it will be interesting to elucidate whether miRNA mediates the regulation of RNPS1 expression in cancer cells. Understanding the transcriptional and post-transcriptional regulation of splicing factors is crucial for discovering new therapeutic targets.

1.6 Aims and significance of the study

RNPS1 is a well known constitutive and alternative splicing regulator that controls and coordinates the cellular transcriptome. Currently we have a clear understanding of the functional roles of RNPS1 in RNA biology, particularly splicing. Owing to the important role of RNPS1 in RNA metabolism, aberrant RNPS1 function may lead to human diseases. To this end, increasing number of evidences showed that several splicing regulators play critical roles in diseases such as cancer. What remains unclear is whether RNPS1, being a guardian of splicing fidelity, has any role in cancer development. The involvement of RNPS1 in cancer was unknown when I began this investigation. Further, I hypothesized that RNPS1 is tightly regulated so that cells can avert the deleterious outcome of aberrant expression of RNPS1. Hence I set out to fill this knowledge gaps by framing the following main objectives of this study.

1. To investigate the pathophysiological condition in which RNPS1 is regulated.

2. To identify specific miRNAs that post-transcriptionally regulates RNPS1.

3. Identification of candidate RNA binding proteins associated with *RNPS1* mRNA.

In the first objective, the functional roles of RNPS1 in cervical cancer cells were systematically investigated. The role of RNPS1 in the development of cancer hallmarks, including proliferation, migration, and invasion of cervical cancer cells were thoroughly assessed. Further, I have attempted to understand the importance of RNPS1 in modulating the alternative splicing of cancer specific genes, thereby generating oncogenic splice variants.

The findings from various groups that miRNAs are involved in the regulation of splicing factors in cancer led us to hypothesize that miRNAs might control the expression of RNPS1. Hence, in the second objective, the post-transcriptional regulation of RNPS1 via miRNAs was examined. The miRNA that specifically regulates RNPS1 in cervical cancer

cell was identified. Subsequently, the miRNA response element within *RNPS1* mRNA was determined.

Besides miRNAs, RBPs are another important regulator of post-transcriptional gene regulation. However, till date, the regulatory proteins controlling the expression of RNPS1 are unknown. Hence, in the third objective, I developed a system termed as, MS2-tagged RNA affinity capture of RBPs, to screen potential RBPs that interact with *RNPS1* 3' UTR. This approach captured the protein interactome profile of *RNPS1* mRNA and will serve as a knowledge base for future studies into RBP-mediated regulation of RNPS1 expression.

Chapter 2:
Generation of polyclonal antibody specific for RNPS1

The work embodied in this chapter is published.

Deka B, Singh KK. Molecular cloning, expression and generation of a polyclonal antibody specific for RNPS1. 2022. Molecular Biology Reports. 9095–9100. https://doi.org/10.1007/s11033-022-07676-8

Abstract

RNA-binding protein with serine-rich domain 1 (RNPS1) is a member of a splicing-dependent mega Dalton protein complex, known as exon junction complex (EJC). During splicing, RNPS1 acts as a protector of global transcriptome integrity by suppressing the usage of cryptic splice sites. Additionally, RNPS1 functions in almost all stages of mRNA metabolism, including constitutive splicing, alternative splicing, translation and nonsense-mediated mRNA decay (NMD). The aim of this chapter was to generate a highly specific polyclonal antibody against human RNPS1 protein. The rabbit antiserum was obtained by immunizing rabbits with the purified recombinant RNPS1 protein (22-305 amino acids). The antiserum was then purified by antigen-immunoaffinity chromatography. The sensitivity and the specificity of the polyclonal antibody were assessed by enzyme-linked immunosorbent assay (ELISA) and Western blot. ELISA demonstrated that the antibody has a high binding affinity for RNPS1. The antibody was able to detect endogenous RNPS1 protein from mammalian cell lysates in Western blot. The antibody also efficiently detected the decrease in RNPS1 expression in siRNA induced knockdown assay, indicating the specificity of the antibody. The polyclonal antibody against RNPS1 will be a useful tool for performing further functional studies on RNPS1.

2.1 Introduction

Antibodies are one of the critical reagents in most of the research methodologies. These are typically characterized as either polyclonal or monoclonal. Polyclonal antibodies have binding affinity for several different antigen epitopes, while monoclonal antibodies have affinity for single antigen epitope. The majority of polyclonal antibodies are produced in laboratories using rabbits. The use of rabbits has advantages over other species. Rabbits have appropriate body size and accessible marginal ear vein and central auricular artery for withdrawing blood samples. Moreover, rabbits provide high reactivity to a range of antigens and have a single primary Immunoglobulin G isotype.

Deciphering the functions of RNPS1 in mRNA metabolism requires the use of various molecular and biochemical techniques. Many of the widely used techniques to study the function of a gene at the protein level are highly dependent on the availability of a sensitive and specific antibody against the protein. In this regard, most of the commercially available anti-RNPS1 antibodies show non-specific reactivity or low sensitivity. Additionally, non-commercial RNPS1 antibody is available in one of the laboratory but its accessibility is limited. In order to study the function and regulation of RNPS1 at the protein level, efficient anti-RNPS1 antibody was required and the non-availability of anti-RNPS1 antibody became the bottleneck for carrying out studies on RNPS1. Therefore, in this study, I report the production of a specific polyclonal anti-RNPS1 antibody. The generated polyclonal antibody was useful for the study of the expression of RNPS1 at the protein level, particularly via Western Blot analysis.

2.2 Materials and methods

2.2.1 Plasmid Construction

pGEX6p3-GST RNPS1 was a kind gift from Prof.Niels Gehring, University of Cologne, Germany. This plasmid was used as the template to PCR amplify truncated RNPS1 gene (22-305 amino acids) using the forward primer containing BamHI restriction site (5'-GCTGGATCCAGCACTAGGGCTCCTTCACCTACC-3') and the reverse primer containing XhoI

restriction site (5'-CTCTCGAGTTATCGGGAGGAGTTGGAGCTGGAGCG-3'). PCR was performed using Pfu polymerase (Biobharati, BB-E0020) under the following conditions: initial denaturation at 94°C for 5 min, followed by 30 cycles of 30 s at 94°C, 60 s at 58°C, 2 min at 72°C and the final extension at 72°C for 7 min. The purified PCR product of size 855bp was digested with BamH1 and XhoI restriction enzymes and ligated into digested pHis-TEV expression vector (Biobharati, BB-V0020) which contains N-terminal His-tag. The ligated plasmids were transformed into *E.coli* Dh5α (Novagen). Positive colonies were confirmed by restriction enzyme digestion and DNA sequencing.

2.2.2 Expression and purification of RNPS1 fusion protein

The recombinant plasmid encoding His-tag truncated RNPS1 protein (22-305 amino acids) was transformed into *E.coli* Rosetta (DE3). A single colony was inoculated in Luria Bertani (LB) medium supplemented with 100 µg/ml ampicillin and incubated overnight at 37°C. Overnight grown culture was then inoculated into fresh LB medium (1:100 dilution) containing ampicillin (100 µg/ml) and grown at 37°C until an OD_{600} of 0.6-0.8 was reached. The recombinant protein expression was then induced by adding IPTG at a final concentration of 0.5 mM IPTG and incubating the bacterial culture at 20°C for 16 hours.

The cells were harvested by centrifugation at 4000 g for 20 min at 4°C. Cell pellet was re-suspended in a lysis buffer containing 20mM Tris-Cl (pH 7.5), 150mM NaCl, 0.5mM EDTA, 0.1% Triton-X 100 and 0.1mM PMSF, followed by sonication. The cell lysate was centrifuged at 13000 rpm for 30 min at 4°C. The supernatant was then added to pre-equilibrated Ni-NTA agarose (Biobharati, Cat. No. BB-NA002) and incubated on a cyclic rotor for 2 h at 4°C. The column was then washed. First wash contained no imidazole whereas second wash contained 20 mM Imidazole to remove the loosely bound non-specific proteins. Finally, the bound proteins were eluted with elution buffer_1 (25mM Tris pH 7.5, 100mM NaCl, 100mM Imidazole), termed as elution 1. Subsequently, the protein of interest was eluted with elution buffer_2 (25mM Tris pH 7.5, 100mM NaCl,

250mM Imidazole), termed as elution 2. The eluted protein was analysed on SDS-PAGE. Elution 2 was used for immunization and antibody generation.

2.2.3 Production and purification of polyclonal antibody against RNPS1

Recombinant RNPS1 protein (250µg/rabbit, Bradford assay) was mixed with Freund's complete adjuvant in 1:1 ratio and injected subcutaneously into 175 and 190 days old New Zealand white rabbits (two batches). After 2 weeks of initial immunization, the rabbits were given 8 booster injections with protein (150 µg/rabbit) mixed with an equal volume of Freund's incomplete adjuvant at 2-week intervals. After last immunization, antiserum was collected after 7 days and its specificity was verified by ELISA. Antiserum was then affinity purified using 6% beaded agarose (Bio Bharati Life Science, BB-NA002) coupled to RNPS1 protein. 5 mL of antisera was applied to the 1 ml of affinity agarose bead. 100mM glycine (pH 2.8) was used to elute the purified antibody followed by neutralization with Tris-Cl (pH8.8). The final volume after concentrating the elute was 2mL (Concentration: 1.5 mg/mL calculated by Bradford Assay).

2.2.4 Antiserum titer determination by ELISA

Each well of the 96-well ELISA plate was coated with 200 pg/well of purified RNPS1 protein by overnight incubation at 4°C. Next day, the wells were washed with phosphate buffered saline containing 0.05% Tween-20 (Sodium chloride 0.137M, Potassium chloride 0.0027M, Sodium Phosphate Dibasic 0.01M and Potassium Phosphate Monobasic 0.0018M) and then blocked with 5% skimmed milk in PBST for 1 h at 37°C. The wells were then washed and incubated with 100 µL antiserum or purified antibody of different dilutions (1:500 to 1:40,000) for 2.5 h at 37°C. After washing with PBST, HRP-conjugated secondary antibody (1:5000 diluted in PBST containing 1% skimmed milk) was added to each well and incubated at 37°C for 1 h. The wells were then washed and 100 µL of TMB (3,3',5,5'- Tetramethylbenzidine) solution was added, followed by addition of an acidic stop solution. The absorbance was measured at 450 nm.

2.2.5 Cell culture

HeLa cell line was maintained in Dulbecco's Modified Eagle's Medium with high glucose (Himedia) supplemented with 10% (v/v) fetal bovine serum (Himedia) and 1% penicillin/streptomycin (Himedia) at 37°C in a humidified atmosphere containing 5% CO_2.

2.2.6 siRNA transfection

For siRNA transfection, cells were seeded in 6-well plates at 50% confluency and reverse transfected using Lipofectamine RNAi MAX (Life technologies). In brief, 60pmol siRNA and 3µl RNAiMAX were each diluted in 100 µl OptiMEM (Life technologies). Both dilutions were mixed, vortexed briefly, incubated for 15 minutes at RT and added to the cells. 24 hours after transfection, the medium was replaced and the cells were harvested after 72 hours. siRNA target sequence for RNPS1 are: RNPS1_1 5'-AAGGAAGACCAGTAGGAAA-3' and RNPS1_2 5'-GGATCGCTCAAAAGATAAA-3'. The luciferase siRNA: 5'-AACGUACGCGGAAUACUUCGATT-3' served as the negative control.

2.2.7 Western blotting

Cell lysates were prepared by sonication in ice-cold RIPA buffer. Total protein was estimated using Bradford reagent (Pierce) and resolved on SDS-PAGE. The proteins were transferred to 0.22 µm nitrocellulose membrane (Himedia) using the semi-dry system (BioRad) and followed by blocking in 5% skimmed milk in TBST for 1 hr at RT. The membrane was incubated with the respective primary antibody (Appendix II) at 4°C overnight. After washing with TBST, the membrane was incubated with HRP-conjugated secondary antibody anti-rabbit IgG (H+L) (CST, 7074S) or anti-mouse IgG (H+L) (CST, 7076S), 1:5000 diluted in TBST containing 1% skimmed milk for 1 h at room temperature. The antibodies were detected using ECL (Invitrogen) and visualized on the ChemiDoc XRS+ system (BioRad).

2.2.8 Mass spectrometry

Rosetta lysates were resuspended in SDS-gel loading buffer and loaded onto 10% SDS-PAGE. After running, the gel was stained with coomassie brilliant blue. The region corresponding to ~40-50 kDa was sliced out using a clean scalpel. The gel pieces were transferred to mass spec compatible tubes and supplemented with 300 µl deioinized water. The mass spectrometry analysis was outsourced to V proteomics, New Delhi, India. The excised band was used for in-gel trypsin digestion and analysed byNano ESI LC-MS/MS using the Orbitrap platform.

2.3 Results

2.3.1 Construction of Expression plasmid pHis-TEV-RNPS1

In order to generate expression construct pHis-TEV-RNPS1, the nucleotide sequence encoding truncated RNPS1 protein (22-305 amino acids) was amplified from the template vector pGEX6p3-GST RNPS1. Amplification by PCR generated a single DNA fragment of size 855 bp (Figure 2.1A). The resulting PCR fragment was then cloned between BamHI and XhoI sites of pHis-TEV vector in frame with N-terminal hexahistidine coding sequence present in the multiple cloning sites of the vector. The positive clone was identified by restriction enzyme analysis and DNA sequencing of full-length cloned PCR product (Figure 2.1B).

2.3.2 Expression and affinity purification of the recombinant protein

For heterologous expression of RNPS1 fusion protein, the expression plasmid pHis-TEV-RNPS1 was transformed into *E.coli* Rosetta (DE3) and induced with IPTG. The size of the expressed truncated RNPS1 protein (22-305 amino acids) is around 40 kDa.

After Ni-NTA affinity purification of His-tagged RNPS1 protein, the purity of protein samples at each step was analysed by SDS-PAGE. The result showed a single band of His-tagged RNPS1 protein (Figure 2.2A, lane 5).

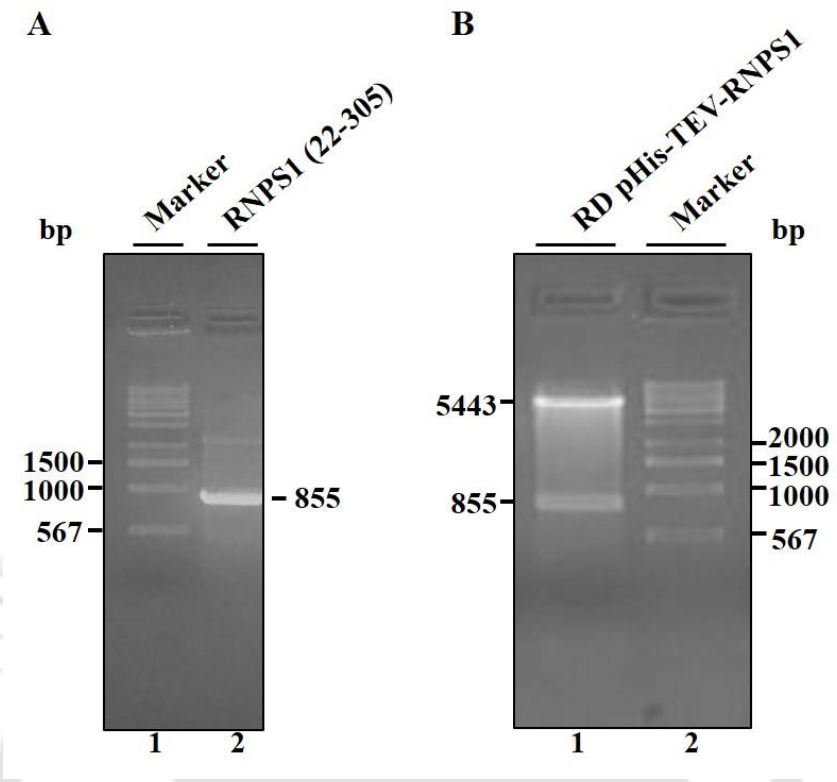


Figure 2.1: Cloning of expression plasmid pHis-TEV-RNPS1 for the generation of polyclonal antibody. (A) PCR amplification encoding truncated RNPS1 protein (22-305 amino acids) from the template vector pGEX6p3-GST RNPS1. Lane 1 contains molecular weight marker and lane 2 contains DNA fragment of size 855 bp encoding RNPS1 protein (22-305 amino acids) (B) Restriction enzyme digestion of recombinant RNPS1 gene in pHis-TEV-RNPS1 plasmid with BamHI and XhoI. Lane 1: double digested product of pHis-TEV-RNPS1 plasmid showed release of insert of size 855 bp encoding RNPS1 protein (22-305 amino acids) and Lane 2: molecular weight marker.

2.3.3 Titer analysis of polyclonal antibody by ELISA

The purified recombinant RNPS1 fusion protein (Figure 2.2A, lane 5, elution 2) was used as an antigen to immunize two rabbits (referred as two batches) and the antiserum against RNPS1 protein was prepared successfully. The titre of the raised RNPS1 antiserum was determined by ELISA. The antiserum at different dilutions (1:500 to 1:40,000) was reacted with an equal amount of purified RNPS1 protein (200 pg). Both the batches (2 rabbits) of RNPS1 antiserum showed a decrease of titre value with increase in dilution compared to the negative control, pre-immune sera (Figure 2.2B). In comparison with two batches, the Batch 1 showed more titre, thus batch 1 antiserum was used for further purifications.

RNPS1 antiserum was purified by affinity chromatography using 6% beaded agarose coupled to RNPS1 protein. The yield of purified anti-RNPS1 antibody was about 3 mg (2ml) from 5 ml of antiserum. The purified antibody was further subjected to quality checking by ELISA. The antibody at different dilutions (1:500 to 1:40,000) was reacted with an equal amount of purified RNPS1 protein (200 pg). ELISA assay reveals that the purified antibody has binding affinity for RNPS1 (Figure 2.2C). Altogether, the results indicate the reactivity of the anti-RNPS1 antibody.

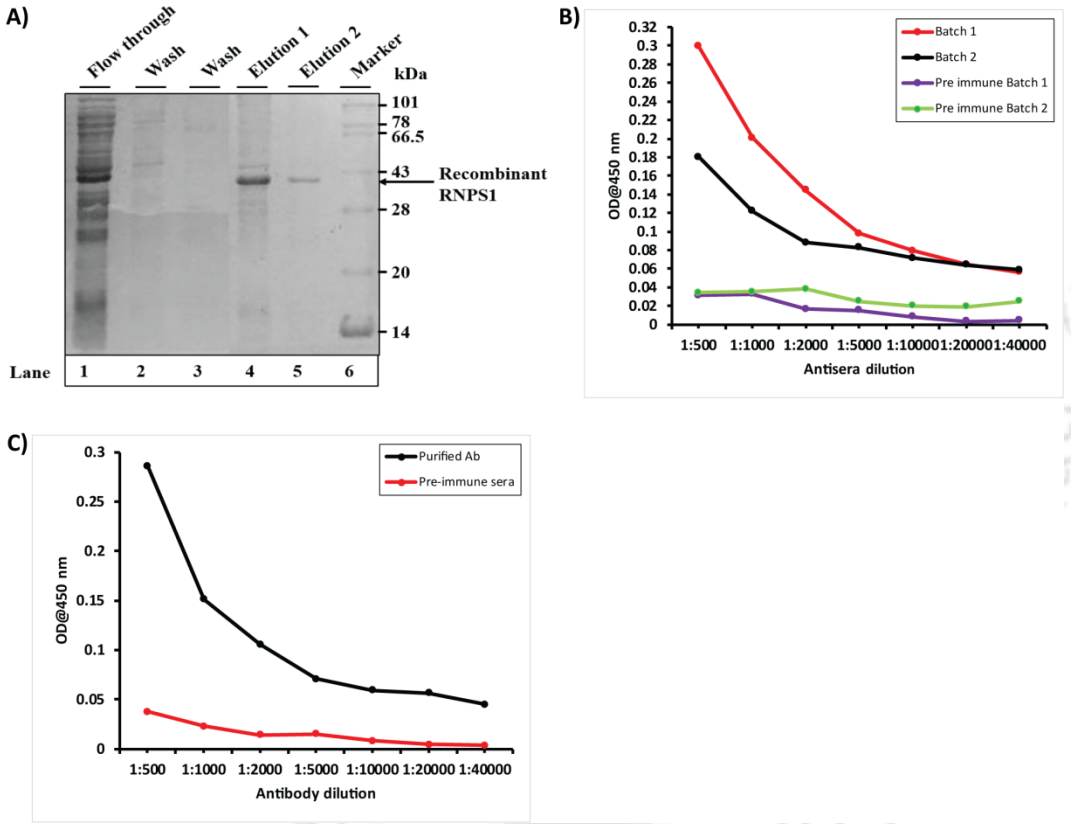


Figure 2.2: Expression and purification of the recombinant RNPS1 protein antigen and analysis of generated antiserum and antibody titre by ELISA. (A) Purification profile of recombinant RNPS1 by Ni-NTA affinity chromatography. Lane 1: Flow-through; Lane 2-3: Wash; Lane 4: Elution 1; Lane 5: Elution 2; Lane 6: Protein ladder. (B) The titres of antisera of two rabbits (Batch 1 and batch 2) were determined by ELISA. Pre-immune sera from respective rabbits served as the negative control. (C) The titre of purified antibody assayed by ELISA.

2.3.4 Specificity analysis of polyclonal antibody by Western Blot

To determine the specificity of the anti-RNPS1 antibody, Western Blot analysis was performed. Western Blot was carried out using a varying amount of purified RNPS1 protein (1-100ng). The results demonstrated that the polyclonal purified antibody could detect up to 1 ng of recombinant RNPS1 protein expressed in *E.coli* Rosetta (DE3) (Figure 2.3A). The purified RNPS1 protein that was used an immunogen might contain contaminating *E.coli* protein. Therefore, to rule out the possibility that the anti-RNPS1 antibody recognizes contaminating *E.coli* immunogenic protein, I assessed total protein lysate from *E.coli* in which RNPS1 was induced (pHis_TEV_RNPS1), total protein lysate from *E.coli* that lacks RNPS1 gene (negative control, empty parent vector, pHis-TEV) and total protein lysate from untransformed *E.coli* by immunoblotting. The result showed that the anti-RNPS1 antibody detects RNPS1 in lysate expressing RNPS1 and the band is absent in other two lysates that lacks RNPS1 gene (Figure 2.3B, lane 3), suggesting that the anti-RNPS1 antibody efficiently detects RNPS1. However, the antibody also detected an additional non-specific protein of *E.coli* in all three lysates (Figure 2.3B, lane 1-3). To identify the probable non-specific *E.coli* proteins, I performed mass spectrometry analysis. Proteins such as Rho or Enolase are probable candidate *E.coli* proteins that non-specifically bind to anti-RNPS1 antibody (Appendix III).

Additionally, Western Blot analysis was performed in cell lysates derived from rat liver tissue and RAW 264.7 (mouse) cell line. It was observed that the anti-RNPS1 purified antibody could detect endogenous RNPS1 protein from crude cell extracts (Figure 2.3C). Taken together, the results confirmed that the anti-RNPS1 antibody could detect a single prominent RNPS1 band on Western Blot analysis of both purified protein and total proteins from cell lines.

2.3.5 Polyclonal antibody validation by knockdown assay

siRNA mediated knockdown assay is one of the crucial validation strategies to ensure the absence of nonspecific binding of polyclonal antibody. Many nonspecific antibodies bind off-target proteins whose molecular weight is in close proximity to the target

protein leading to errors in research. Therefore, to verify the specificity of the anti-RNPS1 antibody, RNPS1 was depleted from HeLa cells using two siRNAs. A siRNA targeting luciferase (Luc) transcript served as a negative control. Western blot analysis

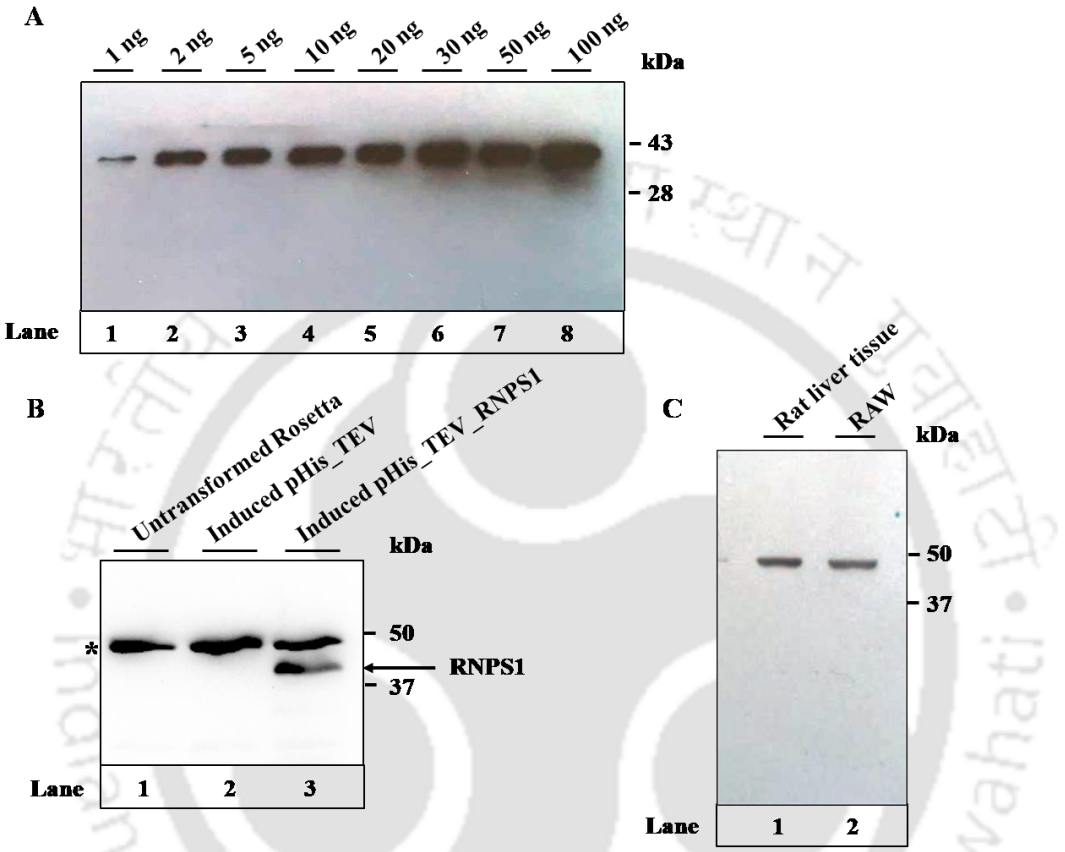


Figure 2.3: Western Blot analysis of the anti-RNPS1 purified antibody. (A) Immunoblotting of purified RNPS1 protein. Lanes 1-8: 1, 2, 5, 10, 20, 30, 50 and 100ng of purified RNPS1 protein, respectively. (B) Immunoblotting of Rosetta lysate. Lane 1: Total protein from Rosetta, Lane 2: Total protein from Rosetta containing empty pHis_TEV parent vector after induction, Lane 3: Total protein from Rosetta containing pHis_TEV_RNPS1 after induction. Bands corresponding to non-specific Rosetta proteins are labelled with an asterisk (C) Immunoblotting of total cell lysate using purified anti-RNPS1 antibody. Lanes 1-2: Total protein from rat liver tissue and RAW cell line, respectively.

revealed that the decrease in RNPS1 protein abundance in RNPS1-depleted cells was efficiently detected by anti-RNPS1 antibody compared to the strong signal from negative control cells (Figure 2.4). Comparing lane 1 with lane 4 (Figure 2.4), in which 100% lysate was loaded in both lanes, I observed the level of RNPS1 protein in RNPS1-depleted cells was less than 50% of that in Luc siRNA (negative control) transfected cells.

Together, the data demonstrate that the generated polyclonal antibody anti-RNPS1 is specific and is useful for ELISA and Western Blot analysis.

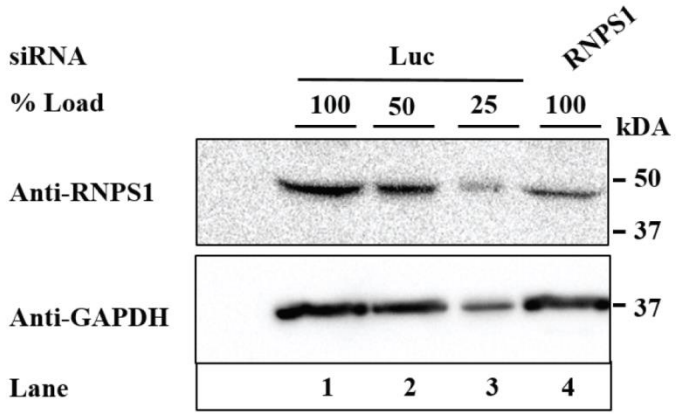


Figure 2.4: Specific knockdown of RNPS1 protein. HeLa cells were transfected with siRNAs targeting RNPS1 (siRNAs: RNPS1_1 and RNPS1_2). Expression levels of RNPS1 protein was determined by Western Blot and GAPDH as loading control. The knockdown efficiency was assessed with anti-RNPS1 antibody and compared with Luc-siRNA transfected cells (negative control). Dilution corresponding to 100% refer to the initial protein amount, 50µg, (lane 1 or lane 4) of negative control or siRNA RNPS1 sample. Serial dilutions corresponding to 50% or 25% (lane 2-3) of the initial protein amount (lane 1) of negative controlwere loaded to assess the efficiency of the RNPS1 depletion. The level of RNPS1 protein in cells treated with siRNA against RNPS1 (lane 4) is less than 50% of that in negative control cells.

2.4 Discussion

Considering the vital functions of RNPS1 in multiple post-transcriptional processes, it is clear that the antibody against RNPS1 is as an indispensable tool for studying in depth the regulatory roles of RNPS1. In the present study, I report the efficient production of a specific anti-RNPS1 antibody. In this regard, RNPS1 gene (22-305 amino acids) was successfully cloned and efficiently expressed as a His-tagged RNPS1 fusion protein in *E.coli.* Furthermore, the fusion protein was efficiently purified using Ni-affinity purification technique. The fusion protein acted as an immunogen to raise high-titer antiserum against RNPS1. The antiserum was then purified by antigen-immunoaffinity chromatography, a highly selective purification method. ELISA assay demonstrated the reactivity of the purified polyclonal anti-RNPS1 antibody. Moreover, the antibody could specifically recognize both recombinant RNPS1 protein and endogenous RNPS1 protein, indicating the efficacy of the antibody. The specificity of the antibody was critically

evaluated in a siRNA-induced knockdown assay, wherein anti-RNPS1 could efficiently detect the decrease in RNPS1 protein expression in Western blot.

The probable reason behind successful generation of anti-RNPS1 antibody via this approach might be attributed to the protein immunogen (truncated RNPS1 protein) that was used instead of peptide antigen. Most of the commercial antibodies used peptide sequence as an immunogen, which might not be suitable for generating anti-RNPS1 antibody. Since the antibody generated using protein immunogen recognizes multiple epitopes, therefore in line with this, I have used appropriate negative control in our target assay.

Additionally, antibodies are also required for applications such as immunoprecipitation (IP) and immunofluorescence (IF) in which the antibody needs to recognize the target protein in its native form. Furthermore, probing the immunoprecipitated samples on Western blot depends on the concentration of the primary antibody. Based on our experience, the anti-RNPS1 antibody needs to be further concentrated, which might extend its potential applications in IP and IF. Nevertheless, the developed antibody would facilitate investigating the functional roles of RNPS1 via Western blot assay.

Chapter 3: Functional roles of RNPS1 in cervical cancer cells

The work embodied in this chapter is published.

Deka B, Chandra P, Yadav P, Rehman A, Kumari S, Kunnumakkara AB, Singh KK. RNPS1 functions as an oncogenic splicing factor in cervical cancer cells. IUBMB Life. 2022; doi: 10.1002/iub.2686.

Abstract

Numerous recent studies suggest that cancer-specific splicing alteration is a critical contributor to the pathogenesis of cancer. RNA-binding protein with serine-rich domain 1, RNPS1, is an essential regulator of the splicing process. However, the defined role of RNPS1 in tumorigenesis still remains elusive. I report here that the expression of RNPS1 is higher in cervical carcinoma samples (TCGA-CESC) compared to the normal tissues. Consistently, the expression of *RNPS1* was high in cervical cancer cells compared to a normal cell line. This study shows for the first time that RNPS1 promotes cell proliferation and colony-forming ability of cervical cancer cells. Importantly, RNPS1 positively regulates migration-invasion of cervical cancer cells. Intriguingly, depletion of RNPS1 increases the chemosensitivity against the chemotherapeutic drug doxorubicin in cervical cancer cells. Further, I characterized the genome-wide isoform switching stimulated by RNPS1 in cervical cancer cells. Mechanistically, RNA-sequencing analysis showed that RNPS1 regulates the generation of tumor-associated isoforms of key genes, particularly *Rac1b*, *RhoA, MDM4*, and *WDR1,* through alternative splicing. RNPS1 regulates the splicing of *Rac1* pre-mRNA via a specific alternative splicing switch and promotes the formation of its tumorigenic splice variant, *Rac1b*. While the transcriptional regulation of RhoA has been well studied, the role of alternative splicing in RhoA upregulation in cancer cells is largely unknown. Here, I have shown that knockdown of RNPS1 in cervical cancer cells leads to the skipping of exons encoding the RAS domain of RhoA, consequently causing decreased expression of RhoA. Collectively, I conclude that gain of RNPS1 expression may be associated with tumor progression in cervical carcinoma. RNPS1 mediated alternative splicing favors an active Rac1b/RhoA signaling axis that could contribute to cervical cancer cell invasion and metastasis. Thus, this thesis work unveils a novel role of RNPS1 in the development of cervical cancer.

3.1 Introduction

Cancer develops primarily as a result of sequential genetic changes and genomic instability, causing constitutive expression of oncogenes and repression of tumor-suppressor genes. As a result, cancer cells gain unique capabilities such as increased replicative potential, insensitivity to growth-inhibitory signals, delayed apoptosis, persistent angiogenesis, and tissue invasion.

Gynecological cancers are the second most diagnosed cancer in women, following breast cancer. Cervical cancer is the fourth most frequent cancer globally and the most common malignancy in developing countries [139]. Cervical cancer is linked to a number of risk factors, including sexually transmitted infections such as HPV (human papillomavirus) and HIV (human immunodeficiency virus), smoking, alcohol consumption, prolonged use of oral contraceptives, and a family history of cervical cancer [140]. Long-term infection with HPV is the major risk factor for cervical cancer [141]. HPV infections trigger a variety of cellular alterations, including alternative splicing, which leads to malignant transformation. Aberrant splicing events in cancer-related genes also result in chemo- and radioresistance [142].

Interestingly, a recent systematic analysis by Kahles et al. revealed that alternative splicing events are more common in cancer tissues than normal tissues. Alterations in splicing factor expression levels appear to be the main drivers of abnormal splicing profiles [143]. The key trans-acting regulatory factors in splicing- SR proteins and the hnRNPs are altered in cervical cancer. As a representative example, SR protein, SRSF10, controls the splicing of the *IL1RAP* gene and generates the oncogenic splice isoform, MIL1RAP, in cervical cancer. As a result, cancerous cells can avoid macrophage phagocytosis and evade the immune system [144]. The FAS receptor gene is another example that undergoes alternative splicing in cancer cells. The FAS ligand secreted by cytotoxic T cells activates the FAS receptor, which results in apoptosis via a death-signaling cascade. FAS receptor has three short mRNA variants that lack the transmembrane domain. As a result, the protein isoforms translated from these variants

are probably released by the cancer cells as decoy receptors for the FAS ligand, thus aiding cancer cells in avoiding apoptosis [145, 146].

Alternative splicing also plays a critical role in promoting the invasion or migration capability of cancer cells. Cell migration is a highly orchestrated multistage process involving changes in the cytoskeleton, cell-substrate adhesions and remodeling of the extracellular matrix. The cancer cells acquire a pro-migratory phenotype to invade the stroma, migrate towards blood vessels and enter the bloodstream. Rho GTPases are one of the vital regulators of cytoskeletal dynamics, thereby playing pivotal roles in cell migration event. Most Rho GTPases remain inactive when bound to GDP and get activated by exchanging the bound GDP nucleotide for GTP. The most highly conserved Rho family members across eukaryotic species are Rho, Rac, and Cdc42. Briefly, Rac and Rho regulate the formation of lamellipodia, whereas Cdc42 promotes the development of filopodia and the coordinated action helps the cell in directional migration [147]. Intriguingly, *Rac1*, *RhoA* and *Cdc42* undergo alternative splicing to alter the expression and/or functions of Rho GTPases. The *Rac1* pre-mRNA produces two splice variants, *Rac1* and *Rac1b*. The splice variant *Rac1b*, produced from the inclusion of an alternative exon 3b, encodes a constitutively active GTPase protein [148, 149]. Rac1b is considered a pro-tumorigenic GTPase that promotes cellular transformation [150]. Likewise, alternative splicing of *Cdc42*, resulting in the switching of the *CDC42-v2* variant to the *CDC42-v1* variant, is associated with malignant transformation [151].

Similar to SR protein, RNPS1 also functions in constitutive and alternative splicing. It is also a critical effector molecule in promoting splicing fidelity. However, the role of RNPS1 in tumorigenesis remains elusive.

In this chapter, I have checked whether RNPS1 contributes to the development of cervical cancer. Our findings revealed for the first time that RNPS1 plays an important role in regulating key hallmarks of cancer, such as survival, migration and invasion.

Interestingly, I found that RNPS1 modulates the alternative splicing of cancer-specific genes to generate oncogenic isoforms.

1.2 Materials and methods

3.2.1 Cell culture

HeLa, SiHa and HDF cell lines were maintained in Dulbecco's Modified Eagle's Medium with high glucose (Himedia) supplemented with 10% (v/v) fetal bovine serum (Himedia) and 1% penicillin/streptomycin (Himedia) at 37°C in a humidified atmosphere containing 5% CO_2.

3.2.2 siRNA transfection

Please refer to Chapter 2, section 2.2.6.

3.2.3 RNA isolation

Total cellular RNA was extracted with TRIzol reagent (Invitrogen) and chloroform separation. In brief, growth medium is removed from cells in a 6-well plate. 500 µl TRIzol was added, the cells were lysed by pipetting and the suspension was transferred to a 1.5 ml reaction tube. Phase separation was induced by adding 100 µl chloroform, followed by vortexing and 10 min incubation at RT. The sample was centrifuged for 15 minutes at 12,000 × g at 4°C. After centrifugation, the upper aqueous phase was transferred to a new 1.5 ml reaction tube. 250 µl isopropanol was added, mixed by vortexing and incubated for 10 min at RT. Subsequently, the RNA was precipitated by centrifuging for 10 minutes at 12,000 × g at 4°C. The supernatant was removed and the pellet was washed once with 500 µl of 75% ethanol, followed by centrifuging for 10 min at 12,000 × g at 4°C. After the last washing step, the ethanol was removed completely and the RNA pellet was air-dried for a few minutes. The pellet was resuspended in nuclease-free H2O.

3.2.4 Removal of genomic DNA

In order to remove genomic DNA contamination from isolated RNA, total RNA was incubated with DNase I (Promega). 2 units of DNase was added to 2 ug of total RNA and incubated at 37°C for 30 mins. Then, DNase was inactivated with 1ul of DNase stop solution and incubated at 65°C for 10 mins.

3.2.5 cDNA synthesis

cDNA was prepared by reverse transcribing DNase treated total RNA using the high capacity cDNA reverse transcription kit (Applied Biosystems). 2ug of total RNA was mixed with 0.8 ul dNTPs, 2 ul random primers, 2 ul reaction buffer, and 1 ul MultiScribe Reverse Transcriptase in a total volume of 20 ul. The mixture was incubated at 25°C for 10 mins, followed by 37°C for 120 mins. Afterward, the enzyme was heat-inactivated at 85°C for 5 mins.

3.2.6 End-point PCR

1ul of diluted cDNA (1:3 dilution of the cDNA) was used as templates in end-point PCR reactions with 0.15 µM gene-specific primers (Appendix I) and EmeraldAmp MAX HS PCR Master Mix (Taqara). The PCR products were analyzed on 2% agarose gels and the images of ethidium bromide-stained bands were visualized using the ChemiDoc XRS+ system (BioRad).

3.2.7 Quantitative real-time PCR

qRT-PCR was performed using PowerUp Sybr green master mix (Invitrogen). DNase I treated cDNA was used as a template for the qPCR. For each reaction, 5 ul of 2X Sybr green master mix was mixed with 0.25 uM of each primer, 1 ul of diluted cDNA (1:3 dilution of the cDNA) and the rest were filled with nuclease-free water. Specific primers (Appendix I) were used for quantifying gene expression and normalized with β-actin expression. Relative gene expression was calculated using the $\Delta\Delta C_T$ method. A melting

curve was determined after the last amplification step to ensure that the amplification products represent a single specific product.

3.2.8 Western blotting

Please refer to Chapter 2, section 2.2.7.

3.2.9 Wound healing assay

Control siRNA and RNPS1 knockdown cells were seeded in 6-well plates and allowed to grow until confluency and then serum-starved for 12 h. Confluent monolayers were scratched with a 200 ul pipette tip. Plates were washed with PBS to remove non-adherent cells, and fresh medium containing 2% FBS was added. The wound was photographed at regular intervals at 5X magnification. The extent of wound closure was quantified using TScratch software. The percentage of wound closure was calculated using the formula $[(A_0 – A_t)/ A_0] X 100$, where A_0 is the wound area at 0 hr and A_t is the area of the wound at t hr.

3.2.10 Transwell migration and invasion assay

For transwell migration, 24-well cell culture inserts (Nunc, Thermo Fischer Scientific) were used, and cells were seeded ($0.8 X 10^5$ cells per insert) on the top chamber of the inserts with the incomplete media. Cells were allowed to migrate through the insert's membrane to the lower side containing complete media. After 24 h, inserts were removed, and cells migrated on the bottom layer of the inserts were fixed with 100% methanol and stained using 0.1% crystal violet. The upper side of the membrane was wiped with cotton swabs to remove non-migrated cells. Images of the migrated cells were taken using a microscope. For quantification, 10% acetic acid solution was used to destain the insert membrane, and absorbance measurements were taken at 595 nm. Invasion assays were performed as described above, but the cells were allowed to migrate through a membrane insert pre-coated with matrigel (Corning).

3.2.11 Colony formation assay

Control siRNA and RNPS1 knockdown cells were seeded in a 6-well plate at a density of 1 X 10^3 cells/well. The cells were grown for seven days, and then colonies were fixed with 100% methanol and stained with crystal violet. Pictures of individual wells were taken and were analyzed using ImageJ software, and the surviving fraction was calculated.

Plating efficiency = no. of colonies counted / no. of cells plated

Surviving fraction = (no. of colonies counted /no. of cells plated) / plating efficiency of control

3.2.12 Cell cycle analysis

Cells were washed with PBS, trypsinized and transferred to a 1.5 ml reaction tube followed by centrifugation. Cell pellets were washed with PBS and resuspended in 150 µl PBS. Fixation and permeabilization of the cells were done with 1 ml ice-cold 70% ethanol at -20°C overnight. The next day, cells were washed with PBS followed by centrifugation at 700 x g for 15 mins. Cell pellets were treated with 150 µl RNase solution (stock: 200 µg/ml) at 37°C for 30 mins. The cells were then stained with 150 µl propidium iodide (stock: 60 µg/ml) to stain the DNA and analyzed by flow cytometry. The data were analyzed using FCS Express 5 software.

3.2.13 Flow cytometric analysis of apoptotic cell death

Control siRNA and RNPS1 knockdown cells were treated with 5µM of Doxorubicin for 48hrs. Dead cells in the media were harvested and adherent cells were trypsinized. Cell pellets were then washed with PBS and resuspended in binding buffer (Invitrogen). The cells were then stained with Annexin V/PI (Invitrogen) for 30 min in the dark at 4°C and analyzed by flow cytometry.

3.2.14 RNA sequencing and computational analysis

Available RNA-Seq datasets of RNPS1 knockdown in HeLa cells were used for RNA-Seq data analysis (E-MTAB-6564). Adaptor contamination was removed from raw reads using cutadapt (v3.3). The reads were aligned against the human genome (version 38, GENCODE release) and Salmon was used to compute estimates for transcript abundance using the -validateMappings parameters. Differential transcript usage was computed with IsoformSwitchAnalyzeR and the DEXSeq method. Significance thresholds were delta isoform fraction |dIF| > 0.1 and adjusted p-value (isoform_switch_q_value) < 0.05.

3.2.15 Statistical analysis

All the quantitative data are represented as mean ± SD. Agarose gel images and chemiluminescence images were quantified by ImageJ software. In all the statistical tests, $p < 0.05$ was considered as significant. Statistical significances were calculated with the Student's t-test using the GraphPad PRISM.

3.3 Results

3.3.1 Effect of RNPS1 knockdown on proliferation and clonogenic potential of cervical cancer cells

To determine whether RNPS1 has a role in cervical cancer tumorigenesis and progression, I first analyzed the expression of RNPS1 in TCGA samples using the UALCAN web resource. The data revealed that RNPS1 is overexpressed in Cervical Squamous Cell Carcinoma and Endocervical Adenocarcinoma (TCGA-CESC) samples compared to normal tissue (Figure 3.1A). I then investigated the expression of RNPS1 in HeLa, SiHa, and human dermal fibroblast (HDF) cell lines. Consistent with the TCGA data, *RNPS1* was highly expressed in cervical cancer cells compared to the normal cell line, HDF (Figure 3.1B).

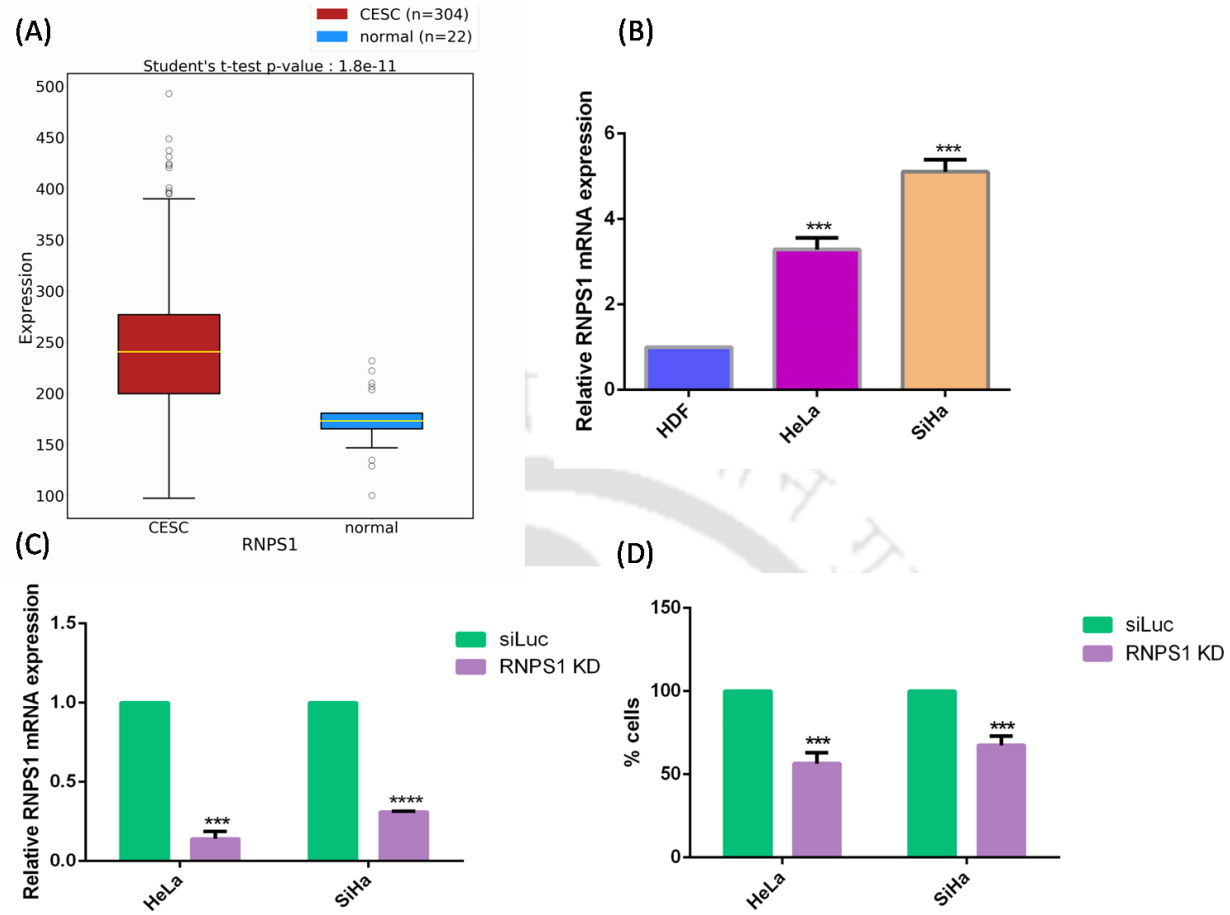


Figure 3.1: Expression levels of RNPS1 in cervical cancer cell lines and tissues and effect of its knockdown on cervical cancer cell proliferation. (A) Expression of RNPS1 in normal vs primary cervical cancer tissue from TCGA database using OncoDB web resource. (B) qRT-PCR shows the relative expression of *RNPS1* in human cervical cancer cell lines (HeLa and SiHa) and normal cell line, HDF. β-actin was used as a normalization control. (C) The expression level of *RNPS1* mRNA upon knockdown of RNPS1 by siRNA in HeLa and SiHa by qRT-PCR. (D) Estimation of cell proliferation by Trypan blue staining in control and RNPS1 knockdown cervical cancer cells after 72hrs of siRNA treatment. Values are depicted as mean ± S.D (n=3) and p-values are depicted as ***p ≤ 0.001, ****p ≤ 0.0001.

Since RNPS1 is significantly upregulated in cervical cancer, I hypothesized that RNPS1 might function as an oncogene. Therefore, to understand the role of RNPS1 in cervical cancer, the expression of RNPS1 was silenced in cervical cancer cell lines, HeLa and SiHa. RNPS1 was knocked down efficiently using two specific siRNAs targeting different regions of *RNPS1* mRNA (Figure 3.1C). Next, the effect of RNPS1 knockdown on the proliferation of cervical cancer cells was examined and observed a decrease in the cell numbers of HeLa and SiHa cells (Figure 3.1D). Together, these suggest that RNPS1 is involved in the positive regulation of the proliferation of cervical cancer cells.

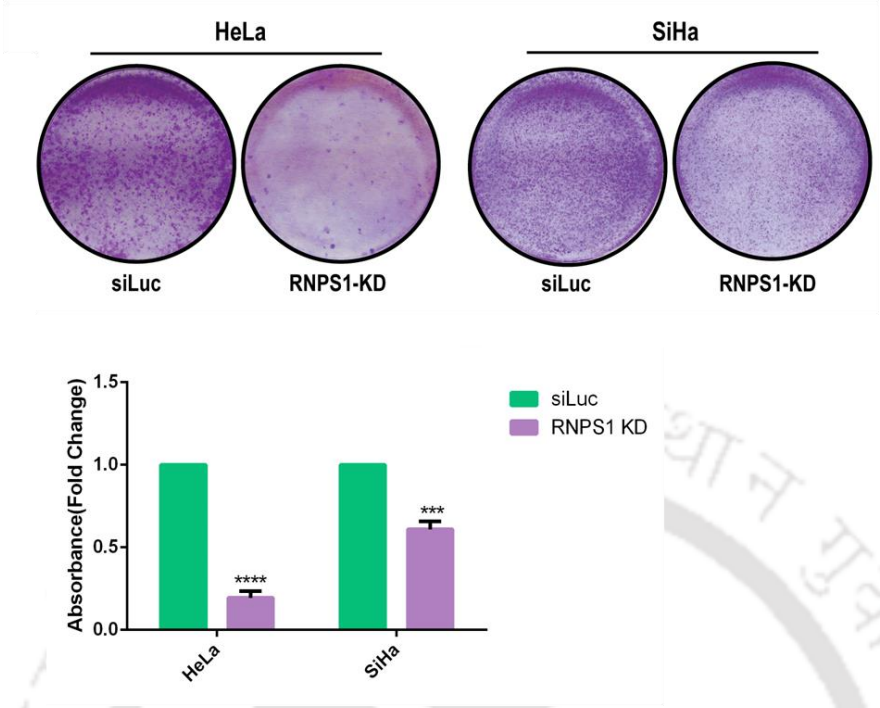


Figure 3.2: Effect of RNPS1 knockdown on the clonogenic potential of cervical cancer cells, HeLa and SiHa using colony formation assay after 7 days of siRNA treatment. The graphical representation of estimation of colony forming ability of RNPS1 knockdown cervical cancer cells. The average absorbance was represented as fold change. Absorbance from control cells were normalized to 1, whereas clonogenic potential in RNPS1 depleted cells was measured as a fold change. Values are depicted as mean ± S.D (n=3) and p-values are depicted as ***p ≤ 0.001, ****p ≤ 0.0001.

The effect of RNPS1 knockdown on the survival of cervical cancer cells was determined using a colony formation assay. This assay basically assesses the ability of a cell to undergo “unlimited” division. The knockdown of RNPS1 decreases the number of colonies compared to the control (Figure 3.2). This indicates that RNPS1 promotes the clonogenic ability of HeLa and SiHa cells. Thus, RNPS1 provides a survival advantage to cervical cancer cells.

3.3.2 Downregulation of RNPS1 alters cell cycle progression of cervical cancer cells

Next, to gain deeper insight into the observed altered proliferation of RNPS1-KD cells, I evaluated the effect of loss of RNPS1 on cell cycle progression by flow cytometry. The cell cycle analysis showed that HeLa cells were distributed more in G2/M-phase and less in the S phase after depletion of RNPS1, which suggests that RNPS1-KD HeLa cells were

arrested at the G2/M phase (Figure 3.3A). Similarly, SiHa cells were distributed less in the S-phase after the knockdown of RNPS1 (Figure 3.3B). Moreover, qPCR analyses of M phase-related genes showed a significant reduction in the mRNA expression of *ANAPC5* and *ANAPC7*. In addition, RNPS1 knockdown enhanced the mRNA level of another essential cell cycle gene, *CDKN2B* (Figure 3.3C-D). Taken together, these results suggest that RNPS1 is an important regulator of the cell cycle.

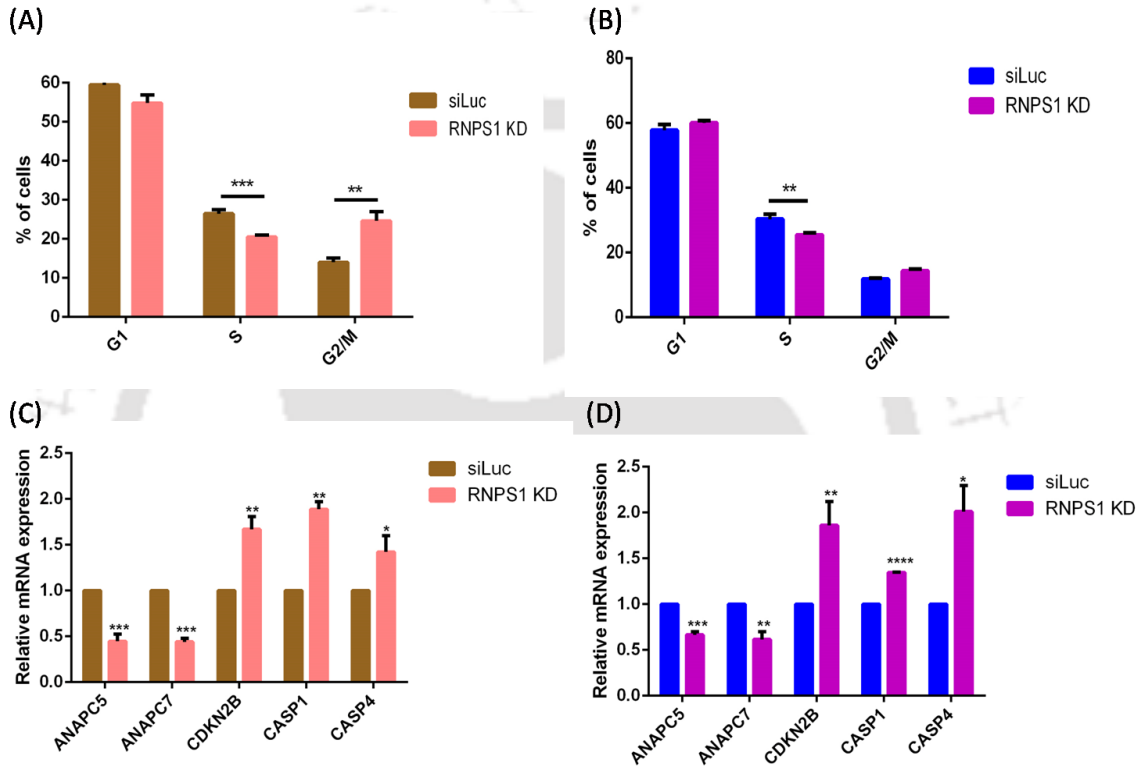


Figure 3.3: Effect of RNPS1 knockdown on cell cycle progression. (A-B) Cell cycle analysis was performed using flow cytometry in (A) HeLa and (B) SiHa cells transfected with siRNA-RNPS1 after 72 hrs of siRNA treatment and its graphical representation in terms of % distribution of cells. (C-D) qRT-PCR of genes related to cell cycle and caspase family in RNPS1 knockdown (C) HeLa and (D) SiHa cells normalized to β-actin. Values are depicted as mean ± S.D (n=3) and p-values are depicted as *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001, ****p ≤ 0.0001.

In cancer, the apoptotic pathway is typically inhibited and faulty regulation of apoptosis is an important event in the development of cancer. Therefore, the effect of RNPS1 depletion on the apoptosis of HeLa and SiHa cells was assessed. Flow cytometry analysis with Annexin V and PI staining revealed that the apoptotic rates of knockdown HeLa cells were higher than those of the control group cells (Figure 3.4A). In line with

previous reports, the data suggest that RNPS1 regulates apoptosis in HeLa cells [62, 152]. On the contrary, depletion of RNPS1 in SiHa cells did not trigger apoptosis (Figure 3.4B). Hence, RNPS1 probably regulates apoptosis in a cancer cell-specific manner. Nevertheless, the levels of Caspase family members, particularly *CASP1* (Caspase 1) (Figure 3.3C-D) and CASP4 (Caspase 4) (Figure 3.12C-E), were increased upon knockdown of RNPS1 in both HeLa and SiHa cells.

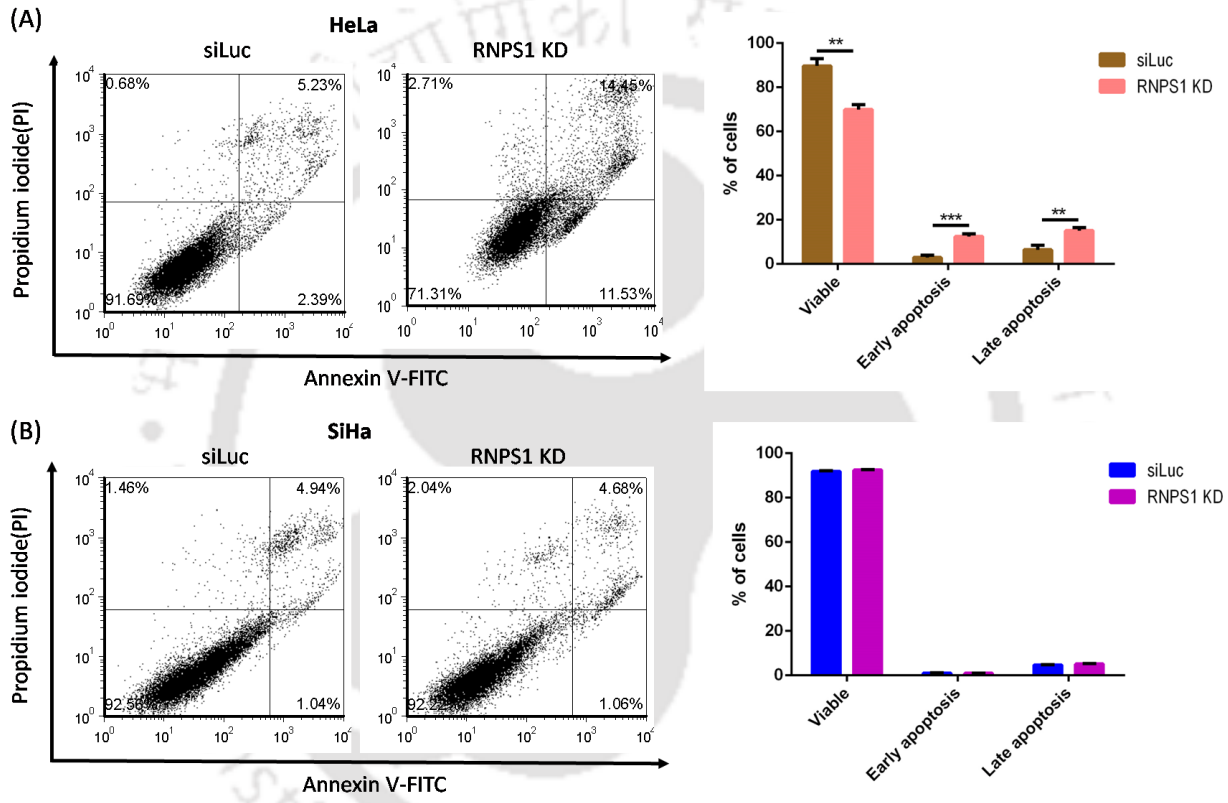


Figure 3.4: Effect of RNPS1 knockdown on apoptosis of cervical cancer cells. (A) Apoptosis rate was quantified using flow cytometry in HeLa cells transfected with the indicated siRNA after 72hrs of siRNA treatment and its graphical representation in terms of % distribution of cells (right panel) (B) Apoptosis rate was quantified using flow cytometry in RNPS1 knockdown SiHa cells and its graphical representation in terms of % distribution of cells (right panel). Values are depicted as mean ± S.D (n=3) and p-values are depicted as **p ≤ 0.01, ***p ≤ 0.001.

3.3.3 Silencing of RNPS1 decreases migration and invasion of cervical cancer cells

Migration is a critical event during the malignant transformation of cancer cells. To assess the functional effect of RNPS1 knockdown on the migration of cervical cancer

cells, HeLa and SiHa, wound healing assay and transwell migration assay was performed. In wound healing assay, I found that migration was inhibited in cervical cancer cells upon RNPS1 knockdown (Figure 3.5). The migration potential of RNPS1-KD cervical cancer cells was further assessed using serum as a chemoattractant by transwell migration assay. Crystal violet staining of the membrane showed that the number of migrated cells in RNPS1-KD cells was less compared to the control cells (Figure 3.6A-B). This suggests that RNPS1-KD cells have lower migration potential compared to the control cell.

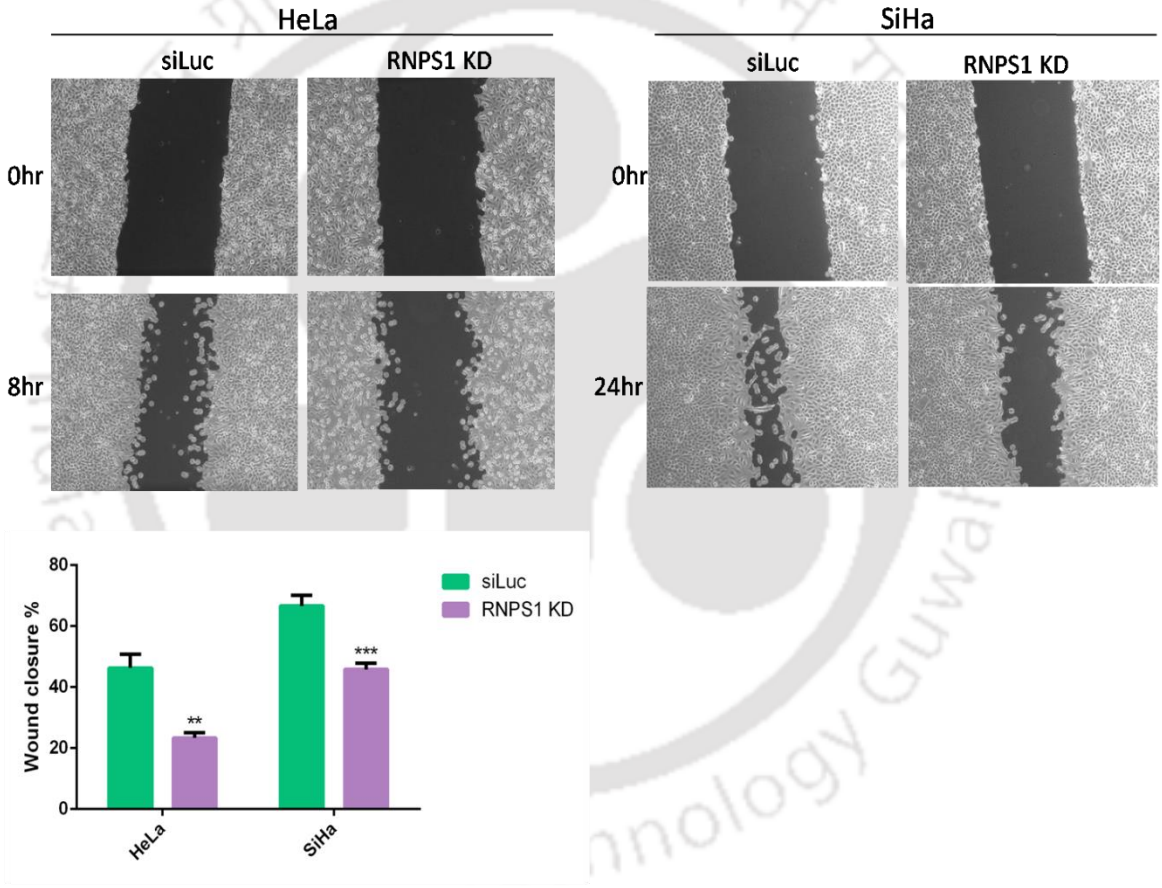


Figure 3.5: Effect of RNPS1 knockdown on migration potential of cervical cancer cells. Wound healing assay to determine the migration of HeLa and SiHa cells upon knockdown of RNPS1 (after 72 hrs of siRNA treatment). Graphical representation of the % of wound closure of RNPS1 knockdown cervical cancer cells in comparison to control cells. Values are depicted as mean ± S.D (n=3) and p-values are depicted as **p ≤ 0.01, ***p ≤ 0.001.

Besides migration, invasion of neighboring tissues is another essential hallmark of metastatic cancer cells. In this regard, RNPS1 knockdown cells exhibited markedly

decreased invasive ability than the negative control cells. The number of cells that invaded the lower part of the matrigel coated transwell insert was substantially reduced in RNPS1 KD cells compared to the control cells (Figure 3.6A, C). Hence, our data indicate that RNPS1 facilitates the invasion of cervical cancer cells.

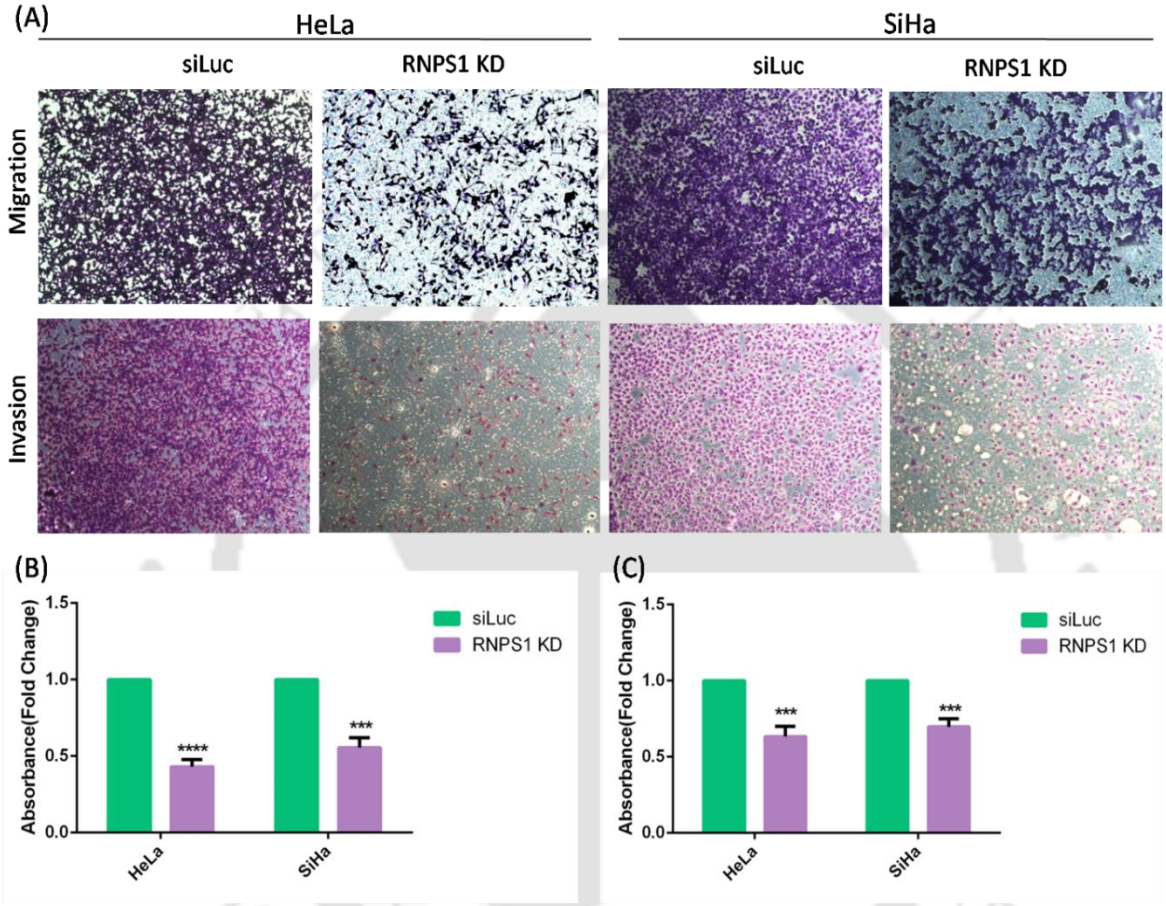


Figure 3.6: Effect of RNPS1 knockdown on migration and invasion potential of cervical cancer cells. (A) Effect of RNPS1 knockdown (after 72 hrs of siRNA treatment) on migration and invasion of cervical cancer cells using Boyden chamber assay. (B) Graphical representation of the migration potential of RNPS1 knockdown cervical cancer cells in comparison to control cells. The average absorbance was represented as fold change. Absorbance from control cells were normalized to 1, whereas migration in RNPS1 depleted cells was measured as a fold change. (C) Graphical representation of the invasive potential of RNPS1 knockdown cervical cancer cells in comparison to control cells. Values are depicted as mean ± S.D (n=3) and p-values are depicted as ***$p \leq 0.001$, ****$p \leq 0.0001$.

Since RNPS1 influences the migration and invasion potential of cervical cancer cells, I next investigated the effect of RNPS1 knockdown on the expression of genes related to migration and invasion. Interestingly, I observed a significant decrease in the expression of genes such as *CTSV* (Cathepsin V) (Figure 3.7) and N-Cadherin protein (Figure 3.12C-E)

upon knockdown of RNPS1. The knockdown of RNPS1 also decreased the cleaved form of MMP9 (matrix metalloproteinases). MMPs and cathepsins are proteolytic enzymes that degrade the extracellular matrix (ECM) and help the cancer cell migrate and invade secondary sites [153]. However, no change in the mRNA level of *TIMP3* (tissue inhibitors of matrix metalloproteinases), an inhibitor of MMPs was detected (Figure 3.7). These data indicate that RNPS1 might facilitate cell migration and invasion in cervical cancer cells. In summary, these results demonstrate that RNPS1 promotes cell proliferation, colony-forming ability, migration and invasion potential, thereby supporting the notion that RNPS1 exerts an oncogenic role in human cervical cancer cells.

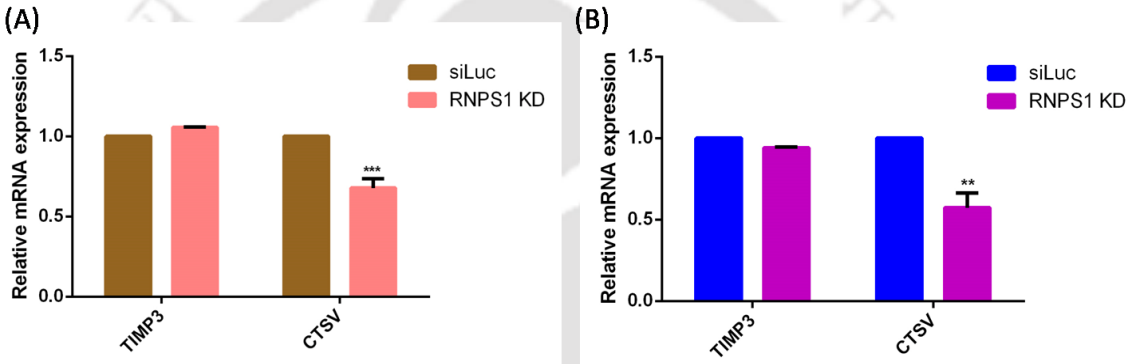


Figure 3.7: qRT-PCR of genes related to migration and invasion in RNPS1 knockdown cells (A) HeLa and (B) SiHa normalized to β-actin. Values are depicted as mean ± S.D (n=3) and p-values are depicted as **p ≤ 0.01, ***p ≤ 0.001.

3.3.4 Silencing of RNPS1 enhances chemosensitivity against drug doxorubicin

Growing evidences suggest that drug resistance is one of the main obstacles to the successful treatment of cancer. In this regard, I determined the role of RNPS1 in the chemoresistance of cervical cancer cells. RNPS1 knockdown HeLa and SiHa cells were treated with doxorubicin (Dox). Dox is an anthracycline and a widely used anticancer drug in chemotherapy to treat various types of cancer. Notably, Dox treatment dramatically reduced the viability and concomitantly increased the apoptosis of RNPS1-KD cervical cancer cells compared to the control cells (Figure 3.8). Taken together, the data suggest that depletion of RNPS1 enhances the chemosensitivity against the drug doxorubicin in cervical cancer cells.

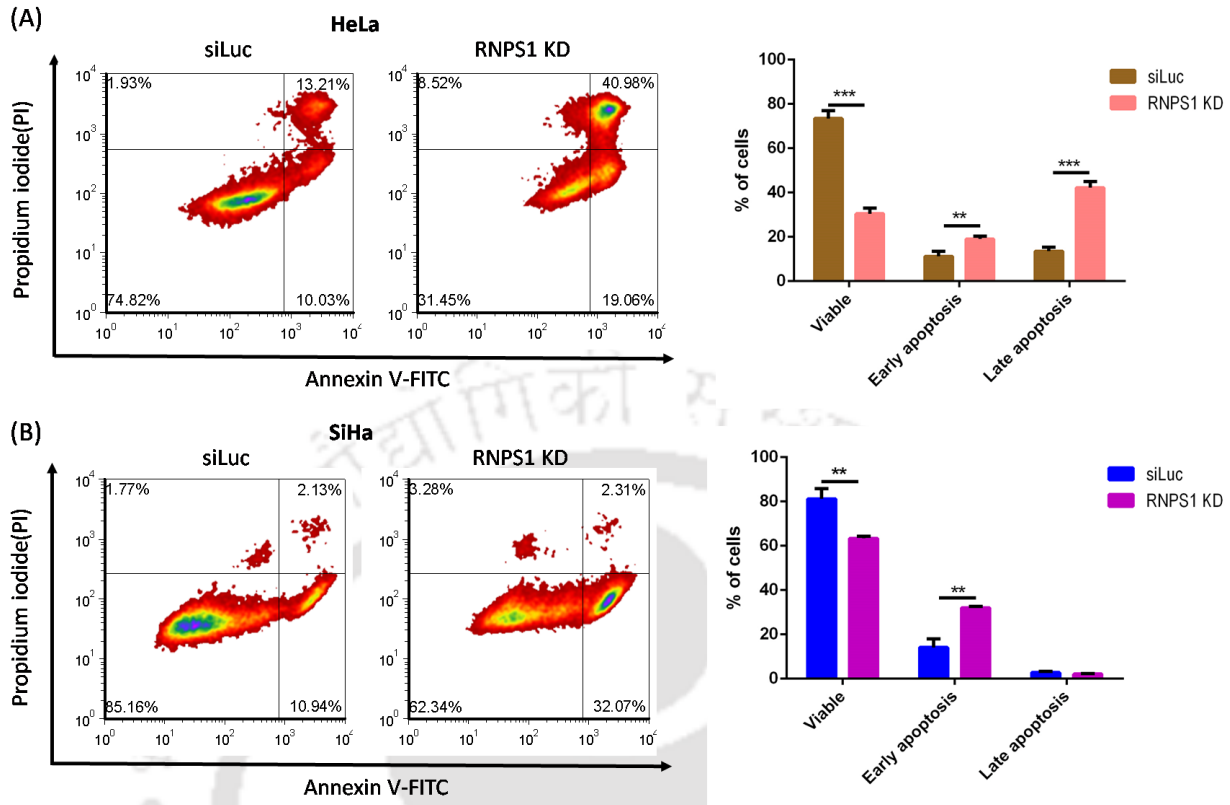


Figure 3.8: Silencing of RNPS1 enhances chemo sensitivity against doxorubicin. RNPS1 knockdown cells (after 72 hrs of siRNA treatment) were treated with 5µM of Doxorubicin for 48hrs. Apoptosis rate was quantified using flow cytometry (left panel) and its graphical representation in terms of % distribution of cells (right panel) (A) HeLa cells (B) SiHa cells. Values are depicted as mean ± S.D (n=3) and p-values are depicted as **p ≤ 0.01, ***p ≤ 0.001.

3.3.5 RNPS1 regulates the generation of cancer specific splice isoforms

RNPS1 is known to regulate alternative splicing in *D. melanogaster* and human cells. In this regard, overexpression of RNPS1 in cervical cancer cells might deregulate splicing and probably leads to differential usage of transcript isoforms (isoform switching) of the same gene. To investigate the role of RNPS1 in modulating the splicing events of cancer-specific splice isoforms, existing RNA-Seq datasets of RNPS1-KD in HeLa cells was analyzed (Appendix IV). Differential transcript usage analysis was performed using the IsoformSwitchAnalyzeR package. The result revealed 219 isoform switching after RNPS1 knockdown.

Intriguingly, RNA-Seq analysis revealed that the knockdown of RNPS1 leads to the downregulation of tumor-associated isoform of *Rac1*, *Rac1b* (Figure 3.9B). Rac1b is a hyperactive variant of the small GTPase Rac1 [154]. Rac1 is a member of the Rho family of GTPases that acts as a molecular switch during signal transduction. Rho GTPases play a central role in regulating cell protrusions, adhesion, and polarization and are therefore involved in cell migration processes [155]. Rac1b differs from Rac1 by the inclusion of an additional exon (exon 3b) that results in an impaired GTP-hydrolysis and increased GDP to GTP exchange rates (Figure 3.9A). In line with RNA-Seq analysis, qPCR and RT-PCR analysis demonstrated that RNPS1 knockdown results in the skipping of exon 3b in *Rac1* mRNA and consequently repressed the expression of *Rac1b* in cervical cancer cells (Figure 3.9C-D). Consistently, Western blot analysis showed downregulation of Rac1b protein level in RNPS1 depleted cervical cancer cells, HeLa and SiHa (Figure 3.12C-E). Importantly, RNPS1 knockdown did not alter the expression of *Rac1* mRNA in cervical cancer cells (Figure 3.9E). Taken together, these suggest that RNPS1 modulates the alternative splicing of *Rac1b*, a cancer-specific splice variant of *Rac1*.

Another essential member of the Rho family of GTPases is RhoA. Remarkably, Isoform-Switch analysis showed that RhoA is also a target of RNPS1 (Figure 3.9F-G). RhoA controls the cytoskeletal organization, cell migration, cytokinesis, and cell cycle through interaction with downstream effectors [155]. RhoA controls the contractility of actomyosin, which is necessary to produce the traction forces that pull the cell body in the direction of migration. Depletion of RNPS1 leads to a switch of the *RhoA* splice variant from a coding variant to non-coding variant lacking exons 3 and 4. The skipping of exons 3 and 4 leads to a loss of Ras domain in the non-coding variant of *RhoA*.

In accordance with the RNA-Seq analysis, the qPCR analysis showed that the knockdown of RNPS1 results in the downregulation of the *RhoA* coding isoform (Figure 3.9H). Further, Western blot analysis showed downregulation of RhoA protein level in RNPS1 depleted cervical cancer cells, HeLa and SiHa (Figure 3.12C-E). Together, these findings

confirmed that RNPS1 controls the activation of the Rho-GTPase signaling cascade in cervical cancer cells by regulating the alternative splicing of *Rac1b* and *RhoA*.

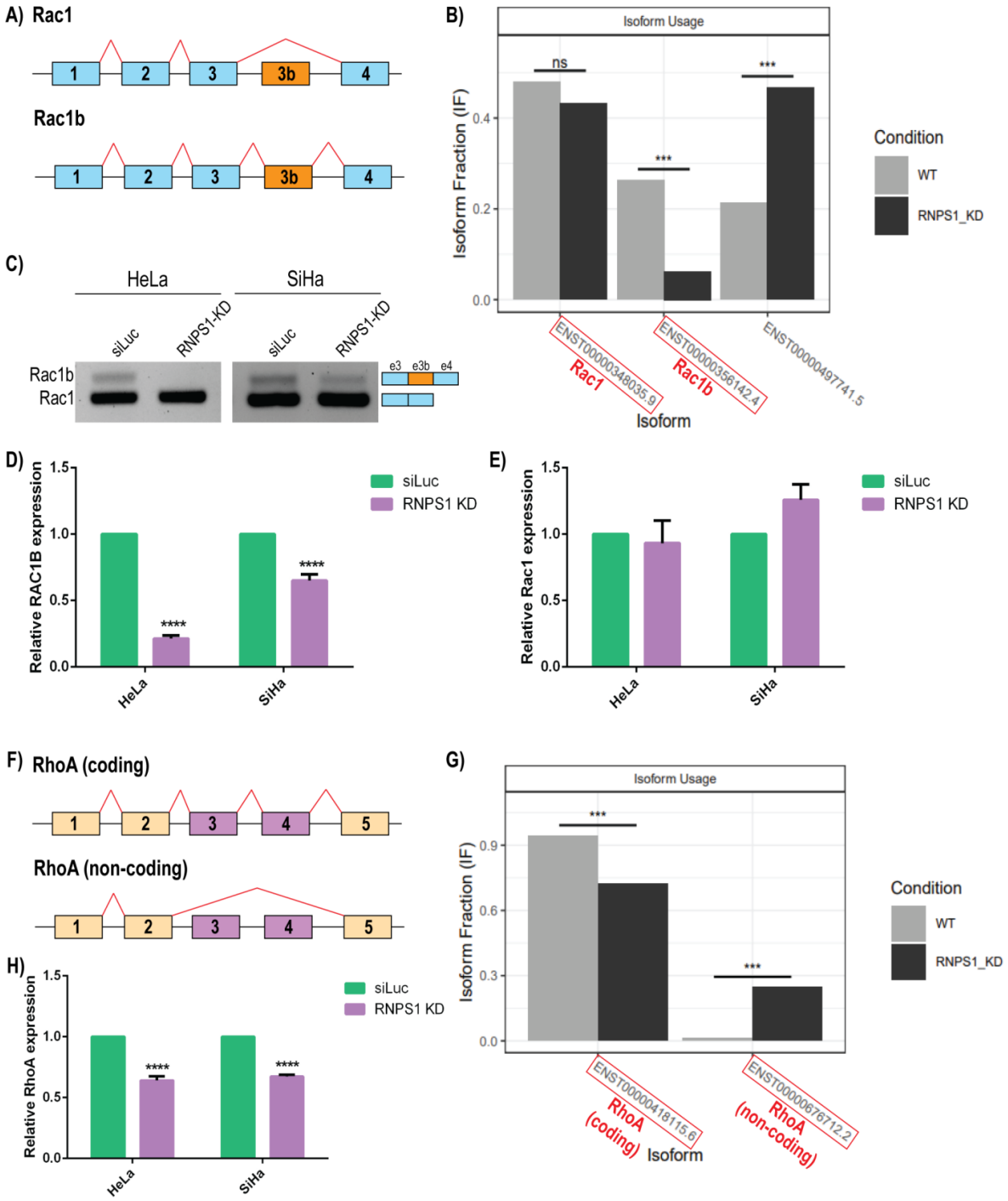


Figure 3.9: Silencing of RNPS1 modulates alternative splicing of cancer specific genes, *Rac1b* and *RhoA*. (A) Schematic representation of *Rac1* and *Rac1b* isoforms showing the alternate exon 3b. (B) Quantification of isoform fraction of *Rac1b* in wild type (Luc siRNA) and RNPS1 knockdown HeLa cells using IsoformSwitchAnalyzeR. Knockdown of RNPS1 leads to downregulation of tumor-associated isoform of

Rac1, *Rac1b* (ENST00000356142.4). The x-axis represents different transcript isoforms of *Rac1 pre*-mRNA. ENST00000348035.9 represents *Rac1* mRNA, ENST00000356142.4 represents *Rac1b* mRNA and ENST00000497741.5 represents a lesser known isoform of *Rac1 pre*-mRNA (C) RT-PCR analysis of *Rac1* exon 3b skipping with RNA from cervical cancer cells transfected with the indicated siRNA. (D) The expression level of *Rac1b* mRNA upon knockdown of RNPS1 in HeLa and SiHa by qRT-PCR. (E) The expression level of *Rac1* mRNA upon knockdown of RNPS1 in cervical cancer cells by qRT-PCR. (F) Schematic representation of isoform switching in *RhoA* mRNA showing the alternate exons 3 and 4. (G) Quantification of isoform fraction of *RhoA* using IsoformSwitchAnalyzeR. Knockdown of RNPS1 in HeLa cells leads to downregulation of coding isoform of *RhoA* (ENST00000418115.6) and upregulation of non-coding isoform (ENST00000676712.2). The x-axis represents different transcript isoforms of *RhoA pre*-mRNA. ENST00000418115.6 represents coding isoform of *RhoA* and ENST00000676712.2 represents non-coding isoform of *RhoA.* (H) The expression level of *RhoA* mRNA upon knockdown of RNPS1 in cervical cancer cells by qRT-PCR. Values are depicted as mean ± S.D (n=3) and p-values are depicted as ***p ≤ 0.001, ****p ≤ 0.0001.

Furthermore, I explored the isoform switching on MDM4, a negative regulator of p53. I found that RNPS1 knockdown leads to a switch of the *MDM4* splicing isoform from stable full-length *MDM4-FL* (coding) to unstable *MDM4-S* (non-coding) lacking exon 6 (Figure 3.10A-B). Next, RT-PCR and qPCR analysis were performed to validate the splicing event of MDM4 mRNA in RNPS1-KD cervical cancer cell lines. Consistent with the RNA-Seq analysis, RT-PCR analysis showed that depletion of RNPS1 leads to the skipping of exon 6 in the *MDM4* mRNA (Figure 3.10C). Similarly, the qPCR analysis revealed that knockdown of RNPS1 caused a marked reduction of the stable *MDM4-FL* isoform and upregulation of the unstable *MDM4-S* isoform in HeLa and SiHa cells (Figure 3.10D). Interestingly, I observed that total *MDM4* mRNA levels remained unaffected in RNPS1-KD cervical cancer cells (Figure 3.10E), indicating that RNPS1-mediated posttranscriptional event probably regulates MDM4 abundance in RNPS1-KD cells.

Another essential target of RNPS1 is WDR1 (WD-repeat domain 1), which is required for actin dynamics in processes such as cell migration and cytokinesis [156]. RNA-Seq data showed that depletion of RNPS1 in cervical cancer cells results in skipping of exons three, four and five in *WDR1* mRNA, yielding a truncated isoform of *WDR1* named *WDRΔ35* (Figure 3.10F-G). This finding was further validated by qPCR analyses (Figure 3.10H).

Additionally, RNPS1 mediates isoform switching of *CDKN2C*, *KIFC1,* and *CEP72* transcripts, as evidenced by differential usage of isoforms of the target genes upon

RNPS1 knockdown (Figure 3.11). Conclusively, these results imply that RNPS1 regulates the splicing events of cancer-associated genes and enables cervical cancer cells to generate protein isoforms favoring tumorigenesis.

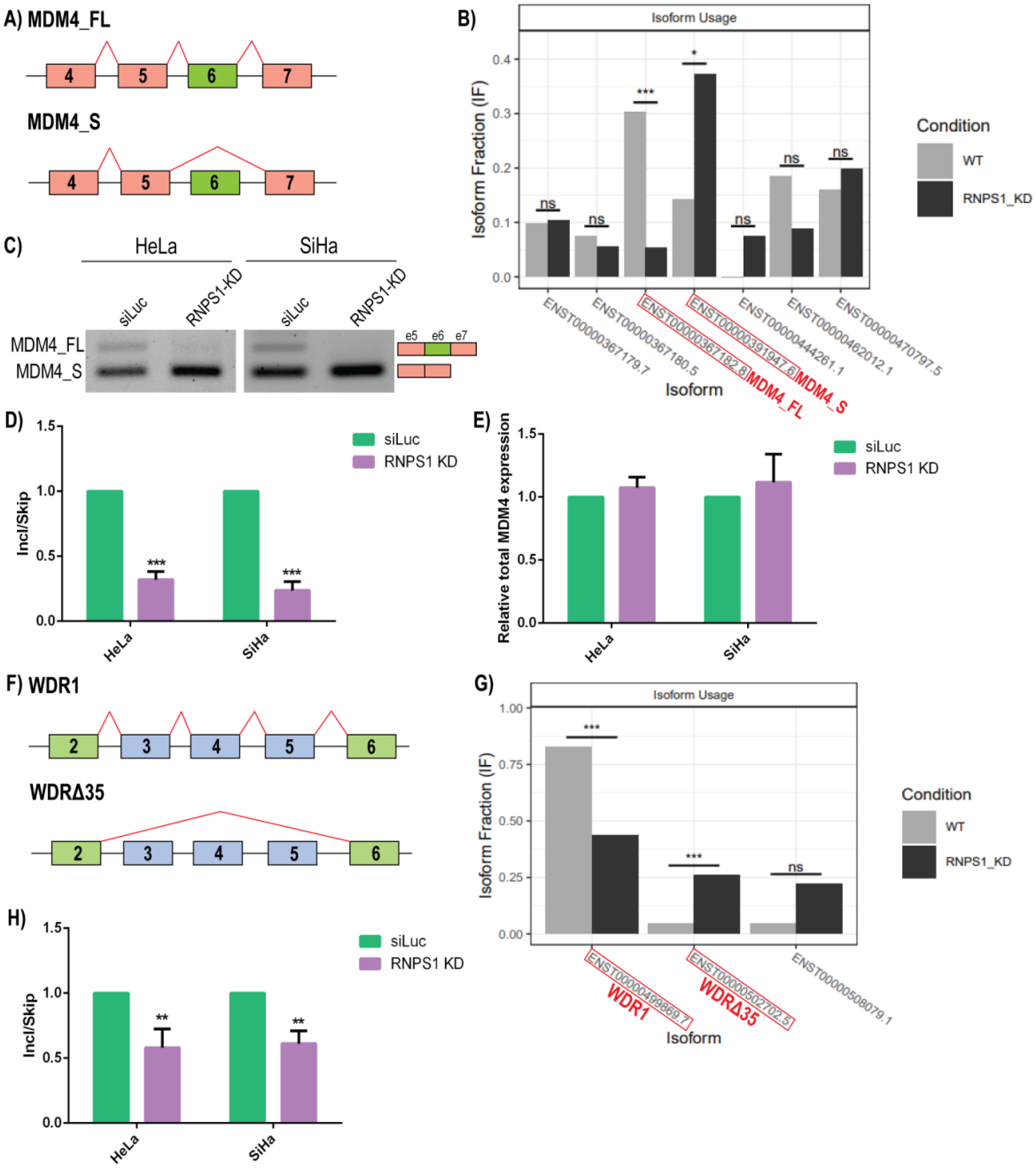
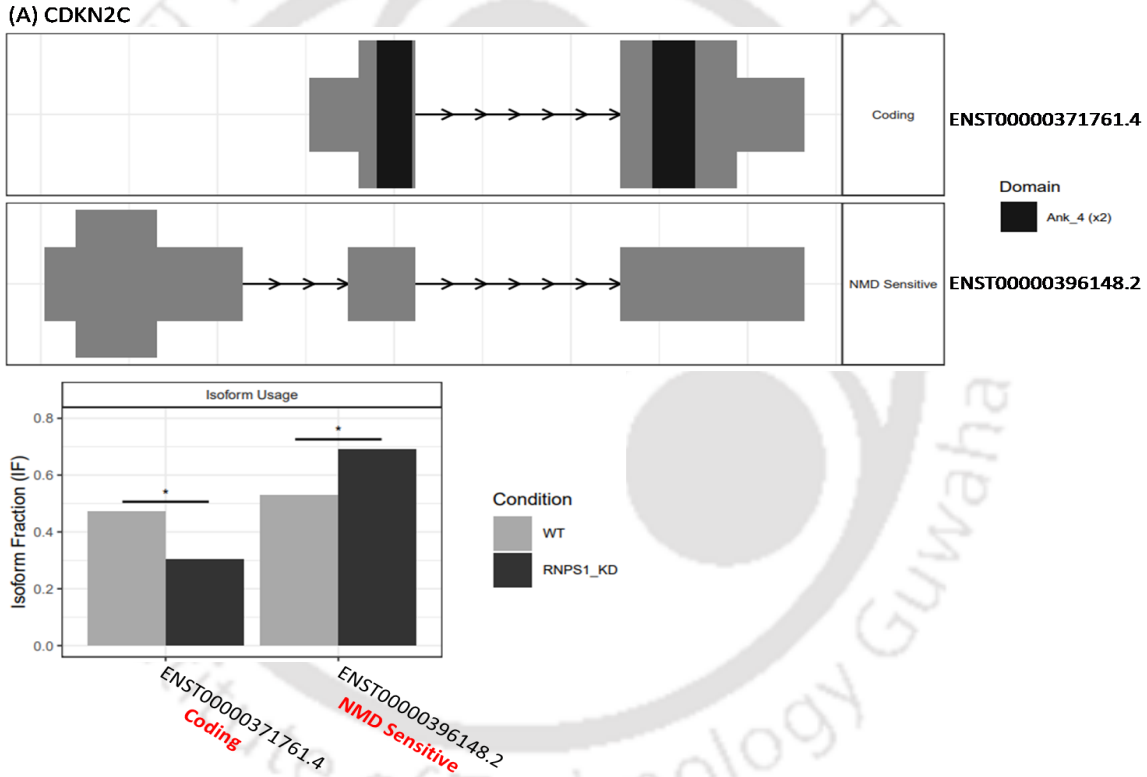
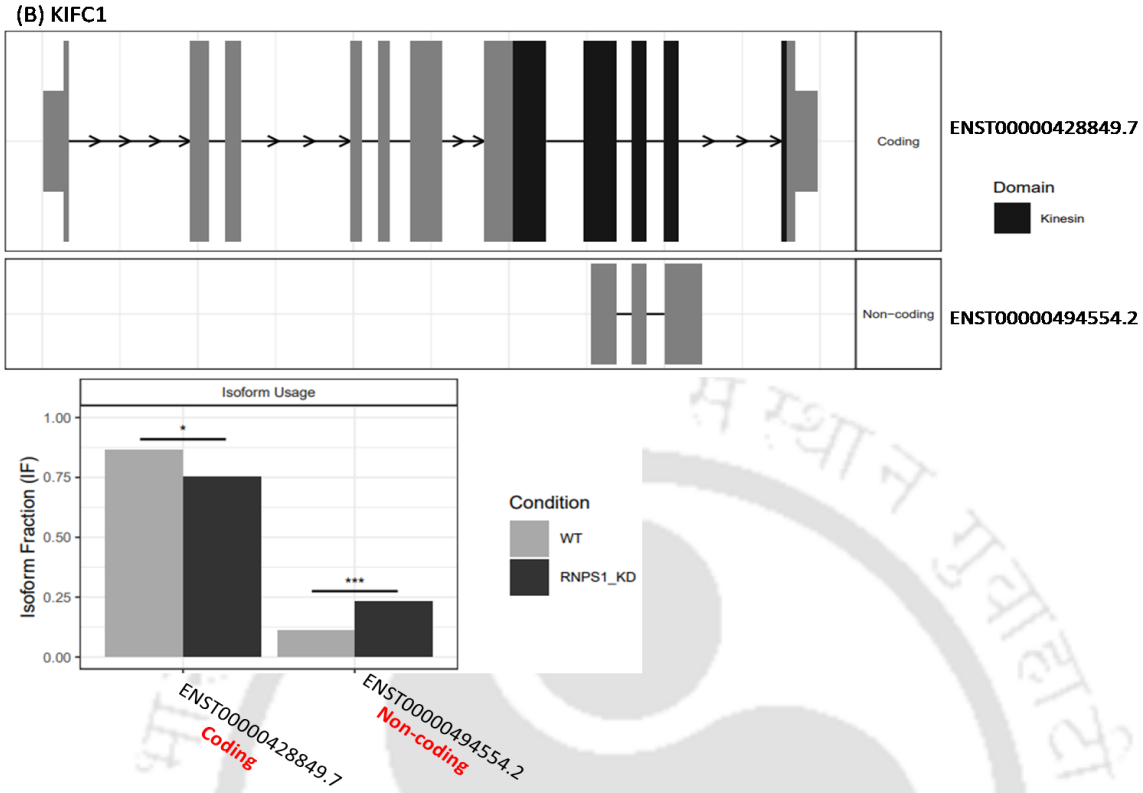


Figure 3.10: Silencing of RNPS1 modulates alternative splicing of cancer specific genes, *MDM4* and *WDR1*. (A) Schematic representation of *MDM4_FL* and *MDM4_S* isoforms showing the alternate exon 6. (B) Quantification of isoform fraction of *MDM4* in wild type (Luc siRNA) and RNPS1 knockdown HeLa cells using IsoformSwitchAnalyzeR. Knockdown of RNPS1 leads to downregulation of *MDM4_FL* (ENST00000367182.8) and upregulation of *MDM4_S* (ENST00000391947.6). The x-axis represents

different transcript isoforms of *MDM4 pre*-mRNA. ENST00000367182.8 represents *MDM4_FL* mRNA and ENST00000391947.6 represents *MDM4_S* mRNA. ENST00000367180.5 and ENST00000470797.5 represent lesser known coding isoforms of *MDM4* pre-mRNA. ENST00000367179.7, ENST00000444261.1, and ENST00000462012.1 represent non-coding isoforms of *MDM4* pre-mRNA. (C) RT-PCR analysis of *MDM4* exon 6 skipping with RNA from HeLa and SiHa cells transfected with the indicated siRNA. (D) A relative inclusion/skipping ratio (Incl/Skip) is plotted indicating that Incl/Skip ratio of *MDM4* decreases when RNPS1 is depleted. (E) The expression level of total *MDM4* mRNA upon knockdown of RNPS1 in cervical cancer cells by qRT-PCR. (F) Schematic representation of *WDR1* isoform switching showing the alternate exons 3, 4 and 5. (G) Quantification of isoform fraction of *WDR1* using IsoformSwitchAnalyzeR. Knockdown of RNPS1 in HeLa cells leads to downregulation of *WDR1* (ENST00000499869.7) and upregulation of *WDRΔ35* (ENST00000502702.5). The x-axis represents different transcript isoforms of *WDR1 pre*-mRNA. ENST00000499869.7 represents *WDR1* mRNA and ENST00000502702.5 represents WDR*Δ35* transcript isoform. ENST00000508079.1 represents non-coding isoform of *WDR1 pre*-mRNA. (H) A relative inclusion/skipping ratio (Incl/Skip) is plotted indicating that Incl/Skip ratio of *WDR1* decreases when RNPS1 is depleted. Values are depicted as mean ± S.D (n=3) and p-values are depicted as *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001.





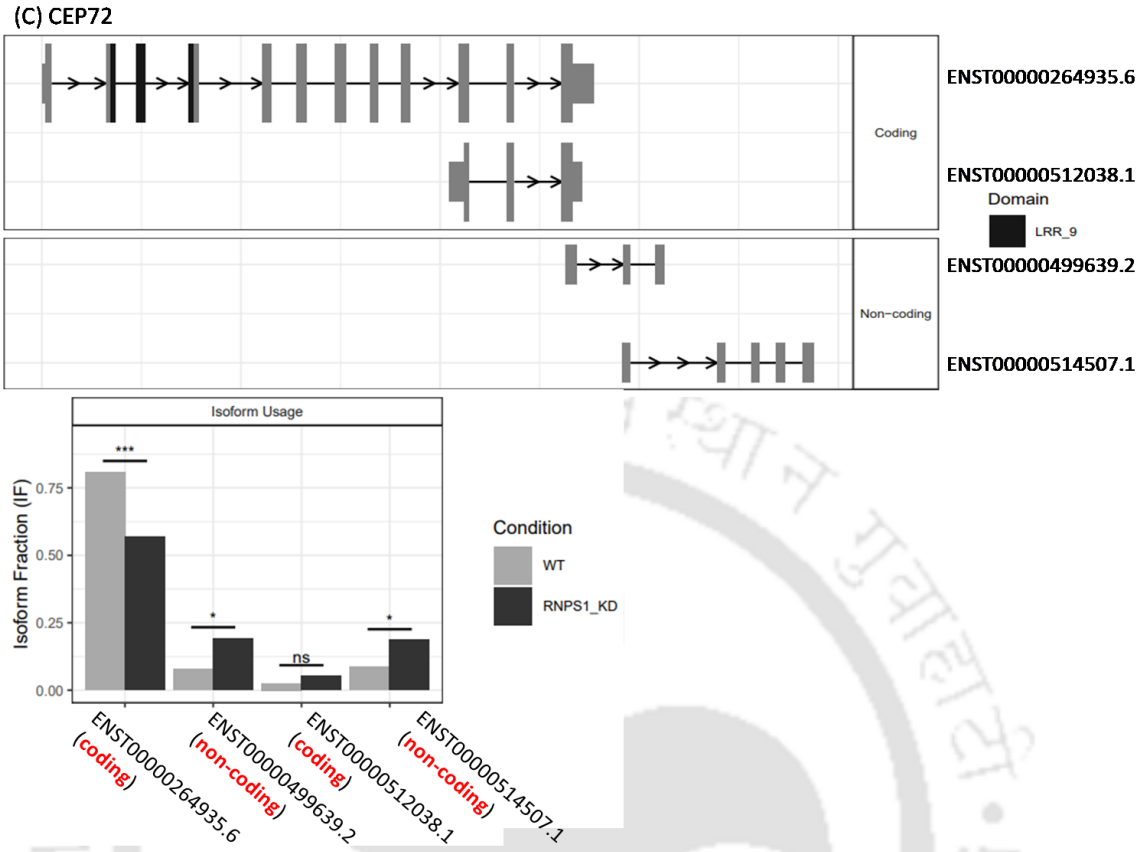


Figure 3.11: Silencing of RNPS1 modulates alternative splicing of cancer specific genes, *CDKN2C, KIFC1* and *CEP72.* (A) Quantification of isoform fraction of *CDKN2C* in wild type (Luc siRNA) and RNPS1 knockdown HeLa cells using IsoformSwitchAnalyzeR. (B) Quantification of isoform fraction of *KIFC1* in RNPS1 knockdown HeLa cells. (C) Quantification of isoform fraction of *CEP72* in RNPS1 knockdown HeLa cells.

3.3.6 Knockdown of RNPS1 modulates Notch1 and JNK signaling molecules

Several signaling pathways contribute to the development of cancer. Therefore, I further explored the downstream targets of RNPS1 by analyzing the expression of genes involved in diverse signaling pathways in cancer cells. I found that the knockdown of RNPS1 upregulates the expression of *Notch1* (Figure 3.12A-B). In line with this, it has been previously reported that down-regulation of Notch1 expression is a crucial event in HPV-induced carcinogenesis of aggressive cervical cancers as well as cervical carcinoma cells such as HeLa and SiHa [157]. Thus, RNPS1 negatively modulates the expression of *Notch1* and likely aids in the tumorigenesis of cervical carcinoma.

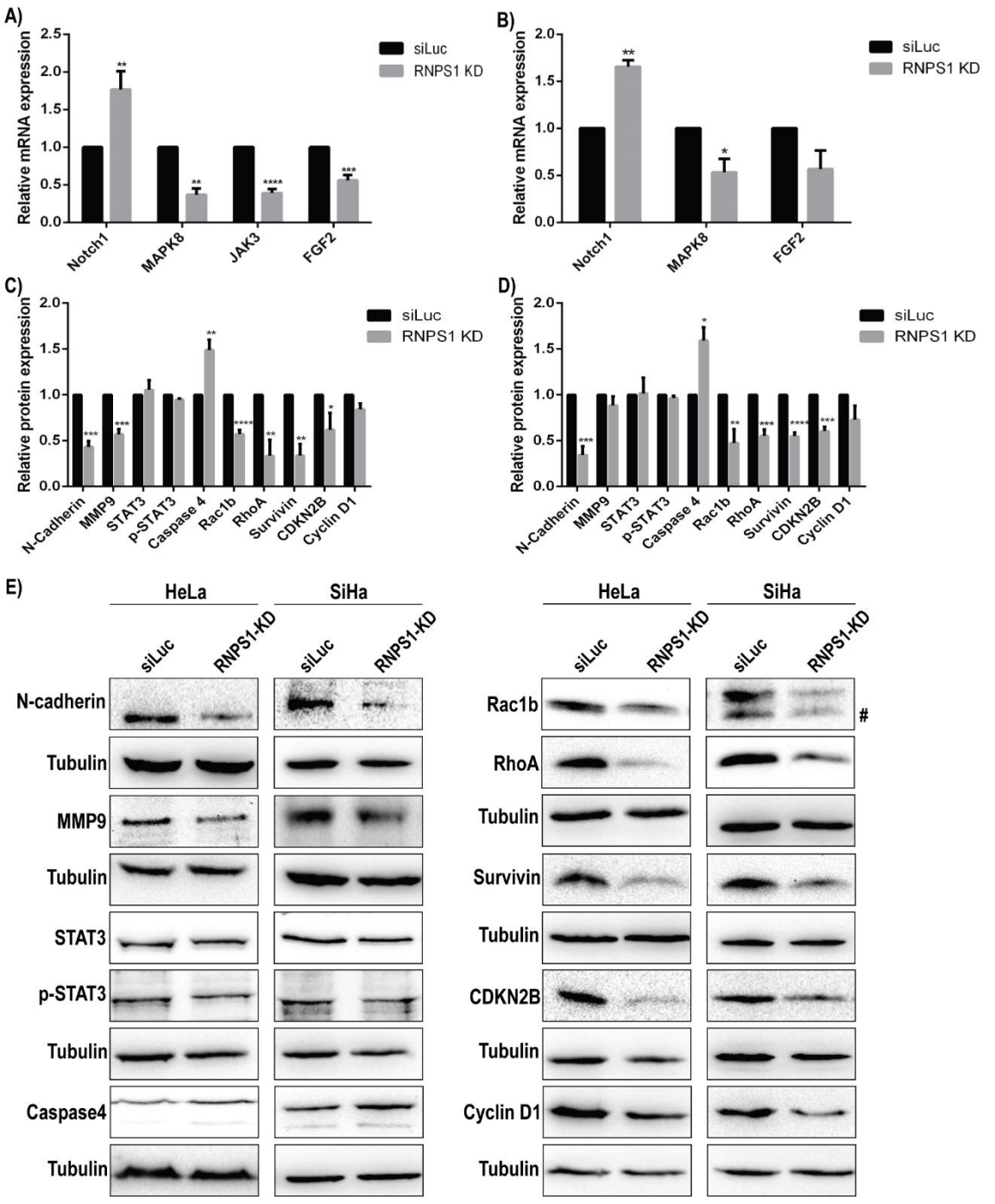


Figure 3.12: Effect of RNPS1 knockdown on various signalling molecules and markers of EMT (Epithelial-Mesenchymal transition). (A-B) qRT-PCR of genes related to Notch1 and JNK signaling in RNPS1 knockdown (A) HeLa and (B) SiHa cells normalized to β-actin. (C-D) Graph showing fold changes in protein levels of EMT markers and signaling molecules by western blotting with cell lysates from (C) HeLa and (D) SiHa cells transfected with the indicated siRNA. (E) Effect of RNPS1 knockdown on EMT markers and signaling molecules by western blotting in HeLa and SiHa cells. Tubulin was used as a normalization control. Values are depicted as mean ± S.D (n=3) and p-values are depicted as *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001, ****p ≤ 0.0001. A band corresponding to non-specific protein is labelled with #.

Aberrant activation of JNK signaling contributes to cervical cancer progression and malignancy. Intriguingly, RNPS1 substantially controls the expression of key genes of the JNK signaling pathway, such as *MAPK8* and *FGF2* (Figure 3.12A-B). Of note, *JAK3* mRNA in SiHa cells fell below the detection limit of our qPCR assay. However, further investigation showed that depletion of RNPS1 has no effect on the expression of transcription factor STAT3 and its phosphorylated form phospho STAT3 (p-STAT3) (Figure 3.12C-E). Collectively, RNPS1 probably regulates the expression of various signaling molecules and contributes to cervical cancer development.

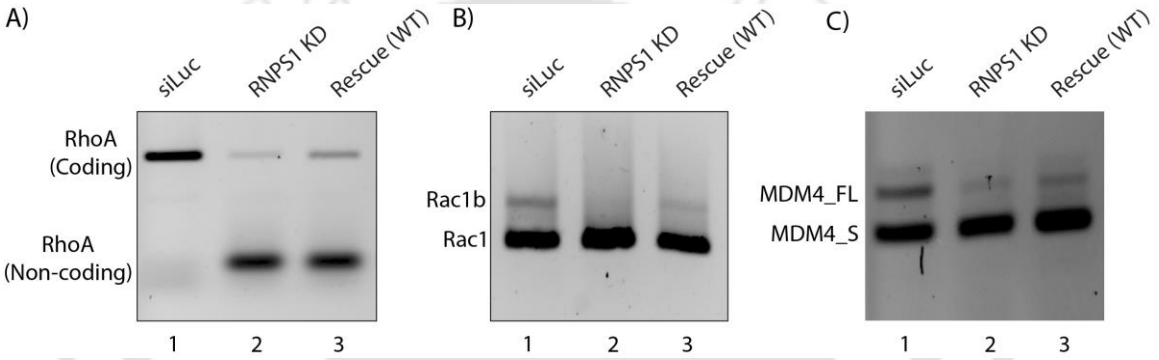


Figure 3.13 Wildtype RNPS1 partially rescues RNPS1-dependent isoform switching events. RT-PCR analysis of (A) *RhoA* (B) *Rac1* (C) *MDM4* isoform switching with RNA from HeLa cell transfected with the indicated siRNA and RNPS1-encoding siRNA resistant plasmid. Lane 1: Luc siRNA, Lane 2: RNPS1 siRNA transfected cell, Lane 3: HeLa cell cotransfected with siRNA against RNPS1 and RNPS1 encoding plasmid.

3.4 Discussion

The regulation of splicing patterns is driven by specialized splicing factors and RNA binding proteins. Emerging reports revealed that the expressions of splicing factors are frequently deregulated in multiple types of cancer. The aberrant expression and/or function of splicing regulators in tumorigenesis have become an important scientific discovery.

In the current study, I found that RNPS1 is significantly upregulated in cervical cancer tissues compared to the normal tissue in the TCGA dataset. The goal of this study was to comprehensively characterize the functions of RNPS1 in cervical carcinoma and dissect the mechanisms involved. Functionally, I found that the knockdown of RNPS1 could significantly repress the proliferation and survival of cervical cancer cells. Similar to our

findings, RNPS1 was recently reported to promote the proliferation of UCEC (uterine corpus endometrial carcinoma) tumor cells [138]. In addition, our results suggest RNPS1 promotes the clonogenic potential of cervical cancer cells. Further, I found that the down-regulation of RNPS1 resulted in G2/M phase cell cycle arrest in HeLa cells. The anaphase-promoting complex/cyclosome (APC/C) is one of the key regulators of cell cycle progression. The APC/C mediates ubiquitin-dependent degradation of cell cycle regulatory proteins to control sister chromatid segregation and cytokinesis, hence crucial for the transition from prophase to telophase in mitosis [158]. Our data suggest that RNPS1 knockdown inhibits the expression of APC/C subunits *ANAPC5* and *ANAPC7*, which presumably leads to cell cycle alteration upon depletion of RNPS1.

Furthermore, RNPS1 knockdown suppressed the migration and invasive potential of cervical cancer cells as evinced by the decrease in the number of cells that migrated or invaded the lower part of the transwell insert. These findings were further supported by the decreased expression of *CTSV* mRNA, N-Cadherin and cleaved MMP9 protein upon RNPS1 knockdown. MMP9 (gelatinase B) plays a critical role in proteolytic degradation of extracellular matrix (ECM), alteration of cell-ECM and cell-cell interactions, resulting in ECM remodeling. This remodeling is essential during tumor invasion, metastasis and modulation of the tumor microenvironment [159]. The knockdown of RNPS1 also decreased the expression of *MAPK8*, *JAK3* and *FGF2*. The JNK pathway is involved in essential cellular processes, including proliferation and survival. This pathway is constitutively expressed in many cancers and results in enhanced proliferation, malignant transformation and drug resistance [160, 161]. Accordingly, knockdown of RNPS1 plausibly modulates JNK signaling pathway, resulting in enhanced chemosensitivity of RNPS1 KD cervical cancer cells.

Importantly, this study has shown for the first time that RNPS1 participates in the regulation of several oncogenic AS events. I found that RNPS1 is involved in the regulation of *Rac1* alternative splicing. RNPS1 knockdown caused skipping of exon 3b in the *Rac1* mRNA, resulting in the downregulation of its active isoform, *Rac1b*. Rac1b

exists primarily in the active GTP-bound state, rendering it constitutively active. Rac1b is overexpressed in multiple types of cancer as compared to the normal tissues. Several studies documented that overexpression of Rac1b plays a pivotal role in tumor cell survival and malignant transformation [150, 154, 162]. Our findings indicate that RNPS1 is a critical regulator of the abundance of Rac1b in cervical cancer cells. In line with our study, a previous report has shown that the expression of Rac1b in colorectal cancer cells is based on an alternative splicing event by the SR proteins SRSF1 and SRSF3. Notably, Rac1b also confers chemoresistance in colorectal cancer cells against chemotherapeutic drugs 5-FU and OXA [163]. Hence, it is tempting to speculate that the chemosensitivity of cervical cancer cells in the absence of RNPS1 is also partly due to the depletion of the Rac1b protein.

Additionally, I demonstrate for the first time the role of a splicing factor in driving RhoA expression through an AS-based mechanism. I show that RNPS1 mediates the inclusion of exons 3 and 4 in the *RhoA* transcript giving rise to the coding splice isoform. In contrast, the absence of RNPS1 leads to skipping of exons 3 and 4, resulting in non-coding variant and a decrease in RhoA expression. Importantly, RhoA is one of the master regulators of cytoskeletal dynamics [155]. Cytoskeletal dynamic is required for invasive cancer metastasis and migration of cancer cells. Therefore, it is perceivable that RNPS1 promotes migration and invasion through modulation of the Rac1b/RhoA signaling axis.

Evidence suggests that MDM4 protein is frequently overexpressed in melanoma cells. However, surprisingly, no correlation exists between MDM4 protein levels and total *MDM4* mRNA levels [164]. The MDM4 gene produces two alternative transcripts, *MDM4-FL* and *MDM4-S. MDM4-FL* harbors exon 6 and encodes the full-length MDM4 protein, whereas the skipping of exon 6 causes the insertion of a premature stop codon and the generation of an unstable MDM4-S protein. A major consequence of skipping exon 6 in *MDM4* mRNA is a reduction in full-length MDM4 protein. This implies that the synthesis of MDM4-FL is majorly regulated at the post-transcriptional stage [165]. As a

result, it has been suggested that MDM4 overexpression in many cancer cells is mediated by an alternative splicing switch that promotes exon 6 inclusion [166]. Consistent with previous reports, I demonstrate that RNPS1 influences the inclusion of *MDM4* exon 6, generating the *MDM4-FL* transcript in cervical cancer cells. Of note, depletion of RNPS1 does not change the level of total *MDM4* mRNA. Taken together, these indicate that the alternative splicing of *MDM4* is regulated by multiple splicing factors, including RNPS1.

WDR1 is a highly conserved protein across all eukaryotes and is known to promote actin dynamics in cellular processes, including cytokinesis and cell migration [156]. Accordingly, WDR1 was found to be upregulated in the high metastatic cell line in gallbladder carcinoma compared to the low metastatic cell line [167]. Likewise, the expression of WDR1 increased in invasive ductal carcinoma and high WDR1 levels correlated with poor survival in breast cancer and lung cancer patients [168, 169]. Here, I show that the knockdown of RNPS1 decreased the expression of *WDR1* in cervical cancer cells and triggered an isoform switch from *WDR1* to *WDRΔ35.* I found out that *WDR1* and *WDRΔ35* are products of alternative splicing events mediated by RNPS1. *WDRΔ35* is a recently discovered truncated isoform of human WDR1. Nevertheless, the functional role of WDRΔ35 remains poorly characterized, which warrants future investigation.

In conclusion, our study uncovers an unknown functional role of RNPS1 in cervical cancer progression. The results suggest that RNPS1 promotes cell survival, invasion and migration of cervical cancer cells plausibly by activating the Rac1b/RhoA signaling axis. Our findings indicate that RNPS1 may function as an oncogene in cervical carcinoma and shed new light on RNPS1-mediated RNA splicing mechanism that is harnessed by the cervical cancer cells to promote its progression. It would be interesting in future to evaluate the interplay between splicing factors during malignant transformation.

*Note: The study performed in Fig. 3.3, 3.9, 3.10 and 3.12 were performed together with Pratap Chandra, Priyanka Yadav, Ayushi Rehman and Sweta Kumari.

Chapter 4:
MicroRNA mediated regulation of RNPS1 in cervical cancer cells

The work embodied in this chapter is under preparation for publication.

Deka B, Rehman A, Singh KK. miR-6893-3p is a bonafide negative regulator of splicing activator, RNPS1. 2022. Cell. Bio. International. (Manuscript Communicated).

Abstract

RNA-binding protein with serine-rich domain 1, RNPS1, is a global guardian of splicing fidelity. In the previous chapter, the expression of RNPS1 was found to be high in cervical cancer cells compared to a normal cell. Silencing of RNPS1 notably reduced migration and invasive potential of cervical cancer cells, suggesting a possible oncogenic role of RNPS1. However, the mechanisms that lead to the dysregulation of RNPS1 expression in cervical cancer cells are not known. In this study, I investigated the role of microRNAs in regulating the expression of RNPS1 in the cervical cancer cell. Using an *in silico* approach, I predicted a miRNA, miR-6893-3p, that can target and downregulate RNPS1. Interestingly, the expression of miR-6893-3p is downregulated in cervical cancer cells compared to normal cells and its level is negatively correlated with the expression of RNPS1. miR-6893-3p was validated as a miRNA that negatively regulates RNPS1 through a combination of qPCR, Western blot analysis, luciferase reporter assays and abrogation of miRNA response elements (MRE) via site-directed mutagenesis. Ectopic expression of miR-6893-3p reduced the endogenous mRNA and protein levels of RNPS1 in HeLa cells. miR-6893-3p mediated regulation of RNPS1 is dependent on the binding of miR-6893-3p to a microRNA response element in the 3'UTR of RNPS1 mRNA. Mechanistic analysis showed that targeted negative regulation of RNPS1 by miR-6893-3p occurs via enhanced mRNA degradation. Importantly, ectopic expression of RNPS1 partly reversed miR-6893-3p mediated inhibition on the migration ability of HeLa cells. Collectively, the study report for the first time that miR-6893-3p is involved in the regulation of RNPS1 expression in cervical cancer cells.

4.1 Introduction

MicroRNAs are an abundant class of short (20-22 nucleotide) non-coding RNAs that mediate post-transcriptional regulation of genes by binding to target transcripts. The miRNA biogenesis comprises sequential processing of precursor transcripts containing stem-loop structures. miRNAs are transcribed by RNA pol II into primary miRNA (pri-miRNA), which are then cleaved by the RNase III enzyme, Drosha, to generate a stem-loop precursor, pre-miRNA. The pre-miRNA is exported from the nucleus to the cytoplasm and finally processed by the RNase III enzyme Dicer to produce a miRNA duplex. The mature miRNA species may be produced from the 5' and 3' arms of the pre-miRNA, which are referred to as miRNA-5p and miRNA-3p, respectively. One of the mature miRNA species is then loaded into the RISC complex to regulate gene expression by eliciting translational repression or mRNA degradation of target mRNA.

Since miRNA-5p species is usually expressed at a higher level compared to its 3p counterpart, it has long been assumed that the poorly expressed miRNA-3p species is biologically insignificant. However, deep sequencing studies have revealed the co-existence of miRNA-5p/3p species in almost half of the analyzed miRNA populations [170, 171]. Similarly, studies from *Drosophila* showed that although miRNA-3p species are less abundant than their miRNA counterpart, they generally exist at physiologically relevant concentrations [172]. Further, most of the seed regions of miRNA-3p species are conserved across vertebrate 3' UTR evolution [173]. Moreover, luciferase reporter assays have demonstrated that both miRNA strands are functional [174].

In 2002, Dr. Croce and his group first reported the involvement of miRNA in cancer. They identified two miRNAs- miR-15 and miR-16 located at the chromosome 13q14 region, a region frequently deleted in B-cell chronic lymphocytic leukemia (CLL) [175]. miR-15 and miR-16 function as tumor suppressors by targeting and downregulating Bcl-2, an anti-apoptotic protein highly expressed in malignant non-dividing B cells [176]. Since the discovery of miR-15 and miR-16 deletions in CLL, a number of evidences have revealed the involvement of miRNA in the progression of cancer. Importantly, miRNA

profiling studies have shown that cancer cells have dramatically different miRNA profiles than respective normal cells [177-180]. MicroRNA expression is dysregulated in cancer due to various underlying mechanisms, including gain or loss of miRNA genes, epigenetic alteration, and defects in the transcriptional or miRNA biogenesis machinery. MicroRNAs can act as oncogenes or tumor suppressors depending on cancer cell types. For instance, miR-370 is frequently downregulated in different types of cancer, including ovarian cancer, cervical cancer, laryngeal squamous cell carcinoma, and colon cancer [181-184]. Some direct targets of miR-370 are MDM4, PTEN, FoxM1, Endoglin and TRAF4. miR-370 functions as a tumor suppressor by inhibiting tumor proliferation and promoting apoptosis. On the other hand, miR-370 is positively associated with the progression of human prostate carcinoma and gastric carcinoma by directly repressing the tumor suppressor FOXO1 and transforming growth factor-β receptor II (TGFβ-RII), respectively [185, 186]. Likewise, miR-6893 functions as a tumor suppressor by targeting S100A1 in cervical cancer cells.

Several recent studies have suggested that miRNAs can also influence the expression of splicing factors. For instance, virtually all SR and hnRNP orthologs in *C. elegans* are targets of miRNAs, and depletion of Argonaute protein causes changes in splice junction usage [187]. Similarly, transcriptome-wide splicing errors have been observed in Dicer knockout mice [188]. On the other hand, many SR proteins have been found to be dysregulated in various diseases, particularly cancer [127]. Interestingly, dysregulation of splicing factors in cancer is partly associated with the altered expression of complementary miRNAs [117, 118].

In the previous chapter, the oncogenic role of RNPS1 in cervical cancer cells was demonstrated. However, to date, the mechanism contributing to the regulation of RNPS1 in cancer cells is not understood. Considering the findings that miRNAs are established regulators of splicing factors, it is worth exploring whether miRNAs control the expression of RNPS1. This could provide an understanding of the probable mechanism by which RNPS1 is regulated, thereby affecting tumor progression. In this

chapter, I show that miR-6893-3p is a novel negative regulator of RNPS1 in cervical cancer cell, HeLa.

4.2 Materials and methods

4.2.1 Preparation of competent cells

Glycerol stock of E.coli DH5α or Top10 was revived on the LBA plate. A single colony of DH5α or Top10 was inoculated in 5ml LB broth and incubated overnight at 37˚C. 100 µl of overnight grown culture was inoculated into a 250ml conical flask containing 50ml LB media and incubated at 37˚C till the O.D reached ~0.35-0.4 at 600nm. The culture was aliquoted in prechilled tubes and harvested by centrifugation at 4000 rpm for 10 minutes at 4˚C. The pellet was resuspended in 25ml ice-cold CaCl2 buffer (100mM CaCl2 and 10mM Tris pH 7.4) and incubated in ice for 45 minutes. The cells were harvested by centrifugation at 4000 rpm for 10 minutes at 4˚C. The pellet was resuspended in 2.5 ml ice-cold CaCl2 buffer with glycerol, aliquoted into prechilled eppendorf tubes and stored at - 80°C.

4.2.2 Transformation of recombinant vector into competent cells

20 µl of the ligation mixture or plasmid was added to 200 µl of the competent cells. The tubes were incubated in ice for 30 minutes. Heat shock was given at 42˚C for 45 seconds and immediately transferred to the ice for 2 minutes. The cells were plated on LB agar with ampicillin (100 µg/ml) and incubated at 37˚C for 16h hrs. Positive colonies were verified by colony PCR, followed by restriction enzyme digestion and sequencing.

4.2.3 Plasmid isolation

Plasmid DNA was extracted using the Qiaprep Plasmid Miniprep kit (Qiagen) or the NucleoBond Xtra Midiprep kit (Macherey Nagel) according to the manufacturer's instructions. In general, 10 ml bacterial overnight culture in LB medium supplemented with 100 µg/ml ampicillin was used for mini-scale plasmid isolation, and 100 ml bacterial overnight culture was used for midi-scale plasmid isolation.

4.2.4 Restriction enzyme digestion

Plasmids and DNA fragments amplified by PCR were digested with suitable enzymes (NEB) according to the manufacturer's instructions. Reactions were carried out in a total volume of 50 µl, using 10 units of each enzyme per µg DNA. The reaction was carried out for 6 to 12 hrs at 37°C. Restriction fragments were visualized by agarose gel electrophoresis and gel purified using the QIAGEN Gel Extraction kit according to the manufacturer's instructions.

4.2.5 Construction of plasmids

To generate plasmids for miRNA overexpression, double-stranded DNA oligo encoding artificial mature miRNA hairpin precursor sequence containing BglII and XhoI restriction sites were ordered from IDT. pSuper basic vector (Oligoengine, VEC-PBS-0002) was double digested with BglII (NEB) and XhoI (NEB). The double digested insert and pSuper basic vector were ligated using T4 DNA ligase (NEB). 200 ng of vector were ligated with (1:5) molar ratio of the insert. Ligations were performed in a 20 µl reaction volume containing 1 µl of ligase and 2 µl of ligase buffer at 16˚C for 20 hours.

For the generation of the luciferase reporter system, full-length RNPS1 3'UTR was PCR amplified from the cDNA of HeLa cells. The PCR reaction was carried out in a total volume of 50 µl containing 2.5 µl of each 10 µM primer, 2ul template DNA, and 25 µl of 2X Q5 Hot Start High-Fidelity 2X Master Mix (NEB). The PCR reaction was carried out in a thermal cycler (Applied Biosystems) by programming the cycling profile consisting of an initial denaturation step of 30 sec at 98˚C followed by amplification for 30 cycles with denaturation for 10 sec at 98˚C, annealing for 40 sec at 67˚C, an initial extension of 1 min at 72˚C and final extension of 10 min at 72˚C for 1 cycle. The PCR product was resolved by 1% agarose gel, followed by gel purification. Purified PCR product was cloned downstream of the firefly luciferase gene in pMirglo Vector (Promega), using NheI and SalI restriction sites. pCI-Flag-RNPS1 was a kind gift from Prof. Niels Gehring, University of Cologne, Germany.

4.2.6 Cell culture

For cell culture, please refer to Chapter 3, section 3.2.1.

4.2.7 Cell transfection

Transfection was done using PEI (Biobharati) according to the manufacturer's instructions. The day before transfection, HeLa cells were seeded in a 6-well plate at the appropriate density to achieve ~75% confluency on the next day. The medium was changed 1-2 hr before the transfection. In brief, 4ug of empty vector as a negative control or miRNA overexpression plasmids and 8µl PEI were diluted in 112 µl OptiMEM (Life technologies). Dilution was pulse vortexed, incubated for 20 mins at room temperature, added to the cell culture dish, and distributed by gentle shaking. 6 hrs after transfection, the medium was replaced to remove PEI precipitates. The cells were harvested after 48 hours for RNA isolation. For experiments involving Western blotting, double transient transfections were performed. HeLa cells were initially transfected with 4ug of miRNA overexpression plasmids. After 48 hrs, cells were re-seeded in a 6-well plate. The next day, the medium was changed, and cells were re-transfected with plasmid at the same concentration. The medium was changed after 6 hrs and the cells were harvested after 48 hrs. A plasmid expressing mVenus (pCI-mVenus) was included in each experiment and served as a transfection control.

4.2.8 RNA isolation

Please refer to Chapter 3, section 3.2.3.

4.2.9 cDNA synthesis

For cDNA synthesis of mRNA, please refer to Chapter 3, section 3.2.5.

For cDNA synthesis of miRNA, the poly(A) method was employed [189]. Briefly, 1 µg of total RNA in a final volume of 10 µl, including 1 µl of 10x poly(A) polymerase buffer, 0.1 mM of ATP, and 1 unit of poly(A) polymerase (NEB) was incubated at 37°C for 30 mins

for poly (A) tailing. Next, 1 µM of universal RT primer (5'-CAGGTCCAGTTTTTTTTTTTTTTTVN-3', where V is A, C and G and N is A, C, G and T) was added to the mixture and incubated at 65°C for 5 mins. Finally, 0.1 mM of each deoxynucleotide (dATP, dCTP, dGTP and dTTP), 200 units of reverse transcriptase (Biobharati) and 4 µl buffer were added to the mixture in a final volume of 20 µl and incubated at 25°C for 10 mins, followed by 42°C for 50 mins. Afterward, the enzyme was heat-inactivated at 70°C for 15 mins.

4.2.10 Quantitative real-time PCR

For mRNA-specific qPCR, please refer to Chapter 3, section 3.2.7.

The poly(A) method was performed for miRNA-specific qPCR, and primers were designed using the software miRprimer [190]. The input for miRprimer is miRNA sequence in fasta format and is saved as a text file named input_miRs.txt. It is saved in the same folder with miRprimer. The next step is to execute miRprimer by double-clicking program's icon. The output file termed result_best_primer_pairs.txt was used to select miRNA-specific primers. qRT-PCR was performed using PowerUp Sybr green master mix (Invitrogen). For each reaction, 5ul of 2X Sybr green master mix was mixed with 0.25 uM of each primer, 1ul of diluted cDNA (1:3 dilution of the cDNA), and the rest was filled with nuclease free water. Specific primers were used for quantifying gene expression and normalized with RNU6 expression. Relative gene expression was calculated using the $\Delta\Delta C_T$ method.

4.2.11 mRNA decay analyses

HeLa cells were transfected with empty vector or miRNA overexpression plasmids. After 48 hrs, mRNA decay experiments were initiated by adding actinomycin D (ActD, final conc. 10µg/mL) (Sigma) to the transfected cells. 3, 6, 8, 9 hr after actinomycin D treatment, total RNA was harvested by using TRIzol (Invitrogen). First-strand cDNA synthesis was performed as described in Chapter 3, section 3.2.5. Real-time PCR was performed with PowerUp Sybr green master mix (Invitrogen). RNPS1 mRNA relative

abundance was determined by normalizing to β-actin using the $\Delta\Delta C_T$ method. The ΔC_T value of each time point was normalized to the ΔC_T value of t = 0 to obtain the $\Delta\Delta C_T$ value. The relative amount of RNPS1 mRNA without actinomycin D treatment at 0 hr was set to 1. The mRNA fraction remaining at each time point was calculated by $2^{(-\Delta\Delta CT)}$. Excel program was used to fit the best exponential curve to the data points. The curve will have the general equation: $C = C_0\ e^{-k_{decay}\ t}$, where C_0 is the concentration of the mRNA at time 0 before the decay starts. C is the concentration of mRNA at time t. The half-life of an mRNA ($t_{1/2}$) is inversely proportional to its decay rate constant (k_{decay}) [191].

$t_{1/2} = \ln(2) / k_{decay}$

4.2.12 Site-directed mutagenesis of the miRNA target site

To induce point mutations in pMirglo-luc-RNPS1_3'UTR, site-directed mutagenesis was performed (NEB). The wild-type RNPS1 3'UTR expression plasmid served as a template in the PCR. Primers were designed using the NEB online design software, NEBaseChanger. For each PCR reaction, 10 ul of Q5 Hot Start High-Fidelity 2X Master Mix was mixed with 1 ul of each 10uM primer, 10ng of template DNA and the rest was filled with nuclease-free water. The cycling profile consists of an initial denaturation step of 30 sec at 98˚C followed by amplification for 25 cycles with denaturation for 10 sec at 98˚C, annealing for 1 min at 63˚C, an initial extension of 4.5 min at 72˚C and final extension of 5 min at 72˚C for 1 cycle.

After exponential amplification, 1ul of PCR product was treated with kinase, ligase and DpnI (KLD) and incubated at room temperature for 15 mins. The KLD mix was then transformed into Top10 competent cells. Positive clones were verified by sequencing.

4.2.13 Luciferase assay

Transient co-transfection of cells with pSuper construct and luciferase reporter constructs (pmirGLO) containing the RNPS1 3'UTR were accomplished using PEI. After 72 h, the activities of firefly and renilla luciferase were measured using the Dual-Glo-

Luciferase Assay system (Promega) according to the manufacturer's instructions. Briefly, 75 µl Dual-Glo luciferase buffer was added to cells in 75 µl medium (96 well plate) followed by incubation at room temperature for 15 mins for cell lysis. Firefly luminescence was measured by a luminometer (Promega). Next, 75 µl Stop & Glo reagent was added and incubated at room temperature for 15 mins. Renilla luminescence was measured. Firefly luciferase measurements were normalized to Renilla luciferase readings.

4.2.14 Western Blotting

Please refer to Chapter 2, section 2.2.7.

4.2.15 Transwell migration assay

Please refer to section Chapter 3, section 3.2.10.

4.2.16 In silico analysis

TargetScanHuman v.7.1 and miRTarBase were used to identify miRNAs that can regulate RNPS1. RNAfold was used to evaluate structural changes in RNPS1 3'UTR upon introduction of mutations to disrupt the miRNA binding site. RNAhybrid was used to evaluate miRNA binding with a 3'UTR sequence by calculating the free energy of binding.

4.2.17 Statistical analysis

Please refer to section Chapter 3, section 3.2.15.

4.3 Results

4.3.1 In silico prediction of potential miRNAs targeting RNPS1 mRNA

To investigate the potential role of microRNA in regulating RNPS1 gene expression, targetscan and miRTarBase databases were used to predict miRNAs that might bind to

the RNPS1 mRNA. TargetScan predicts targets of miRNAs based on the complementarity between the seed sequence and seed matches and also takes into account factors such as conservation of seed sequence across organisms. miRTarBase database contains microRNA-Target interactions supported by experiments including, microarray, next-generation sequencing, western blot, and reporter assay.

MicroRNAs: miR-370-3p, miR-6893-3p, miR-33b-5p, miR-590-3p, miR-149-5p and miR-490-3p were selected in this study for further analyses. All these miRNAs are predicted to target 3'UTR of RNPS1. The predicted binding sites of miR-490-3p and miR-33b-5p in the 3'UTR of RNPS1 are conserved across vertebrates. Similarly, the predicted binding site of miR-149-5p in RNPS1 3' UTR is conserved across mammals. Further, miRTarBase predicted the binding of miR-370-3p, miR-149-5p, miR-590-3p and miR-490-3p to RNPS1 mRNA based on CLIP-Seq analysis. Together, these imply the likelihood of these miRNAs to act as regulators of RNPS1.

The candidate miRNAs were also filtered based on the free energy of hybridization between candidate miRNAs and the predicted binding sites within the 3'UTR of RNPS1, which was calculated using the RNAhybrid program. This program considers mRNA:miRNA duplex formation as a thermodynamic feature and calculates the minimum free energy (mfe) changes that take place during the interaction between miRNA and its target mRNA. The free energy change indicates the strength of each binding position.

The minimum free energy of hybridization of miR-33b-5p, miR-590-3p, miR-149-5p, miR-370-3p, and miR-490-3p are -26.3 kcal/mol, - 14.8 kcal/mol, -31.4 kcal/mol, -29.4 kcal/mol, and 30.6 kcal/mol respectively (Figure 4.1). miR-6893-3p showed the most stable base-pairing (-38.3 kcal/mol) with its predicted binding site in RNPS1 3'UTR. Most of these candidate miRNAs have high negative delta G values (mfe ≤ -29 kcal/mol), implying that these miRNAs have a higher probability of binding to the seed regions within the 3' UTR of RNPS1. Taken together, miR-33b-5p, miR-590-3p, miR-149-5p, miR-370-3p, miR-490-3p and miR-6893-3p are probable miRNA regulators of RNPS1.

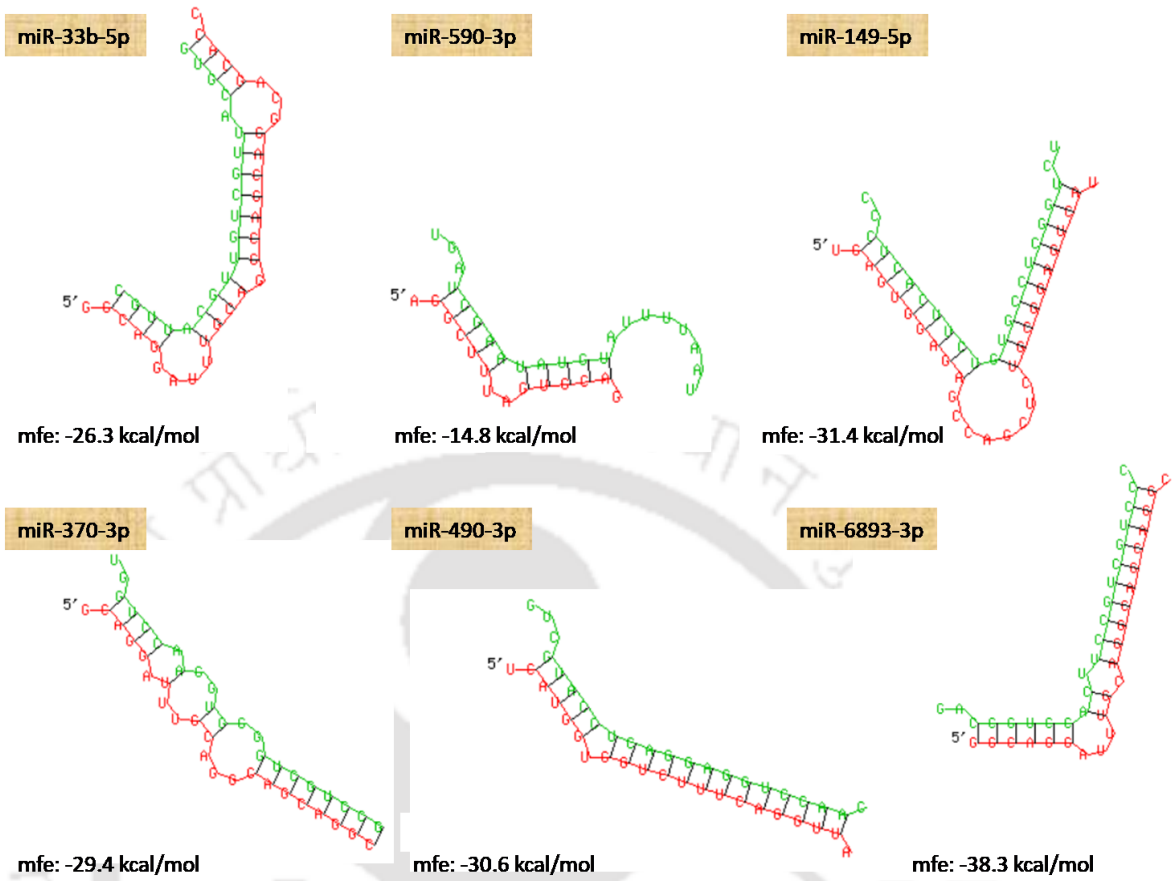


Figure 4.1: The free energy of hybridization of miRNAs. The candidate miRNAs are miR-33b-5p, miR-590-3p, miR-149-5p, miR-370-3p, miR-490-3p and miR-6893-3p. Red line represents binding sites within RNPS1 3'UTR and green line represents miRNA.

4.3.2 Generation of miRNA overexpression plasmids

The widely used method of overexpressing miRNA is cloning genomic sequences corresponding to the miRNA hairpin precursor into an overexpression vector. However, this method is generally unsuitable for specific expression of the miRNA-3p species, without concurrent expression of the miRNA from the 5' arm (miRNA-5p). Consequently, the function of miRNA-3p species has only recently begun to be recognized compared to its miRNA counterpart.

In order to generate plasmids that can overexpress specific species of miRNA, a suitable alternate strategy was followed to overexpress any mature miRNAs without bias [192]. For this, an artificial miRNA hairpin precursor sequence was designed. Herein, the mature miRNA sequence was placed in the 3' strand of the stem, it's complementary

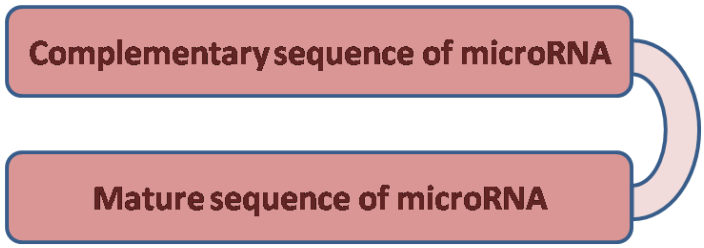


Figure 4.2: Schematic diagram of the stem-loop miRNA precursor for the overexpression of miRNA species. The 3' arm of the precursor contained the mature sequence of miRNA species, while the 5' arm contained the complementary sequence of the miRNA species.

sequence was placed in the 5' strand, and a loop sequence was inserted between the two (Figure 4.2). The DNA sequence corresponding to the miRNA hairpin precursor sequence was cloned into the pSuper basic vector. This vector employs the pol-III H1-RNA promoter and has a termination signal comprising five thymidines in a row, thereby can produce short RNAs without a poly A tail. Hence, mature miRNA sequences of candidate miRNAs were cloned into pSuper expression vector (Figure 4.3).

4.3.3 Identification of candidate miRNA that regulates RNPS1

To examine the possibility of RNPS1 gene regulation by miRNA, HeLa cells were transfected with miRNA overexpression plasmid. The ectopic expression of miR-33b-5p does not regulate the level of RNPS1 mRNA. These indicate that miR-33b-5p does not regulate RNPS1 at the mRNA level. Similarly, overexpression of miR-590-3p does not change the expression of RNPS1 mRNA. Moreover, ectopic expression of miR-370-3p and miR-149-5p also did not alter the level of RNPS1 mRNA (Figure 4.4). Taken together, these indicate that miR-33b-5p, miR-590-3p, miR-370-3p, and miR-149-5p do not regulate RNPS1 at the mRNA level.

I next investigated if RNPS1 is a potential target of miR-490-3p and miR-6893-3p. For this, HeLa cells were transfected with miR-490-3p and miR-6893-3p. qPCR analysis showed that miR-490-3p modestly reduced the endogenous *RNPS1* at the mRNA level, whereas miR-6893-3p strongly downregulated the expression of *RNPS1* mRNA (Figure 4.4).

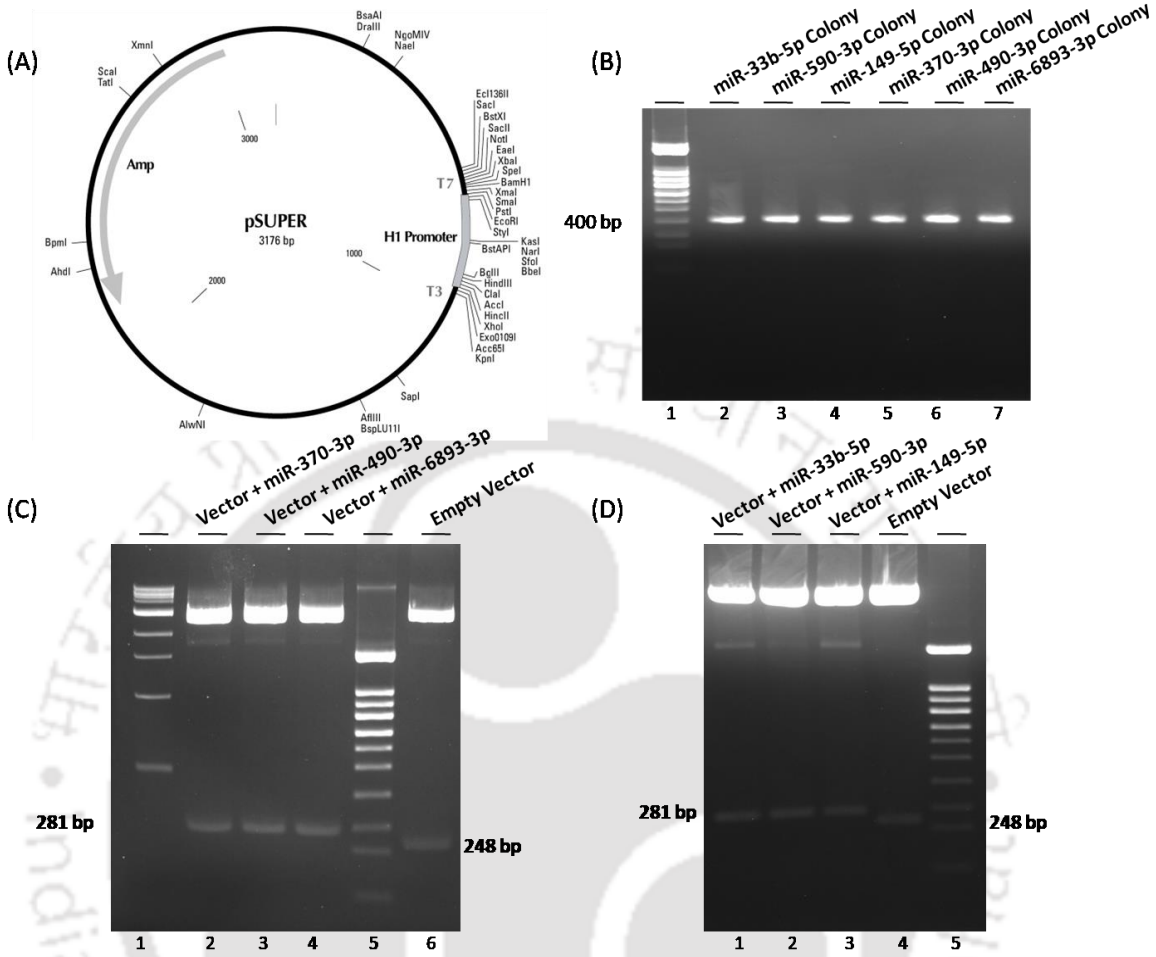


Figure 4.3: Generation of miRNA overexpression plasmids. (A) Schematic diagram of pSuper empty vector. (B) Mature miRNA hairpin containing sequences were ligated into pSuper basic vector. Positive colonies were confirmed by colony PCR. (C-D) miRNA overexpression plasmids were further validated through restriction enzyme digestion. Empty vector served as negative control. The positive clone (vector with insert) releases an insert of size around ~281 bp upon digestion with EcoRI and XhoI, whereas negative clone without insert releases an insert of size ~248 bp.

Further, Western blot was performed and results revealed a significant reduction in RNPS1 protein level upon ectopic expression of miR-6893-3p in HeLa cells (Figure 4.5). Altogether, this suggests that miR-6893-3p is a regulator of RNPS1 that negatively controls RNPS1 at both mRNA and protein levels.

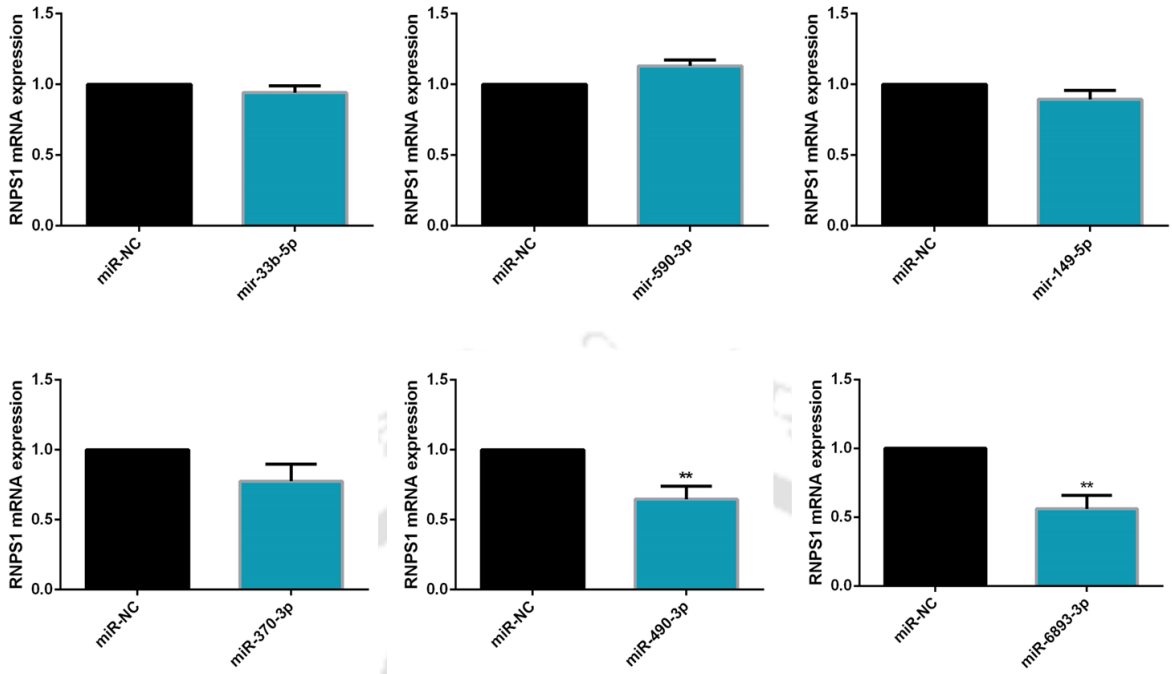


Figure 4.4: Effect of miRNA overexpression on mRNA level of RNPS1. qPCR of RNPS1 mRNA from HeLa cells transfected with pSuper-empty vector or pSuper-miRNA viz., miR-33b-5p, miR-590-3p, miR-149-5p, miR-370-3p, miR-490-3p and miR-6893-3p. RNPS1 mRNA expression was normalized to β-actin mRNA expression. The values represent mean±SD (t-test, n=3). **P<0.01

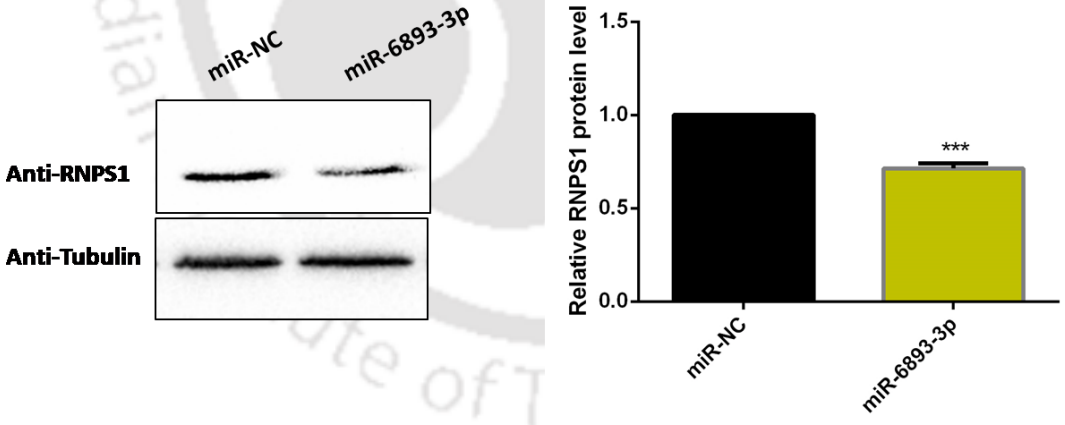


Figure 4.5: RNPS1 is negatively regulated by miR-6893-3p. Western blot of endogenous RNPS1 protein from HeLa cells transfected with pSuper-empty vector or pSuper-miR-6893-3p and normalized to tubulin level. The values represent mean±SD (n=3). ***P<0.001

4.3.4 Expression of miR-6893-3p in cervical cancer cell lines

In the previous chapter, RNPS1 expression was found to be high in cervical cancer cells compared to normal cells. To determine whether miRNAs are partly responsible for the

aberrant expression of RNPS1 in cervical cancer cells, the expression patterns of miRNAs in cervical cancer cells, HeLa and SiHa, were examined. The expression of miR-6893-3p in cervical cancer cells is low compared to the normal cell line, HDF (Primary Human dermal fibroblast) (Figure 4.6). These results suggest that negative regulation of RNPS1 by miRNAs might serve as one of the molecular mechanisms underlying the high expression of RNPS1 in cervical cancer cells.

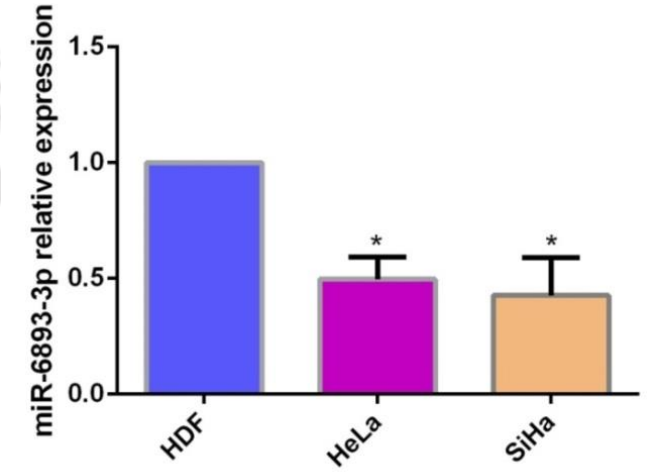


Figure 4.6: miR-6893-3p is downregulated in cervical cancer cells. qPCR shows the relative expression of miR-6893-3p in human cervical cancer cell lines (HeLa and SiHa) and normal cell line, HDF.

4.3.5 miR-6893-3p binds to the seed region on the 3' UTR of RNPS1 mRNA

I next investigated whether RNPS1 3' UTR acts as a post-transcriptional modulator of the expression of RNPS1 by using a 3' UTR luciferase reporter system. In this regard, the full-length RNPS1 3' UTR was inserted downstream of the firefly luciferase gene in the pmirGLO vector (Figure 4.7). Empty vector pmirGLO provides a constitutive expression of the firefly luciferase protein under the PGK promoter and SV40 polyadenylation signal. Renilla luciferase gene in the same vector serves as a control reporter for normalization. The luciferase reporter and miR6893-3p constructs were then co-transfected into HeLa cells and luciferase activity was measured via a luminometer. The result showed that the insertion of RNPS1 3' UTR resulted in a decrease in luciferase activity. However, the luciferase activity of cells co-transfected with pmirGLO empty

vector and miR6893-3p did not change (Figure 4.7). This result suggests that miR-6893-3p directly binds to 3' UTR of RNPS1 mRNA and negatively regulates its expression. This finding also corroborates the prediction of the RNAhybrid tool that miR-6893-3p forms high stable base-pairing with the predicted seed region within RNPS1 3'UTR (Figure 4.1).

The 3' UTR of RNPS1 contains one predicted binding site of miR-6893-3p. To validate the binding sites of miR-6893-3p on RNPS1 3' UTR, I performed site-directed mutagenesis of predicted binding sites of the miRNA on RNPS1 3' UTR. *In silico* analysis revealed that these mutations only had a minor effect on the secondary structure of the 3'UTR while drastically decreasing the negative free energy of binding of miR-6893-3p (Figure 4.8). This suggests a weaker binding potential between miR-6893-3p and the seed region within RNPS1 3' UTR. Dual luciferase assays using cells co-transfected with the mutant luciferase reporter and miR-6893-3p revealed that the suppression of luciferase activity was completely abolished when mismatch mutations were introduced into the seed region of miR-6893-3p on RNPS1 3' UTR (Figure 4.7). The destruction of the binding site derepressed the RNPS1 3'UTRs from negative regulation due to the loss of binding by miR-6893-3p. Thus, these data indicate that the predicted MRE is critical for the direct and specific binding of miR-6893-3p to the RNPS1 mRNA.

4.3.6 miR-6893-3p promotes rapid decay of RNPS1 mRNA

Multiple lines of evidence pointed that miRNA-mediated mRNA degradation to be a predominant molecular mechanism of miRNA-mediated regulation of gene expression. To examine whether miR-6893-3p plays a role in decreasing RNPS1 mRNA stability, HeLa cells were transfected with miR-6893-3p and treated with actinomycin D to halt mRNA transcription and the half-life of the *RNPS1* transcript was calculated. Actinomycin D is an antibiotic that intercalates into DNA and blocks RNA pol II-mediated transcription. Total RNA was isolated at different time intervals after actinomycin D treatment and RNPS1 mRNA levels were quantified by qPCR. The calculated half-life of the *RNPS1* transcript in HeLa cells transfected with the miR-6893-3p was 3.8 hr, compared to a half-life of 7.3 hr in the negative control (empty vector-transfected cells) (Figure 4.9).

These data indicate that *RNPS1* mRNA amounts decreased in a time-dependent manner in actinomycin D treated control cells; however, the *RNPS1* mRNA dramatically reduced in miR-6893-3p transfected cells. miR-6893-3p promoted accelerated decay of endogenous *RNPS1* mRNA, with >3-fold reduction in mRNA half-life compared to control miRNA. Thus, targeted negative regulation of RNPS1 by miR-6893-3p occurs via enhanced mRNA degradation, resulting in a suppression of RNPS1.

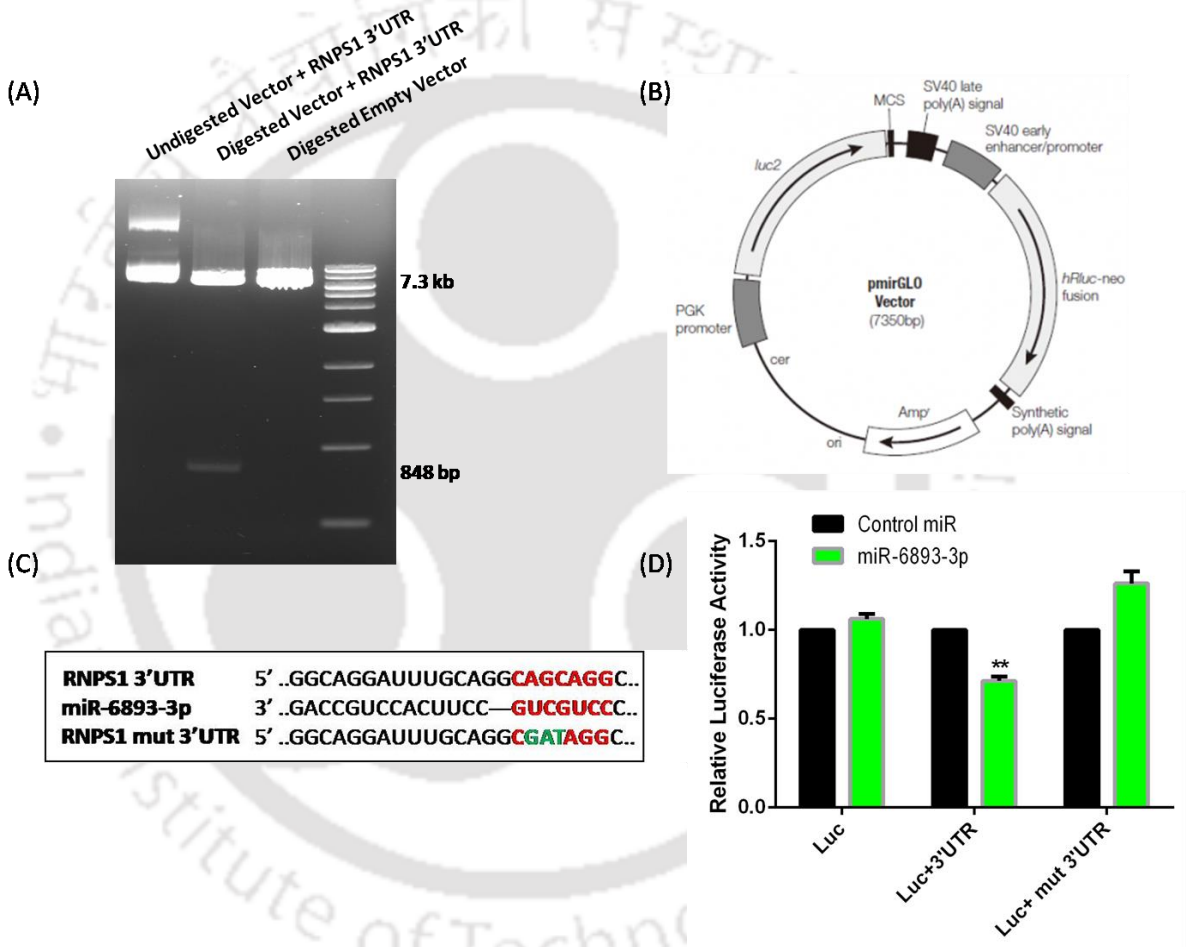


Figure 4.7: miR-6893-3p targets the RNPS1 3'UTR. (A) Generation of luciferase reporter vector with RNPS1 3'UTR. Positive colony was validated through restriction digestion in lane 2, empty vector in lane 3 served as negative control. Lane 1 contains undigested RNPS1 3'UTR luciferase reporter plasmid. (B) Schematic diagram of pMirglo empty vector. (C) Prediction of binding sites of miR-6893-3p in 3'UTR of RNPS1 using Targetscan and site-directed mutagenesis of predicted seed region. (D) HeLa cells were co-transfected with luciferase empty vector or luciferase 3'UTR reporter construct or luciferase mutated 3'UTR reporter and miR-6893-3p (green bars) or control pSuper empty vector (black bars). Luciferase activity was normalized to renilla luciferase. The values represent mean±SD (n=3). **P<0.01.

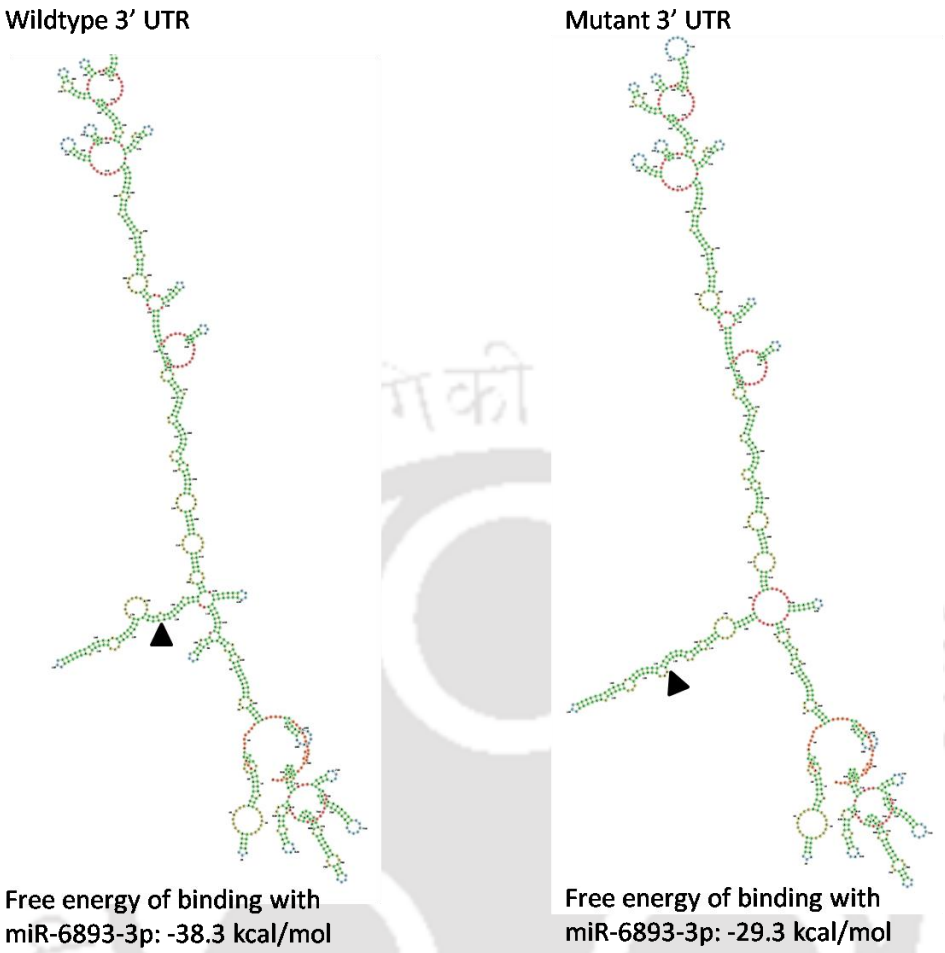


Figure 4.8: Secondary structure analysis of wildtype and mutated RNPS1 3'UTR. Changes in the free energy of hybridization of miR-6893-3p with RNPS1 3'UTR following site-directed mutagenesis of the seed sequence. Black arrow represents the binding site of miR-6893-3p.

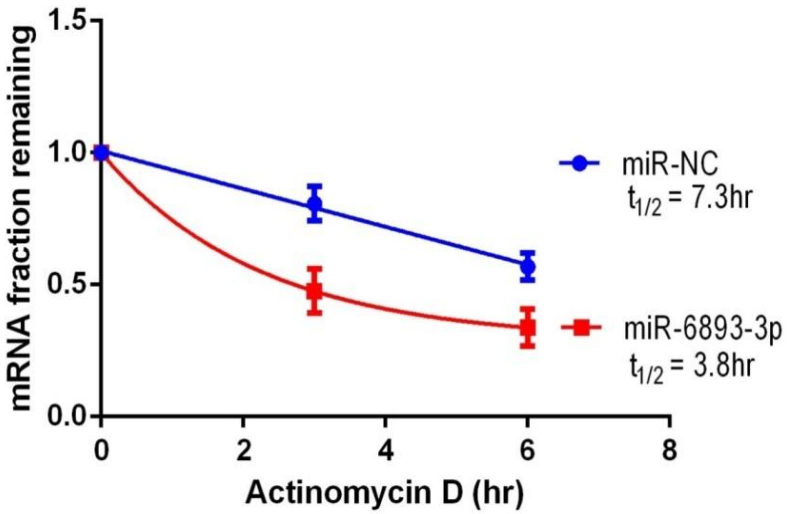


Figure 4.9: miR-6893-3p reduces the stability of RNPS1 mRNA. HeLa cells were transfected with pSuper empty (miR-NC) or pSuper-miR-6893-3p. Actinomycin D was added and endogenous RNPS1 mRNA decay was analysed by qPCR using β-actin for normalization.

4.3.7 RNPS1 reverses the migration inhibiting effect of miR-6893-3p

Recent studies have shown that elevated expression of miR-6893 inhibits the migratory ability of HeLa cells. Moreover, in the previous chapter, I have determined that RNPS1 promotes the migration potential of cervical cancer cells. To further evaluate the functional interaction between miR-6893-3p and RNPS1 in HeLa cells, a rescue experiment was carried out by co-transfection of miR-6893-3p and RNPS1 plasmids. Interestingly, the result revealed that ectopic expression of miR-6893-3p attenuated the migration ability of HeLa cells. On the other hand, co-transfection of miR-6893-3p and RNPS1 appeared to reverse the inhibitory role of miR-6893-3p on the migration of HeLa cells (Figure 4.10). These results suggest that overexpression of RNPS1 restored the miR-6893-3p mediated tumor-suppressive effects, as represented by enhanced cell migration of HeLa cells.

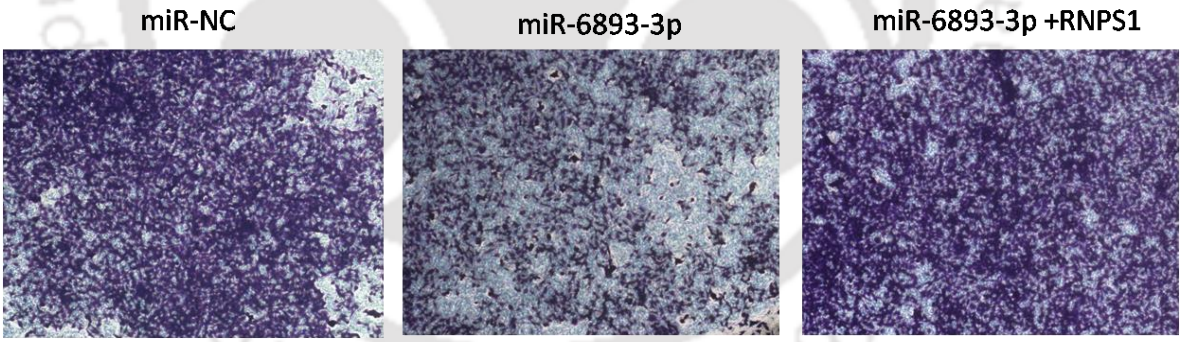


Figure 4.10: RNPS1 reverses the migration inhibiting effect of miR-6893-3p. Overexpression of miR-6893-3p decreases migration of HeLa cells and overexpression of RNPS1 could reverse the inhibitory effect of miR-6893-3p. HeLa cell migration abilities were determined by transwell assays after transfection with the indicated vectors.

4.4 Discussion

Despite the fact that roughly 70% of the human genome is actively transcribed, only 1-2% of genomic sequences encode proteins, whereas a major portion of the sequences possibly expresses non-coding RNA [193]. MicroRNA is one of the extensively studied

non-coding RNAs. Numerous genes are known to be regulated by miRNAs at the post-transcriptional level, and the abnormal expressions of miRNAs contribute to cancer initiation and development. In the present study, I show downregulation of miR-6893-3p expression in cervical cancer cells and miR-6893-3p expression was inversely correlated with RNPS1 level. The data establish miR-6893-3p as a novel negative regulator of RNPS1.

miR-6893 is a known negative regulator of S100A1 [194]. S100A1 is a calcium binding protein and is generally upregulated in malignant tumors. Overexpression of miR-6893 has been shown to reduce S100A1 level, whereas loss of miR-6893 caused upregulation of S100A1. Additionally, increased expression of miR-6893 or knockdown of S100A1 markedly reduced the migration and invasive abilities of HeLa cells compared to the control cells. Furthermore, miR-6893 overexpressing HeLa cells were found to be sensitive to cisplatin treatment compared to the control transfected HeLa cells. The viability of miR-6893-overexpressed cancer cells substantially reduced when treated with cisplatin. Moreover, increased expression of miR-6893 enhanced the apoptotic rate of HeLa cells upon treatment with cisplatin, thus miR-6893 is involved in the suppression of chemoresistance of HeLa cells [194].

In the present study, quantitative RT-PCR assay confirmed the downregulation of miR-6893-3p in cervical cancer cells compared to a normal cell. This is consistent with the results of the TCGA dataset analyses by Chen et al., 2019 [194]. In the previous chapter, I have found overexpression of RNPS1 in cervical cancer cells. The negative correlation of expression between miR-6893-3p and RNPS1 levels suggests a regulatory role for miR-6893-3p in RNPS1 expression. In this regard, miR-6893-3p plasmid was transfected into the HeLa cells to evaluate the effect of overexpression of miR-6893-3p on RNPS1 expression in HeLa cervical cancer cells. The results showed direct evidence that miR-6893-3p regulates RNPS1 expression both at the mRNA and protein level. In order to further validate that miR-6893-3p directly targets the 3’ UTR of RNPS1 mRNA, I generated luciferase expression plasmids containing the putative miR-6893-3p binding

region or mutated seed sequences in the 3' UTR of RNPS1 mRNA. The results demonstrated that the 3' UTR of RNPS1 is a target of miR-6893-3p due to the presence of 7 mer binding site.

Furthermore, transfection of miR-6893-3p decreased the migration ability of the HeLa cell line, while cotransfection of miR-6893-3p and plasmid overexpressing RNPS1 increased the migration potential of HeLa cells. Given that RNPS1 expression and migration are regulated by miR-6893-3p, it may be inferred that the migration potential of HeLa cells modulated by the expression of miR-6893-3p may be partially mediated by RNPS1. In summary, the study revealed a novel mechanism of how RNPS1 is aberrantly expressed in cervical cancer cells. It would be interesting to further examine whether miR-6893-3p-RNPS1 dysregulation represents a factor for malignant transformation.

*Note: Ayushi Rehman helped in performing the study in Figure 4.4 and 4.6.

Chapter 5: Identification of candidate RNA binding proteins associated with RNPS1 mRNA

The work embodied in this chapter is accepted for publication.

Deka B, Singh KK. Identification of candidate RNA binding proteins associated with RNPS1 3'UTR. (Book Chapter) Healthcare Research and Related Technologies - Proceedings of NERC 2022 (Springer Nature).

Abstract

The course of the journey of RNA-from RNA biogenesis to RNA degradation- is dictated by a vast number of proteins that interact with RNA to regulate its fate. The interaction between RNA and protein can be direct via RNA-binding proteins (RBPs) or other proteins, which function cooperatively within multisubunit ribonucleoprotein (RNP) complexes. As a result, identifying the RBP-RNA interaction is crucial for understanding how genes are regulated. The commonly used strategy is to isolate RNAs bound to known RBPs; however, recently, pull-down methods have been developed for the direct precipitation of specific RNA of interest and its associated proteins. In the previous chapter, I have shown that RNPS1 is subjected to post-transcriptional gene regulation via miRNAs. Herein, I performed a systematic approach for identifying the regulatory RBPs associated with *RNPS1* 3'UTR in the physiological context. The method is based on the RNA aptamer MS2 tag followed by affinity purification. The *RNPS1* 3'UTR is tagged with MS2 RNA hairpins and the chimeric RNA is then pulled-down by the simultaneous expression of a fusion protein MS2-Flag. Subsequent evaluation of the pulled down ribonucleoprotein (RNP) complex by mass-spectrometry led to the identification of a list of proteins that interact with *RNPS1* 3'UTR. These proteins plausibly regulate the fate of *RNPS1* mRNA, including transcription, splicing, nuclear export, localization, translation and decay.

5.1 Introduction

The coordinated, multi-layered process of gene expression involves a large number of trans-acting factors. While most genomic studies have focused on transcriptional regulation, post-transcriptional regulation has only lately been studied. RNA-binding proteins (RBPs) and non-coding RNAs are two main players in post-transcriptional regulation that act in a combinatorial or cooperative manner to control the spatio-temporal expression of target proteins. Computational, genetic, and biochemical strategies have been employed to identify and confirm RBP-binding sites on mRNAs and their corresponding RBPs. However, traditional approaches identify only a limited number of RNA-protein interactions. Therefore, exploring the binding characteristics of RBPs at a systems level is essential to comprehensively understand the regulatory mechanisms mediated by RBPs and determine the "RNP code" of an mRNA.

RNP immunoprecipitation (RIP-chip) was the first developed high-throughput method to detect *in vivo* targets of a given RBP using cDNA arrays [195]. RIP is performed under endogenous conditions that preserve RNA-protein interactions. RIP studies have unraveled several putative regulatory elements and RNP components in response to different stimuli. Further, this method provided evidence for the operon/regulon model, in which RBPs synchronize the expression of functionally related proteins [196, 197].

Subsequent high-throughput methods employed cross-linking followed by immunoprecipitation to precisely identify the position of the binding sites of RBPs within mRNA. One of these techniques is PAR-CLIP (photoactivatable ribonucleoside cross-linking and immunoprecipitation). PAR-CLIP employs a long UV wave (365 nm) to cross-link photoactive thiouridine incorporated into RNA [198]. This cross-linking does not induce extensive photo damage to nucleic acid and protein compared to a short UV wave (254 nm). These techniques have been developed to isolate RNAs from particular RNA binding proteins *in vivo*. However, developing appropriate strategies to identify and explore proteins that bind to a specific RNA is challenging.

In previous studies, RNA of interest is chemically modified during *in vitro* transcription to study the interactome of a specific RNA. One of the widely used labels in *in-vitro* studies is the incorporation of biotin within the desired RNA. The tagged RNA is then incubated with cell lysates to allow RBPs to bind RNA, followed by RNA pulldown using streptavidin-coated beads and detection of bound proteins [199]. In order to study RNA-protein interactions within cells or under endogenous conditions, chemically labeled RNAs are expressed in cells. Subsequently, the labeled RNA is isolated from the RNA pool using specific antibodies or affinity compounds such as bromouride. However, the challenges of these approaches are ensuring proper localization of chemically altered transcripts and avoiding nonspecific association with proteins. To this end, alternate strategies have been developed that contain tagging and pulldown of exogenously expressed transcripts. Various small RNA tags, also known as aptamers, are tagged to the RNA of interest and pulldown using molecules that recognize the aptamer. For instance, RNA aptamers S1 and D8 have an affinity for Sephadex and streptavidin, respectively [200-202]. The S1 aptamer was employed to analyze the bound proteins of RNase P by tagging RPR1 RNA, the large subunit of RNase P. The aptamer tag showed that RNaseP occurs as a monomer and is loosely associated with RNase MRP [200].

Although RNPS1 plays an important role in splicing activity, there are no reports on regulatory proteins controlling the expression of RNPS1. To study the RBP-mediated regulation of RNPS1 expression, I developed a method to characterize the interaction between *RNPS1* 3'UTR and proteins. The method utilizes the RNA aptamer, MS2 tag, one of the most widely used tags. It is a viral RNA sequence of 19 nt length that folds into a hairpin loop structure. This loop structure is recognized by the MS2 bacteriophage coat RNA-binding protein with very high specificity and affinity. The MS2 is added to *RNPS1* 3'UTR, followed by co-expression of the MS2-tagged *RNPS1* RNA along with the MS2 coat protein fused to a Flag tag. The MS2 RNP complex is purified using an antibody against Flag, and finally, the proteins bound to *RNPS1* 3'UTR were analyzed via mass spectrometry. Intriguingly, *RNPS1* 3'UTR was found to interact with several RBPs,

including heterogeneous nuclear ribonucleoprotein C (HNRNPC), interleukin enhancer-binding factor 2 (ILF2) and ILF3.

5.2 Materials and methods

5.2.1 Construction of plasmids

The plasmid vector pcDNA5-FRT-TO was a kind gift from Prof. Niels Gehring, University of Cologne, Germany. First, dtTomato was cloned in pcDNA5-FRT-TO using NheI and XhoI restriction sites. Second, 6X MS2 aptamer sequence were cloned in pcDNA5-FRT-TO-dtTomato using XhoI and SalI restriction sites. Finally, *RNPS1* 3'UTR was cloned downstream of MS2 aptamer in pcDNA5-FRT-TO-dtTomato-MS2 using SalI and NotI restriction sites. Cloned gene sequences were further validated by Sanger sequencing.

The plasmid vector PB-Cuo-MCS-IRES-GFP-EF1α-CymR-Puro-Flag-MS2 was a kind gift from Prof. Niels Gehring, University of Cologne, Germany.

5.2.2 Transformation of recombinant vector

For transformtion, please refer to Chapter 4, section 4.2.2.

5.2.3 Cell culture

For cell culture, please refer to Chapter 3, section 3.2.1.

5.2.4 Generation of stable cells

Flp-In T-REx-293 cells were seeded in 6-well plates. 24 hours after seeding, the cells were transfected with pcDNA5-FRT-TO-dtTomato-MS2-RNPS1 3'UTR and pOG44 (FLP recombinase). 48 hours after transfection, the cells were splitted 1:3 into new 6-well plates containing fresh medium supplemented with 150 µg/ml hygromycin B. Selection in hygromycin-containing medium was continued for approximately 3 weeks until single colonies were obtained. Single colonies were picked and expanded. These single stable cells were again seeded in 6-well plates. 24 hours after seeding, the cells were transfected with PB-Cuo-MCS-IRES-GFP-EF1α-CymR-Puro-FLAG-MS2 and Piggybac

transposase vector. 72hrs after transfection, the cells were splitted 1:3 into new 6-well plates containing fresh medium supplemented with 2µg/ml puromycin. Selection in puromycin-containing medium was continued for approximately 2 weeks until single resistant colonies became visible. Single colonies were picked and expanded to get double stable cell lines.

5.2.5 RNA isolation

Please refer to section Chapter 3, section 3.2.3.

5.2.6 cDNA synthesis

For cDNA synthesis of mRNA, please refer to Chapter 3, section 3.2.5.

5.2.7 Quantitative real-time PCR

For mRNA-specific qPCR, please refer to Chapter 3, section 3.2.7.

I calculated fold enrichment of each RIP reaction from qPCR data.

ΔCt [Normalized RIP] = Ct [IP] − Ct [Input]

ΔΔCt [RIP/Control] = ΔCt [Normalized RIP] − ΔCt [Normalized control]

Fold enrichment = $2^{(-\Delta\Delta Ct\ [RIP/Control])}$

5.2.8 Preparation of cell lysates

Double stable Flp-In T-REx-293 cells were at first induced with tetracycline (final conc. 1ug/ul) and next day, the cells were induced with cumate (final conc. 300ug/ml) and tetracycline. After 48 hours, the cells were washed twice with DPBS (Dulbecco's phosphate-buffered saline) and scraped from the cell culture dish in EJC buffer (20 mM HEPES-KOH (pH 7.9), 200 mM NaCl, 2 mM MgCl2, 0.2% Triton X-100, 0.1% Nonidet P40, 0.05% Na-deoxycholic acid). Lysed cells were transferred to 1.5 ml eppendorf tubes on ice followed by sonication (15 cycles at 25% amplitude; 1 second ON, 4 seconds OFF).

The lysate was centrifuged at 13000 rpm for 10 min at 4°C and the supernatant was transferred to a new tube.

5.2.9 RNP Immunoprecipitation

The RNP immunoprecipitation was performed by incubating cell lysate with anti-FLAG-M2 magnetic beads (Sigma-Aldrich) and incubated 5 hours in an overhead shaker at 4°C. After incubation, beads were washed four times with EJC buffer. After the last washing step, FLAG peptides were used to elute the bound protein.

5.2.10 Silver staining

The SDS-PAGE gel was prepared using BIO-RAD mini gel cast system. The input and RNP IP samples were resolved on 10% SDS-PAGE gel in a vertical electrophoresis tank. After SDS-PAGE, the gel was shortly rinsed in water. Fixation solution (30% ethanol, 10% acetic acid, H2O) was added on the gel and incubated for 2 h with gentle shaking. The gel was then washed with gentle shaking with 30% ethanol 3x 10 min followed by washing gel with dest. water 2x 10 min. The gel was incubated with sensitizer solution (Sodium dithionite 25 mg/100 ml) for 1 min followed by washing gel with dest. water 2x 1min. The staining solution (0.2% AgNO3, 3 µl/40 ml formaldehyde solution (36-38%)) was added on the gel and incubated for 25 min. The gel was washed with dest. water for 1 min and developer solution (6% sodium carbonate, 4 µg/ml sodium thiosulfate, formaldehyde) was added and incubated for 2-3 min with gentle mixing. Finally stop solution (4% (w/v) Tris and 2% (v/v) acetic acid) was added before the protein bands get dark stained.

5.2.11 Mass spectrometry

Please refer to section Chapter 2, section 2.2.8.

5.3 Results

5.3.1 Construction of double stable cells for interactome capture of RNPS1 3'UTR

Besides miRNAs, RBPs are other essential regulators of post-transcriptional gene regulation. Therefore, to isolate RBPs associated with *RNPS1* 3'UTR double stable cell lines were generated. The stable cell expresses both MS2-RNPS1 3'UTR and FLAG-tagged MS2 coat protein. Double stable cells were generated using two types of systems: Flp-In T-REx system and PiggyBac transposon system. The Flp-In TREx system is designed to generate cell lines that stably express the gene of interest in an isogenic and inducible manner. As a result, this eliminates any variation in expression levels caused by the integration of the gene of interest into various sites of the chromosome. It also allows cells to be cultured without expressing the gene of interest, thereby averting probable harmful effects on cell growth. Furthermore, the expression level can be regulated to ensure expression of the desired gene at the near-physiological level, thus avoiding false RNA-protein associations resulting from over-expression. The PiggyBac transposon system is an efficient non-viral vector system for stable expression of the gene of interest. This system utilizes the transposon-mediated integration mechanism to "cut and paste" the desired gene into the genome of mammalian cells. It takes advantage of cabbage looper moth-derived PiggyBac transposon and is one of the most efficient transposons for modifying the mammalian genome. Moreover, the expression of proteins can be switched on or induced by adding small molecule cumate to the cells.

A schematic outline of the generation of the double stable cell procedure used in this study is shown in Fig. 5.1. In order to construct MS2-RNPS1 3'UTR expression plasmid, I have used the vector pcDNA5/FRT/TO. It is a 5.1 kb expression plasmid designed for use with the Flp-In T-REx system. It contains FRT (Flp Recombination Target) site linked to the hygromycin resistance gene for Flp recombinase-mediated integration into the mammalian cells. Flp-In T-REx-293 cells were cotransfected with a vector expressing

tdTomato-MS2-RNPS1 3'UTR and pOG44 plasmid (Figure 5.1). Upon cotransfection, the Flp recombinase enzyme produced by pOG44 initiates a homologous recombination event between the FRT sites integrated into the genome and pcDNA5/FRT/TO such that the gene of interest and hygromycin resistance gene are inserted at the FRT site. The stable cells were selected using hygromycin and the expression of chimeric RNA, tdTomato-MS2-RNPS1 3'UTR, was validated via RT-PCR (Figure 5.2). These stable cells were further cotransfected with PiggyBac transposon vector containing FLAG-tagged MS2 coat protein and PiggyBac transposase (Figure 5.1). Upon transfection, transposase

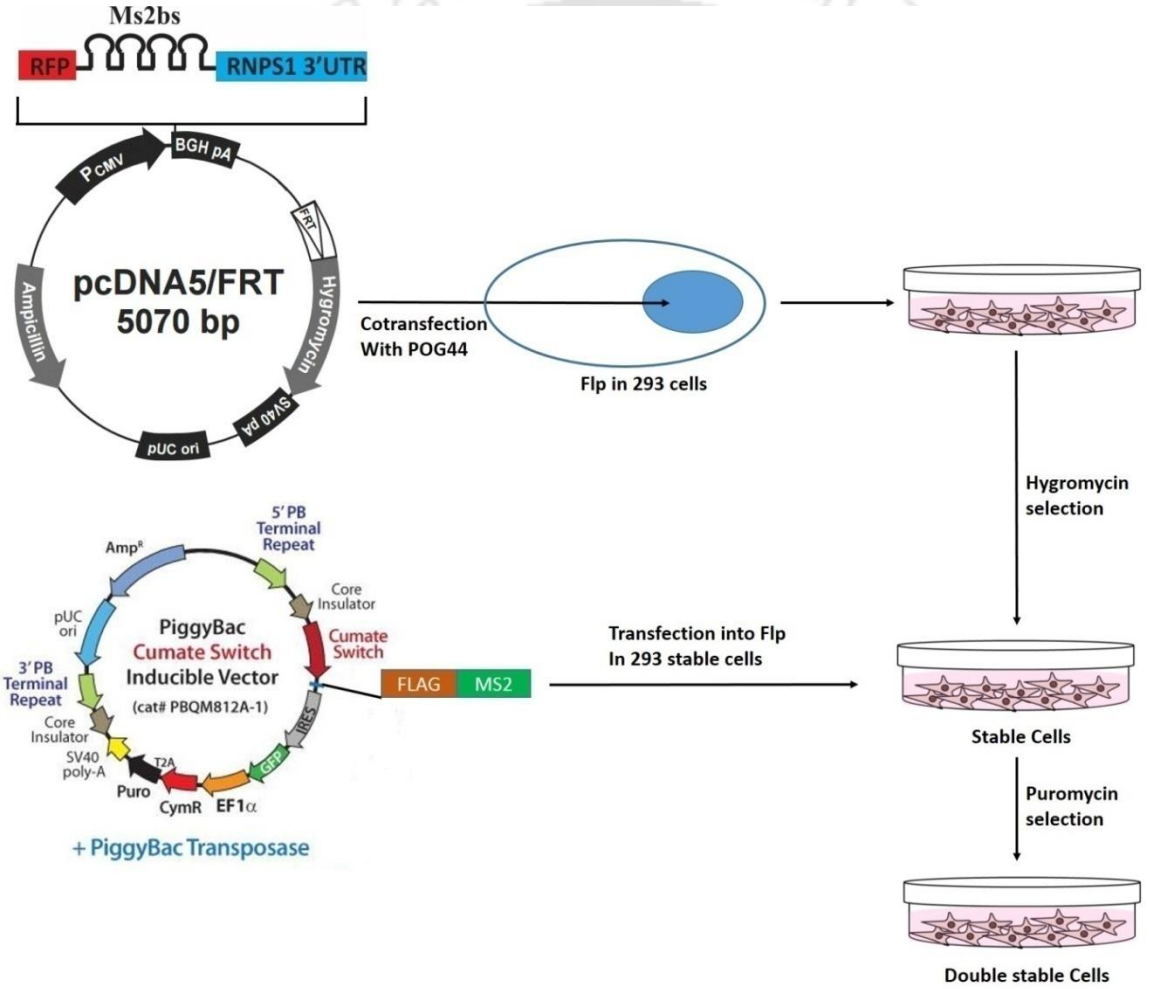


Figure 5.1: Schematic representation of generation of double stable cells. Flp-In T-REx-293 cells expressing simultaneously tdTomato-MS2bs-RNPS1 3'UTR and FLAG-MS2cp. MS2bs stands for MS2 binding site. Flp-In T-REx-293 cells were cotransfected with a plasmid expressing tdTomato-MS2-RNPS1 3'UTR and pOG44 plasmid. The stable cells were selected using hygromycin. These stable cells were further cotransfected with PiggyBac transposon vector containing FLAG-tagged MS2 coat protein and PiggyBac transposase. The stable cells were selected via puromycin resistance.

Identifies the inverted terminal repeat sequences (ITRs) on the transposon vector and successfully transfers the gene of interest from the vector into the TTAA chromosomal locations. The stable cells were selected via puromycin resistance and finally, the double stable cells express both tdTomato-RNPS1 3'UTR MS2 and FLAG-tagged MS2 coat protein. The negative control stable cells express tdTomato-MS2 and FLAG-tagged MS2 coat protein.

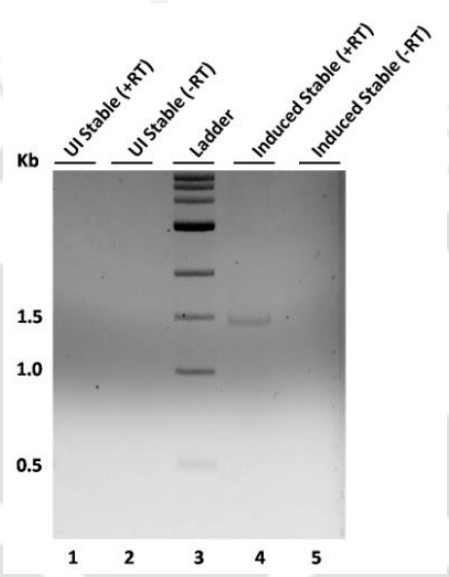


Figure 5.2: Validation of stable cells. Stable Flp-In T-REx-293 cells were induced with 1ug/ml of tetracycline. Stable cells were validated via RT-PCR using td tomato forward primer and *RNPS1* 3'UTR reverse primer (Lane 4). Lane 2 and 5 are non-RT control. UI stands for uninduced stable cells.

5.3.2 MS2 tagged RNA affinity capture of RBPs

Double stable cells were induced with doxycycline and cumate to express tdTomato-RNPS1 3'UTR-MS2 and FLAG-tagged MS2 coat protein, respectively. A schematic outline of the immunoprecipitation–mass spectrometry procedure used in this study is shown in Fig 5.3. Next, lysates were prepared from the double stable cells for RNP-IP and the RBPs that associate with *RNPS1* 3'UTR in vivo were then captured by affinity purification using magnetic beads coated anti-FLAG antibody. A part of the IP complexes was used to isolate total mRNA transcripts pulled down, followed by qRT-PCR to determine the amount of *RNPS1* 3'UTR enriched in the RNP-IP assay. The other part of the IP was analyzed on an SDS-PAGE gel and bands were excised for mass spectrometry analysis. A

negative control RNP-IP was performed on stable cells expressing chimeric RNA tdTomato-MS2 without *RNPS1* 3'UTR. The qRT-PCR result revealed that *RNPS1* 3'UTR was dramatically enriched in the pulldown of stable cells expressing tdTomato-MS2bs-RNPS1 3'UTR- compared to negative control stable cells (Figure 5.4B). Silver staining of the gel loaded with *RNPS1* 3'UTR-IP identified several bands (Figure 5.4A). These data demonstrate that chimeric *RNPS1* 3'UTR can be efficiently and specifically immunoprecipitated from cell extracts.

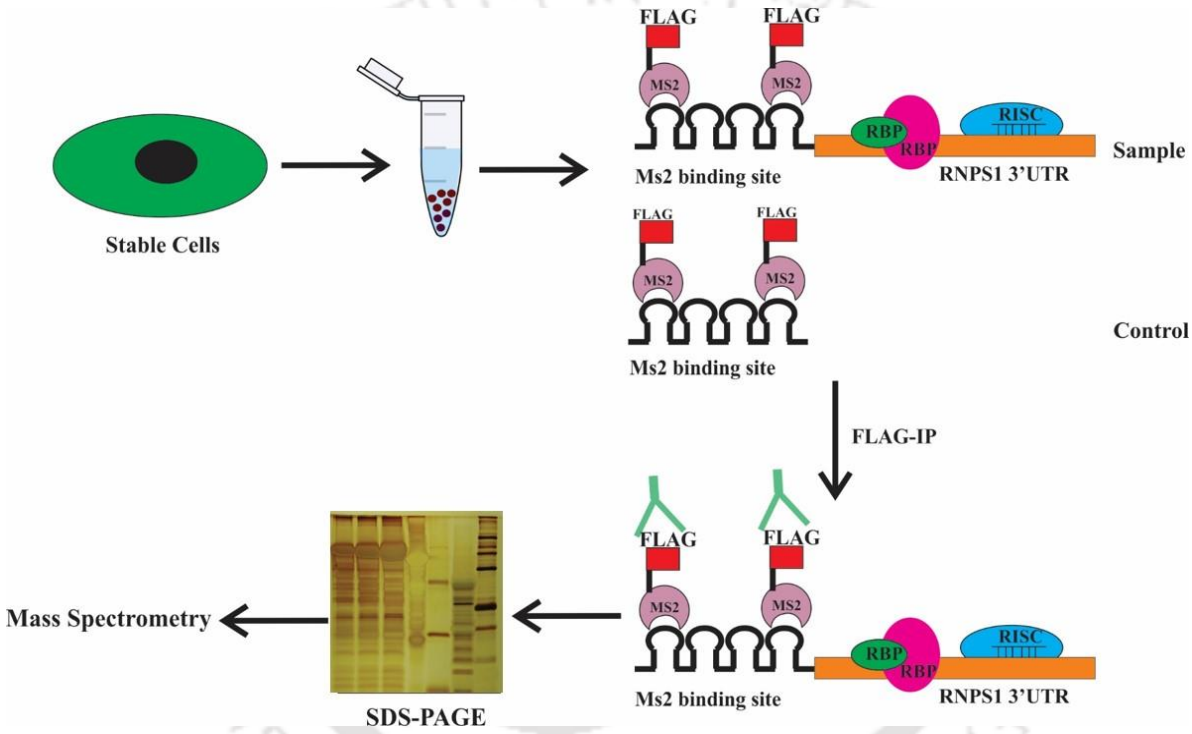


Figure 5.3: Schematic representation of MS2 tagged RNA affinity capture of proteins bound to *RNPS1* 3'UTR. *RNPS1* 3'UTR stable cells express tdTomato-MS2-RNPS1 3'UTR and FLAG-MS2cp, whereas control cells express tdTomato-MS2 and FLAG-MS2cp. Lysates were prepared from the double stable cells for RNP-IP and the RBPs that associate with RNPS1 3'UTR in vivo were then captured by affinity purification using magnetic beads coated anti-FLAG antibody. The IP was analyzed on an SDS-PAGE gel and bands were excised for mass spectrometry analysis.

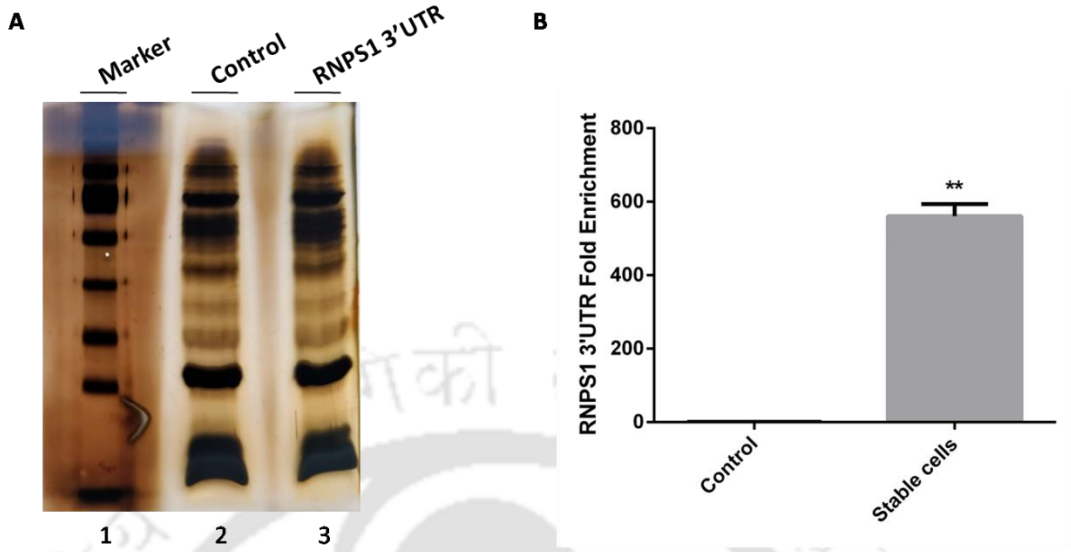


Figure 5.4: Identification of *RNPS1* 3'UTR binding proteins by MS2-tagged RNA affinity purification. (A) Immunoprecipitation complexes from control and *RNPS1* 3'UTR stable cells were separated on 10% SDS-PAGE and the PAGE gel was stained with silver staining. Control cells express tdTomato-MS2 and FLAG-MS2cp, whereas *RNPS1* 3'UTR stable cells express tdTomato-MS2-RNPS1 3'UTR and FLAG-MS2cp (B) The enrichment of *RNPS1* 3'UTR in samples obtained after FLAG RNP IP was measured by qRT-PCR. The values represent mean±SEM (n=3). **P<0.01.

5.3.3 Analysis of candidate RNPS1 3'UTR binding proteins

Mass spectrometry analysis identified 148 unique proteins with a protein False Discovery Rate lower than 1% and log2 fold change ≥ 0.5 (Table 5.1). To determine the potential biological pathways associated with *RNPS1* 3'UTR interacting proteins, I performed enrichment analysis in the Gene Ontology (GO) domain "Biological Pathway" using FunRich tool. The study unraveled that *RNPS1* 3'UTR binding proteins were most significantly enriched in "Gene expression" processes (Figure 5.5). Further, "3' UTR-mediated translational regulation" is also one of the predominant processes in the GO analysis. These findings suggest that RNPS1 expression is probably regulated by many key regulators via its 3'UTR. Interestingly, proteins associated with *RNPS1* 3'UTR are also involved in "Influenza viral RNA transcription and elongation" processes (Figure 5.5). Along the same line, a previous meta-analysis of genome-wide studies has identified RNPS1 as one of the host cell factors essential for influenza virus infection [203].

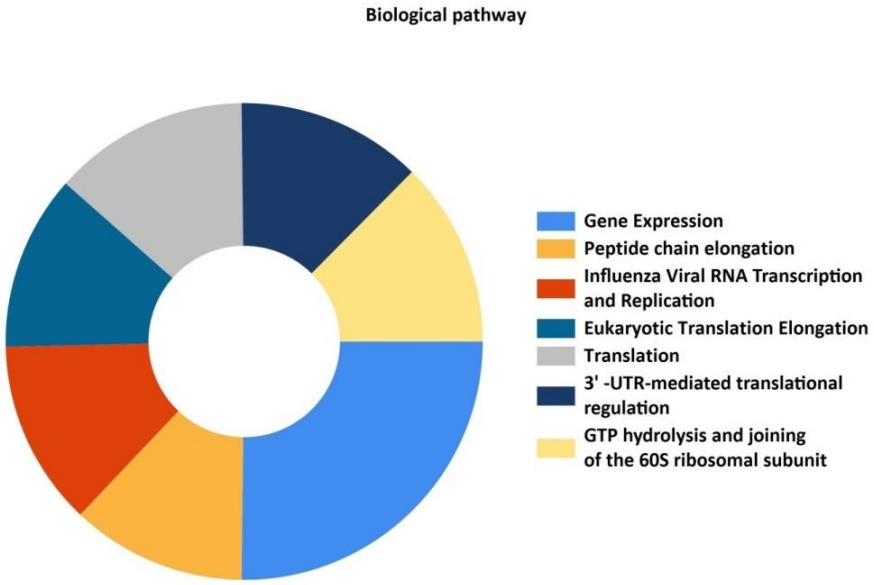


Figure 5.5: Biological pathway category of RNPS1 binding proteins in Gene Ontology analysis by FunRich. The RNPS1 binding proteins were most highly enriched in gene expression (35.6%) followed by translation (18.8%) and 3'UTR-mediated translational regulation (17.8%). A significant number of proteins were also enriched in Influenza Viral RNA Transcription and Replication (17.8%).

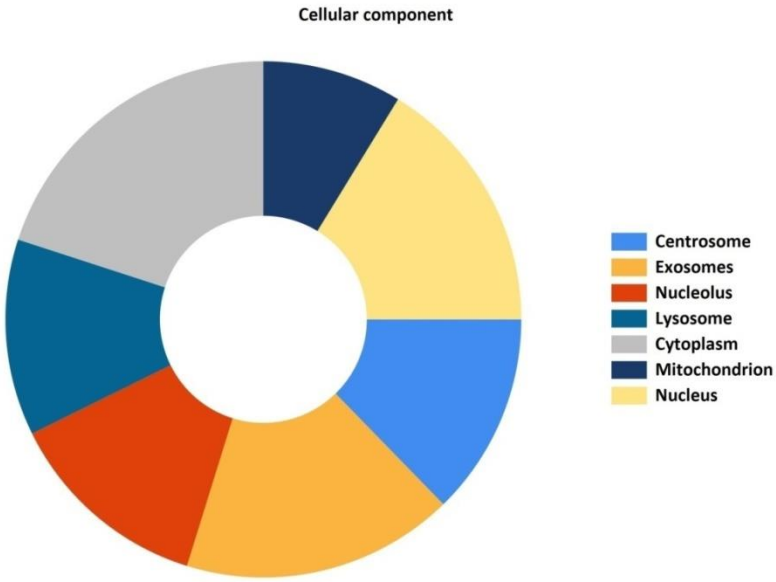


Figure 5.6: Cellular component category of RNPS1 binding proteins in Gene Ontology analysis by FunRich. The RNPS1 binding proteins were most highly enriched in cytoplasm (72.2%), followed by exosome (62.1%) and nucleus (59%). A significant number of proteins were also enriched in centrosome (46.2%) and nucleolus (46.9%).

Additionally, I performed enrichment analysis in the Gene Ontology (GO) domain "Cellular component". RNPS1 mRNA binding proteins were mainly distributed in the cytoplasm, exosome and nucleus (Figure 5.6). Moreover, GO analysis unraveled that candidate RNPS1 binding proteins were significantly enriched in the centrosome, lysosome and mitochondrion compartments. These analyses suggest RNPS1 transcript binding proteins comprise key transport regulators and indicate possible cellular compartments in which RNPS1 transcript localizes.

5.3.4 HNRNPC interacts with RNPS1 mRNA

Among all candidate binding proteins, a few proteins were selected based on their functions in RNA metabolism. One of the candidate proteins is HNRNPC, albeit the log2 fold change of HNRNPC was below 0.5. HNRNPC is a well-known RBP with functions in RNA splicing, RNA export, RNA stability, 3'end processing and translation. To confirm whether HNRNPC is indeed an RNPS1 binding factor, I immunoprecipitated FLAG-tagged HNRNPC from HEK 293 cell lysates and analyzed its interaction with *RNPS1* mRNA. Flag-HNRNPC was overexpressed in HEK293 cells and RNA-IP was performed from whole-cell lysates using anti-FLAG magnetic beads. qRT-PCR result shows that endogenous *RNPS1* mRNA was significantly enriched in FLAG-HNRNPC IP, thereby implying that HNRNPC associates with *RNPS1* mRNA (Figure 5.7). It was previously demonstrated through RNA-IP that HNRNPC associates with *APP* mRNA (Amyloid precursor protein). As expected, the APP mRNA was also detected in HNRNPC RNA IP in the current study (Figure 5.7). I also found many putative HNRNPC binding sites in the RNPS1 3'UTR by 'RBPmap', this suggests a possible direct interaction between *RNPS1* mRNA and HNRNPC (Figure 5.8). It is intriguing to speculate that the binding of HNRNPC to *RNPS1* mRNA probably modulates the stability, splicing or translation of *RNPS1* mRNA. Taken together, these data show the reliability of the IP-MS strategy and confirm association of HNRNPC with *RNPS1* mRNA. The analysis provides proof of concept that MS2-mediated pulldown of *RNPS1* 3'UTR could be used to screen proteins interacting with *RNPS1* transcript in a physiological setting.

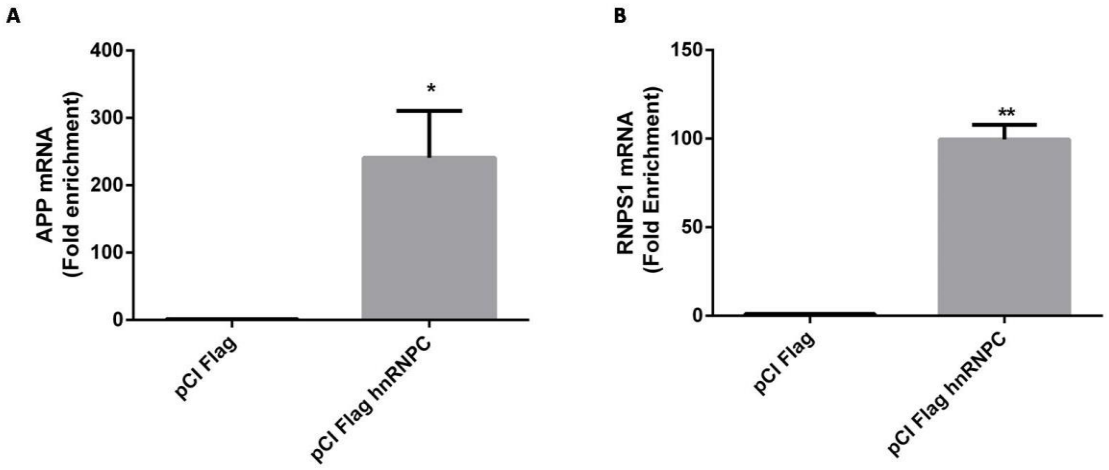


Figure 5.7: Verification of candidate *RNPS1* 3'UTR binding protein. RNA immunoprecipitation was performed with HEK293 cell lysates transfected with pCI-FLAG or pCI-FLAG-HNRNPC using FLAG antibody to enrich RNA bound to HNRNPC. The presence of APP and RNPS1 was detected by qPCR. APP mRNA served as positive control. *P<0.05, **P<0.01.

Position	Motif	Occurrence
49	huuuuuk	cuguaacuuauaccccacccagcucaguuuuugucacuuuucuagccaaaggaagacc
50	huuuuuk	uguaacuuauaccccacccagcucaguuuuugucacuuuucuagccaaaggaagacca
51	huuuuuk	guaacuuauaccccacccagcucaguuuuugucacuuuucuagccaaaggaagaccag
52	uuuuu	uaacuuauaccccacccagcucaguuuuugucacuuuucuagccaaaggaagacca
58	huuuuuk	auaccccacccagcucaguuuuugucacuuuuucuagccaaaggaagaccaguaggaaa
59	huuuuuk	uaccccacccagcucaguuuuugucacuuuuucuagccaaaggaagaccaguaggaaag
60	uuuuu	accccacccagcucaguuuuugucacuuuuucuagccaaaggaagaccaguaggaaa
61	uuuuu	ccccacccagcucaguuuuugucacuuuucuagccaaaggaagaccaguaggaaag
685	huuuuuk	gguuaucuuggcaacauguacauuguuuuuauuuuuuuucuuuuuugcuuucauugu
686	huuuuuk	guuaucuuggcaacauguacauuguuuuuauuuuuuuucuuuuuugcuuucauugua
687	huuuuuk	uuaucuuggcaacauguacauuguuuuauuuuuuuucuuuuuugcuuucauuguac
688	huuuuuk	uaucuuggcaacauguacauuguuuuauuuuuuuucuuuuuugcuuucauuguaca
689	huuuuuk	aucuuggcaacauguacauuguuuuauuuuuuuucuuuuuugcuuucauuguacag
690	huuuuuk	ucuuggcaacauguacauuguuuuauuuuuuuucuuuuuugcuuucauuguacagu
691	huuuuuk	cuuggcaacauguacauuguuuuauuuuuuuucuuuuuugcuuucauuguacaguc
692	huuuuuk	uuggcaacauguacauuguuuuauuuuuuuucuuuuuugcuuucauuguacaguca
693	huuuuuk	uggcaacauguacauuguuuuauuuuuuuucuuuuuugcuuucauuguacagucag
694	uuuuu	ggcaacauguacauuguuuuauuuuuuuucuuuuuugcuuucauuguacaguca
695	huuuuuk	gcaacauguacauuguuuuauuuuuuuucuuuuuugcuuucauuguacagucagua
696	huuuuuk	caacauguacauuguuuuauuuuuuuucuuuuuugcuuucauuguacagucaguac
697	huuuuuk	aacauguacauuguuuuauuuuuuuucuuuuuugcuuucauuguacagucaguacu
698	huuuuuk	acauguacauuguuuuauuuuuuuucuuuuuugcuuucauuguacagucaguacua
699	huuuuuk	cauguacauuguuuuauuuuuuuucuuuuuugcuuucauuguacagucaguacuau
700	huuuuuk	auguacauuguuuuauuuuuuuucuuuuuugcuuucauuguacagucaguacuaua
701	uuuuu	uguacauuguuuuauuuuuuuucuuuuuugcuuucauuguacagucaguacuau
702	uuuuu	guacauuguuuuauuuuuuuucuuuuuugcuuucauuguacagucaguacuaua
734	huuuuuk	uucauuguacagucaguacuauaaaauuucucuuuugaguuuuauaccuuuguagca
735	uuuuu	ucauuguacagucaguacuauaaaauuucucuuuugaguuuuauaccuuuguagc
738	huuuuuk	uuguacagucaguacuauaaaauuucucuuuugaguuuuauaccuuuguagcauuuu
739	huuuuuk	uguacagucaguacuauaaaauuucucuuuugaguuuuauaccuuuguagcauuuua
740	uuuuu	guacagucaguacuauaaaauuucucuuuugaguuuuauaccuuuguagcauuuu
741	uuuuu	uacagucaguacuauaaaauuucucuuuugaguuuuauaccuuuguagcauuuua
746	huuuuuk	ucaguacuauaaaauuucucuuuugaguuuuauaccuuuguagcauuuuagaugaca

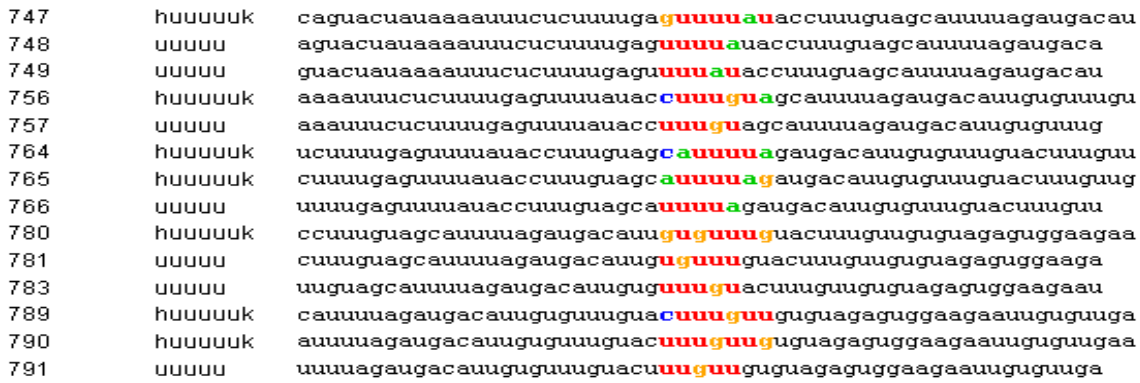


Figure 5.8: Putative binding sites of HNRNPC within RNPS1 3'UTR predicted via RBPmap. Position refers to the starting position of the HNRNPC binding site in the input sequence (RNPS1 3'UTR). The motif refers to the motif that is mapped to the query sequence (RNPS1 3'UTR). Occurrence displays the presence of the motif in the query sequence.

Table 5.1: List of candidate *RNPS1* 3'UTR binding proteins from MS analysis.

Gene Name	Description	Score	Log2 FC	Queries Matched
GPI	Glucose-6-phosphate isomerase	207.36	3.2192	15
KLHL41	Kelch-like protein 41	70.799	3.0059	11
CSTA	Cystatin-A	27.527	2.7133	2
PGK1	Phosphoglycerate kinase 1	106.21	2.6164	12
RPL13A	60S ribosomal protein L13a	46.911	1.8461	5
CKB	Creatine kinase B-type	323.31	1.8057	16
RPL4	60S ribosomal protein L4	33.437	1.7826	8
FKBP4	Peptidyl-prolyl cis-trans isomerase	92.897	1.7727	17
PGAM1	Phosphoglycerate mutase 1	162.54	1.6863	9
RPL22	60S ribosomal protein L22	6.4384	1.6826	2
BAG5	BAG family molecular chaperone regulator 5	17.259	1.6508	5
GDI2	Rab GDP dissociation inhibitor beta	73.143	1.635	14
ALDH9A1	4-trimethylaminobutyraldehyde dehydrogenase	6.6876	1.5634	4
RPL27	60S ribosomal protein L27	8.4243	1.5345	4
CALR	Calreticulin	264.02	1.5323	13
ENO1	Alpha-enolase	211.21	1.5011	14
U2AF1	Splicing factor U2AF 35 kDa subunit	32.904	1.4892	3
ILF2	Interleukin enhancer-binding factor 2	55.794	1.4484	8
PGD	6-phosphogluconate dehydrogenase	136.11	1.4297	12
PDIA4	Protein disulfide-isomerase A4	111.48	1.4063	14

ADSS	Adenylosuccinate synthetase isozyme 2	12.991	1.3858	4
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase	312.96	1.3849	14
LDHA	L-lactate dehydrogenase A chain	59.569	1.3834	10
MTHFD1	C-1-tetrahydrofolate synthase	105.12	1.3396	24
TARS	Threonine--tRNA ligase	60.083	1.3336	15
EIF4A3	Eukaryotic initiation factor 4A-III	24.186	1.3172	11
AGPAT1	1-acyl-sn-glycerol-3-phosphate acyltransferase alpha	28.586	1.2838	1
TKT	Transketolase	124.95	1.2805	16
PPP1CC	Serine/threonine-protein phosphatase	52.844	1.28	7
LDHB	L-lactate dehydrogenase B chain	129.09	1.2766	9
EIF4A1	Eukaryotic initiation factor 4A-I	57.255	1.2588	15
GANAB	Neutral alpha-glucosidase AB	102.1	1.2567	18
ST13	Hsc70-interacting protein	27.435	1.2559	5
ETFA	Electron transfer flavoprotein subunit alpha	22.525	1.1714	5
RPL5	60S ribosomal protein L5	42.995	1.1521	5
LCP1	Plastin-2	28.172	1.1463	11
DLD	Dihydrolipoyl dehydrogenase	13.544	1.1433	5
VDAC1	Voltage-dependent anion-selective channel protein 1	116.19	1.1354	6
API5	Apoptosis inhibitor 5	14.26	1.1313	5
S100A9	Protein S100-A9	47.29	1.1223	3
TPI1	Triosephosphate isomerase	149.9	1.1057	12
RPL23A	60S ribosomal protein L23a	21.525	1.0753	3
HSP90B1	Endoplasmin	238.74	1.0687	23
PKM	Pyruvate kinase PKM	323.31	1.0601	24
IGF2BP1	Insulin-like growth factor 2 mRNA-binding protein 1	26.15	1.0554	9
ALDOA	Fructose-bisphosphate aldolase	110.38	1.0439	11
NCL	Nucleolin	123.84	1.0381	16
SND1	Staphylococcal nuclease domain-containing protein 1	41.573	1.0336	11
ACTN4	Alpha-actinin-4	168.56	1.02	26
NSUN2	tRNA (cytosine(34)-C(5))-methyltransferase	44.164	1.0196	10
JUP	Junction plakoglobin	24.405	1.0052	9
DDX39B	Spliceosome RNA helicase DDX39B	75.249	0.9998	10
CACYBP	Calcyclin-binding protein	111.81	0.9959	9
PFKP	ATP-dependent 6-phosphofructokinase	114.28	0.9941	13
RPLP0	60S acidic ribosomal protein P0	34.394	0.9924	7

RPL26	60S ribosomal protein L26	8.4212	0.9839	4
ANXA6	Annexin A6	64.566	0.9746	10
TRAP1	Heat shock protein 75 kDa	73.977	0.9735	15
APEX1	DNA-(apurinic or apyrimidinic site) lyase	80.274	0.9715	7
HNRNPL	Heterogeneous nuclear ribonucleoprotein L	136.33	0.9623	11
YWHAZ	14-3-3 protein zeta/delta	197.64	0.9549	11
RPL27A	60S ribosomal protein L27a	17.498	0.9431	3
TUFM	Elongation factor Tu	52.299	0.9407	10
HNRNPA3	Heterogeneous nuclear ribonucleoprotein A3	34.951	0.9402	9
FSCN1	Fascin	41.679	0.9394	8
RPL15	60S ribosomal protein L15;Ribosomal protein L15	26.188	0.9363	9
HSP90AB2P	Putative heat shock protein HSP 90-beta 2	50.684	0.9149	8
YWHAB	14-3-3 protein beta/alpha	242.06	0.9081	9
RPL11	60S ribosomal protein L11	16.493	0.893	3
RAB1A	Ras-related protein Rab-1A	8.2647	0.8849	5
CCDC97	Coiled-coil domain-containing protein 97	14.59	0.8801	4
U2AF2	Splicing factor U2AF 65 kDa subunit	42.127	0.8624	6
BSG	Basigin	6.3922	0.8565	2
YWHAH	14-3-3 protein eta	30.372	0.8555	5
ACACA	Acetyl-CoA carboxylase 1	23.434	0.8544	5
MDH2	Malate dehydrogenase	58.18	0.8508	10
HNRNPA1	Heterogeneous nuclear ribonucleoprotein A1	145.23	0.8335	12
PSAT1	Phosphoserine aminotransferase	28.531	0.8327	9
BAT3;BAG6	Large proline-rich protein BAG6	18.689	0.8292	5
PTBP1	Polypyrimidine tract-binding protein 1	77.442	0.8234	6
MAGED2	Melanoma-associated antigen D2	22.616	0.8204	2
HNRNPH1	Heterogeneous nuclear ribonucleoprotein H	230.37	0.811	14
KIAA1033	WASH complex subunit 7	6.1897	0.8087	3
RPL3	60S ribosomal protein L3	61.735	0.808	11
EIF2S3	Eukaryotic translation initiation factor 2 subunit 3	77.706	0.8073	8
PRKCSH	Glucosidase 2 subunit beta	168.56	0.8037	10
DHX9	ATP-dependent RNA helicase A	180.95	0.8032	24
DDX6	Probable ATP-dependent RNA helicase DDX6	20.613	0.7988	5
AHSA1	Activator of 90 kDa heat shock protein ATPase homolog 1	32.245	0.7958	6
G3BP1	Ras GTPase-activating protein-binding protein 1	56.906	0.7668	8

DLAT	Dihydrolipoyllysine-residue acetyltransferase	14.174	0.7421	7
AHCY	Adenosylhomocysteinase	85.067	0.7397	13
RPL35A	60S ribosomal protein L35a	4.768	0.7314	4
RUVBL1	RuvB-like 1	230.86	0.7299	14
CSE1L	Exportin-2	154.71	0.7283	19
HSP90AA1	Heat shock protein HSP 90-alpha	323.31	0.7162	31
RAB11A	Ras-related protein Rab-11A	15.335	0.7158	4
ANP32A	Acidic leucine-rich nuclear phosphoprotein 32 A	79.613	0.7103	8
CTPS1	CTP synthase 1	92.619	0.7015	9
ARHGEF10	Rho guanine nucleotide exchange factor 10	107.87	0.6874	17
HNRNPR	Heterogeneous nuclear ribonucleoprotein R	91.483	0.682	14
POLR2A	DNA-directed RNA polymerase II subunit RPB1	21.389	0.6716	7
MYBBP1A	Myb-binding protein 1A	34.009	0.666	12
RPL23	60S ribosomal protein L23	28.332	0.664	6
STIP1	Stress-induced-phosphoprotein 1	63.365	0.6618	16
ATP2A2	Sarcoplasmic/endoplasmic reticulum calcium ATPase 2	85.469	0.6572	16
SET	Protein SET	323.31	0.6479	7
DSP	Desmoplakin	78.121	0.646	27
PLS3	Plastin-3	51.822	0.6427	15
TMPO	Lamina-associated polypeptide 2	69.45	0.636	6
ATP6V1B2	V-type proton ATPase subunit B, brain isoform	40.983	0.6311	6
SRSF7	Serine/arginine-rich splicing factor 7	20.995	0.6287	4
RPL18	60S ribosomal protein L18	47.041	0.6283	6
YWHAE	14-3-3 protein epsilon	158.28	0.626	16
RPL13	60S ribosomal protein L13	39.28	0.6137	5
BLMH	Bleomycin hydrolase	10.12	0.6073	4
PDHB	Pyruvate dehydrogenase E1 component subunit beta	22.666	0.6037	5
DDOST	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase	21.779	0.6007	7
RPS17	40S ribosomal protein S17	9.9434	0.58	4
EEF2	Elongation factor 2	323.31	0.5741	29
DDX17	Probable ATP-dependent RNA helicase DDX17	91.108	0.5733	16
RPS23	40S ribosomal protein S23	10.907	0.5687	4
CAND1	Cullin-associated NEDD8-dissociated protein 1	71.342	0.5685	13
HNRNPF	Heterogeneous nuclear ribonucleoprotein F	124.62	0.5672	8

SAE1	SUMO-activating enzyme subunit 1	75.896	0.5663	5
UBA1	Ubiquitin-like modifier-activating enzyme 1	135.09	0.5629	17
POLR3C	DNA-directed RNA polymerase III subunit RPC3	23.367	0.5621	4
PAICS	Multifunctional protein ADE2	68.055	0.5611	9
HNRNPA2B 1	Heterogeneous nuclear ribonucleoproteins A2/B1	211.19	0.561	16
ANXA2	Annexin	26.244	0.555	8
RDX	Radixin	30.389	0.5541	12
SSB	Lupus La protein	74.392	0.5489	12
EIF5A	Eukaryotic translation initiation factor 5A-1	42.535	0.5461	6
ILF3	Interleukin enhancer-binding factor 3	116.77	0.5411	14
SFPQ	Splicing factor, proline- and glutamine-rich	127.02	0.5405	16
DSG1	Desmoglein-1	32.702	0.5366	8
XPO1	Exportin-1	67.824	0.5357	14
SUPT5H	Transcription elongation factor SPT5	17.567	0.5326	5
HNRNPM	Heterogeneous nuclear ribonucleoprotein M	67.433	0.5323	22
PA2G4	Proliferation-associated protein 2G4	183.71	0.5313	11
PRMT1	Protein arginine N-methyltransferase 1	39.372	0.5277	9
RPS7	40S ribosomal protein S7	156.5	0.5251	8
EIF2S1	Eukaryotic translation initiation factor 2 subunit 1	16.709	0.5245	6
DCD	Dermcidin	4.2344	0.5082	2
SRSF1	Serine/arginine-rich splicing factor 1	26.9	0.5057	7
WDR61	WD repeat-containing protein 61	45.46	0.5006	5
EEF1A2	Elongation factor 1-alpha 2	34.964	0.5001	12

5.4 Discussion

RNA-binding proteins play a crucial role in the complex regulation of genes and therefore, there is a growing interest in determining these proteins and how they affect gene expression. Several studies were conducted to identify RBPs by utilizing affinity purification. For instance, the Schroeder laboratory has used the streptomycin aptamer tag to validate many predicted sRNA-protein interactions, including 6S RNA/RNAP, CsrB/CsrA, or OxyS/Hfq [204]. Likewise, the aptamer tag approach was employed to

identify RBPs associated with the U1 small nuclear ribonucleoprotein particles (snRNP) [205, 206].

In this chapter, I have used the interaction between MS2 RNA aptamer tagged *RNPS1* 3'UTR and MS2 coat protein to screen proteins associated with *RNPS1* mRNA in the cell. The MS2 aptamer interacts with high affinity with MS2cp with K values of 1-3 nM. This technique has the advantage of identifying protein-RNA interactions in cells. In comparison to in-vitro-based techniques, the expression of aptamer-tagged RNA inside the cell may have advantages. For example, transcription of RNA in the cell might be a prerequisite for the assembly of those RBPs that get deposited co-transcriptionally. Second, reversible formaldehyde cross-linking may be performed to trap transient RBPs that associate weakly or temporally with the tagged RNA. Additionally, the expression of aptamer-tagged RNAs can be used to visualize the spatial expression pattern of the RNA.

The MS2 RNA aptamer was used successfully as an RNA affinity tag to isolate RBPs associated with *RNPS1* 3'UTR. A total of 148 candidate binding proteins were pulled down by affinity purification from HEK 293 cells. Our mass spectrometry results show that *RNPS1* 3'UTR can bind to a variety of proteins, including ILF2, ILF3, HNRNPC, and PTBP1. ILF2, also known as nuclear factor 45 (NF45), forms a stable heterodimer with ILF3 (also called NF90) and acts as a crucial regulatory factor in cellular processes, including transcription, splicing, mRNA stability, translation, cell proliferation, and apoptosis [207-210]. Intriguingly, recent studies report ILF2 as a key host protein that is required in the replication of viruses, including human immunodeficiency virus type 1 (HIV-1) and hepatitis virus [211-213]. In the same line, our gene ontology analysis showed that RBPs bound to *RNPS1* 3'UTR play a role in influenza virus infection. Previous studies determined that RNPS1 interacts with nucleoprotein of Influenza A Virus (H7N9) and RNPS1 is plausibly involved in the replication process of influenza virus [214, 215].

Conclusively, this approach captured the protein interactome profile of *RNPS1* mRNA. Gene ontology analysis demonstrated that the identified proteins have diverse functions inside the cell. The association of *RNPS1* 3'UTR with regulatory proteins identified in this study provides a new direction for future studies on the regulation of RNPS1.

Chapter 6:
Summary and Future Perspective

Summary and Future Perspective

Since the discovery of split genes in eukaryotes, it has been known that a single gene can generate many mRNA and protein isoforms via the process of alternative splicing. This explains how the human genome can encode more than 100,000 proteins while only having about 24,000 protein-coding genes. Mapping alternate exons onto protein structures unraveled that they encode surface regions or loop regions crucial for protein-protein interactions. As a result, alternative splicing is a critical mechanism during cellular signaling and regulation of cell-type-specific functions.

Several studies have indicated that perturbations in alternative splicing contribute to tumor progression. Cancer cells exploit the process of alternative splicing to generate advantageous protein isoforms that drive oncogenesis. To this end, dysregulation in splicing contributes to all the hallmarks of cancer, including sustained proliferation, migration, invasion, angiogenesis, and evasion of apoptosis. In recent years, we have begun to understand that most tumor-related splicing modifications are due to changes in particular components of the splicing machinery.

In my first piece of work, I identified for the first time that RNPS1, a global guardian of splicing fidelity, is involved in the survival, migration, and invasion of cervical cancer cells. RNPS1 is a well-known constitutive and alternative splicing activator. Importantly, recent studies revealed the critical role of RNPS1 in maintaining transcriptome integrity by masking the usage of pseudo-splice sites. However, the pathophysiological significance of RNPS1 is poorly studied. *In silico* analysis of publicly available TCGA dataset unveiled interesting finding. The expression of RNPS1 was higher in cervical carcinoma samples compared to normal tissue. Consistently, the level of *RNPS1* was significantly higher in cervical cancer cells compared to the normal cell line. This alteration in the expression of RNPS1 in cervical cancer provided a rationale to investigate the possible role of RNPS1 in cancer.

Interestingly, knockdown of RNPS1 in cervical cancer cells, HeLa and SiHa, reduced the migration and invasion of cervical cancer cells. I wondered if RNPS1 exerts its effect by modulating the splicing of cancer-associated genes. Using RNA-Sequencing analysis, I identified that RNPS1 facilitates the generation of cancer-specific splice isoforms of genes involved in cell migration and invasion events. Our results showed that RNPS1 is involved in the alternative splicing of components of Rho GTPases, including *Rac1* and *RhoA*. Rho GTPases play a central role in cancer cell migration, invasion and cell polarity by modulating the cytoskeletal dynamics and cell adhesion. RhoA expression levels in tumors have been found to be drastically higher than the surrounding normal tissue; particularly high level of RhoA expression was found in the metastatic region. Overexpression of RhoA in cervical cancer is associated with increased migration and proliferation. It is well known that the expression of RhoA is transcriptionally regulated; however, so far, there is no evidence that RhoA overexpression in cancer cells is mediated by alternative splicing. In this study, I have found that RNPS1 switches alternative splicing of the *RhoA* transcript from a non-coding isoform to the major, translatable isoform, thus resulting in an increase in RhoA protein levels in cervical cancer cells.

Additionally, I have shown that RNPS1 regulates the splicing switch of *Rac1* and favors the formation of pro-tumorigenic isoform *Rac1b. Rac1b* differs from *Rac1* by harboring an extra exon 3b, rendering it constitutively active. The inclusion of exon 3b leads to an in-frame insertion of 19 amino acid residues in the proximity of the switch II domain, a regulatory region of GTPase. Consequently, Rac1b is mainly present in the active GTP-bound state due to its inability to interact with Rho-GDI. Furthermore, previous studies showed that Rac1b is overexpressed in multiple cancers and promotes cell survival and cell proliferation via NFκB signaling. In accordance with this, our data also showed that RNPS1 provides growth and survival advantage to cervical cancer cells plausibly by regulating the expression of Rac1b, RhoA, and other essential genes. Further, I have determined that RNPS1 is also involved in the alternative splicing of *WDR1,* a protein associated with migration and metastasis. RNPS1 controls the switch between the full-

length isoform of *WDR1* and the truncated isoform termed *WDRΔ35* that lacks exons three, four and five. Finally, I also show that RNPS1 confers chemoresistance against the drug doxorubicin in cervical cancer cells. This is the first study to provide detailed molecular insights into the role of RNPS1 in cancer, particularly cervical cancer.

In the second piece of work, I elucidated the mechanism of how RNPS1 is dysregulated in cervical cancer cells and discovered that RNPS1 is subjected to post-transcriptional gene regulation via miR-6893-3p. I found that miR-6893-3p expression is downregulated in cervical cancer cells and an inverse correlation was observed between miR-6893-3p and RNPS1 expression in cervical cancer cells. I hypothesized that the negative correlation of expression between miR-6893-3p and RNPS1 levels might suggest a regulatory role for miR-6893-3p in RNPS1 expression. The results showed that miR-6893-3p negatively regulates RNPS1 by binding to the microRNA response element in the 3'UTR of RNPS1. Further, the targeted negative regulation of RNPS1 by miR-6893-3p occurs via enhanced mRNA decay.

In the third project, MS2 tagged RNA affinity purification was performed to pull down the protein interactome profile of *RNPS1* 3'UTR. This approach identified several novel regulatory proteins, including ILF2, ILF3, HNRNPC, and PTBP1. Developing an atlas of all the protein interactome within *RNPS1* mRNA would help us identify the "central protein" that can be critical for regulating RNPS1. However, the functional role of the interaction between these regulatory proteins and the *RNPS1* transcript was not explored in this study due to the time limitation. Further work would enable the characterization of this functional interaction in cells, which may unveil an additional or complementary mechanism for the regulation of RNPS1 by regulatory proteins.

References

1. Wang Z, Burge CB: **Splicing regulation: From a parts list of regulatory elements to an integrated splicing code**. *RNA* 2008, **14**(5):802-813.
2. Kastner B, Will CL, Stark H, Luhrmann R: **Structural Insights into Nuclear pre-mRNA Splicing in Higher Eukaryotes**. *Cold Spring Harbor Perspectives in Biology* 2019, **11**(11):a032417.
3. Zhuang Y, Weiner AM: **A compensatory base change in U1 snRNA suppresses a 5' splice site mutation**. *Cell* 1986, **46**(6):827-835.
4. Kondo Y, Oubridge C, van Roon A-MM, Nagai K: **Crystal structure of human U1 snRNP, a small nuclear ribonucleoprotein particle, reveals the mechanism of 5' splice site recognition**. *eLife* 2015, **4**:e04986.
5. Liang W-W, Cheng S-C: **A novel mechanism for Prp5 function in prespliceosome formation and proofreading the branch site sequence**. *Genes & Development* 2015, **29**(1):81-93.
6. Kistler AL, Guthrie C: **Deletion of MUD2, the yeast homolog of U2AF65, can bypass the requirement for sub2, an essential spliceosomal ATPase**. *Genes & Development* 2001, **15**(1):42-49.
7. Hausner TP, Giglio LM, Weiner AM: **Evidence for base-pairing between mammalian U2 and U6 small nuclear ribonucleoprotein particles**. *Genes & Development* 1990, **4**(12a):2146-2156.
8. Staley JP, Guthrie C: **An RNA Switch at the 5' Splice Site Requires ATP and the DEAD Box Protein Prp28p**. *Molecular Cell* 1999, **3**(1):55-64.
9. Charenton C, Wilkinson ME, Nagai K: **Mechanism of 5' splice site transfer for human spliceosome activation**. *Science (New York, NY)* 2019, **364**(6438):362-367.
10. Zhan X, Yan C, Zhang X, Lei J, Shi Y: **Structures of the human pre-catalytic spliceosome and its precursor spliceosome**. *Cell Research* 2018, **28**(12):1129-1140.
11. Madhani HD, Guthrie C: **A novel base-pairing interaction between U2 and U6 snRNAs suggests a mechanism for the catalytic activation of the spliceosome**. *Cell* 1992, **71**(5):803-817.
12. Raghunathan PL, Guthrie C: **RNA unwinding in U4/U6 snRNPs requires ATP hydrolysis and the DEIH-box splicing factor Brr2**. *Current Biology* 1998, **8**(15):847-855.
13. Galej WP, Wilkinson ME, Fica SM, Oubridge C, Newman AJ, Nagai K: **Cryo-EM structure of the spliceosome immediately after branching**. *Nature* 2016, **537**(7619):197-201.
14. Schwer B, Guthrie C: **A conformational rearrangement in the spliceosome is dependent on PRP16 and ATP hydrolysis**. *The EMBO Journal* 1992, **11**(13):5033-5039.
15. Fica SM, Oubridge C, Galej WP, Wilkinson ME, Bai X-C, Newman AJ, Nagai K: **Structure of a spliceosome remodelled for exon ligation**. *Nature* 2017, **542**(7641):377-380.
16. Gilbert W: **Why genes in pieces?** *Nature* 1978, **271**(5645):501-501.
17. Celotto AM, Graveley BR: **Alternative splicing of the Drosophila Dscam pre-mRNA is both temporally and spatially regulated**. *Genetics* 2001, **159**(2):599-608.
18. Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ: **Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing**. *Nature genetics* 2008, **40**(12):1413-1415.
19. Blencowe BJ: **Alternative Splicing: New Insights from Global Analyses**. *Cell* 2006, **126**(1):37-47.

20. Galante PAF, Sakabe NJ, Kirschbaum-Slager N, de Souza SJ: **Detection and evaluation of intron retention events in the human transcriptome**. *RNA (New York, NY)* 2004, **10**(5):757-765.
21. 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 USA* 1992, **89**.
22. Cho S, Hoang A, Sinha R, Zhong X-Y, Fu X-D, Krainer AR, Ghosh G: **Interaction between the RNA binding domains of Ser-Arg splicing factor 1 and U1-70K snRNP protein determines early spliceosome assembly**. *Proceedings of the National Academy of Sciences of the United States of America* 2011, **108**(20):8233-8238.
23. Busch A, Hertel KJ: **Evolution of SR protein and hnRNP splicing regulatory factors**. *Wiley Interdisciplinary Reviews RNA* 2012, **3**(1):1-12.
24. Wu JY, Maniatis T: **Specific interactions between proteins implicated in splice site selection and regulated alternative splicing**. *Cell* 1993, **75**.
25. Shen H, Green MR: **A pathway of sequential arginine-serine-rich domain-splicing signal interactions during mammalian spliceosome assembly**. *Mol Cell* 2004, **16**.
26. Shen H, Kan JL, Green MR: **Arginine-serine-rich domains bound at splicing enhancers contact the branchpoint to promote prespliceosome assembly**. *Mol Cell* 2004, **13**.
27. Krainer AR, Conway GC, Kozak D: **Purification and characterization of pre-mRNA splicing factor SF2 from HeLa cells**. *Genes Dev* 1990, **4**.
28. Ge H, Manley JL: **A protein factor, ASF, controls cell-specific alternative splicing of SV40 early pre-mRNA in vitro**. *Cell* 1990, **62**(1):25-34.
29. Maquat LE, Tarn W-Y, Isken O: **The Pioneer Round of Translation: Features and Functions**. *Cell* 2010, **142**(3):368-374.
30. Topisirovic I, Svitkin YV, Sonenberg N, Shatkin AJ: **Cap and cap-binding proteins in the control of gene expression**. *WIREs RNA* 2010, **2**(2):277-298.
31. Lemay J-F, Lemieux C, St-Andre O, Bachand F: **Crossing the borders: Poly(A)-binding proteins working on both sides of the fence**. *RNA Biology* 2010, **7**(3):291-295.
32. Doidge R, Mittal S, Aslam A, Winkler GS: **Deadenylation of cytoplasmic mRNA by the mammalian Ccr4-Not complex**. In: *Biochemical Society Transactions: 2012*; 2012: 896-901.
33. Meyer KD, Jaffrey SR: **The dynamic epitranscriptome: N6-methyladenosine and gene expression control**. *Nature Reviews Molecular Cell Biology* 2014, **15**(5):313-326.
34. Norbury CJ: **Cytoplasmic RNA: a case of the tail wagging the dog**. *Nature Reviews Molecular Cell Biology* 2013, **14**(10):643-653.
35. Subtelny AO, Eichhorn SW, Chen GR, Sive H, Bartel DP: **Poly(A)-tail profiling reveals an embryonic switch in translational control**. *Nature* 2014, **508**(7494):66-71.
36. Mauer J, Luo X, Blanjoie A, Jiao X, Grozhik AV, Patil DP, Linder B, Pickering BF, Vasseur J-J, Chen Q *et al*: **Reversible methylation of m6Am in the 5' cap controls mRNA stability**. *Nature* 2016, **541**(7637):371-375.
37. Le Hir H, Izaurralde E, Maquat LE, Moore MJ: **The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions**. *The EMBO Journal* 2000, **19**(24):6860-6869.
38. Lejeune F, Ishigaki Y, Li X, Maquat LE: **The exon junction complex is detected on CBP80-bound but not eIF4E-bound mRNA in mammalian cells: dynamics of mRNP remodeling**. *The EMBO Journal* 2002, **21**(13):3536-3545.
39. Tange TØ, Shibuya T, Jurica MS, Moore MJ: **Biochemical analysis of the EJC reveals two new factors and a stable tetrameric protein core**. *RNA* 2005, **11**(12):1869-1883.

40. Gerbracht JV, Boehm V, Britto-Borges T, Kallabis S, Wiederstein Janica L, Ciriello S, Aschemeier DU, Kruger M, Frese CK, Altmuller J *et al*: **CASC3 promotes transcriptome-wide activation of nonsense-mediated decay by the exon junction complex**. *Nucleic Acids Research* 2020, **48**(15):8626-8644.
41. Bessonov S, Anokhina M, Will CL, Urlaub H, Luhrmann R: **Isolation of an active step I spliceosome and composition of its RNP core**. *Nature* 2008, **452**(7189):846-850.
42. Jurica MS, Licklider LJ, Gygi SR, Grigorieff N, Moore MJ: **Purification and characterization of native spliceosomes suitable for three-dimensional structural analysis**. *RNA (New York, NY)* 2002, **8**(4):426-439.
43. Reichert VL, Le Hir H, Jurica MS, Moore MJ: **5' exon interactions within the human spliceosome establish a framework for exon junction complex structure and assembly**. *Genes & Development* 2002, **16**(21):2778-2791.
44. Alexandrov A, Colognori D, Shu M-D, Steitz Joan A: **Human spliceosomal protein CWC22 plays a role in coupling splicing to exon junction complex deposition and nonsense-mediated decay**. *Proceedings of the National Academy of Sciences* 2012, **109**(52):21313-21318.
45. Steckelberg A-L, Boehm V, Gromadzka Agnieszka M, Gehring Niels H: **CWC22 Connects Pre-mRNA Splicing and Exon Junction Complex Assembly**. *Cell Reports* 2012, **2**(3):454-461.
46. Barbosa I, Haque N, Fiorini F, Barrandon C, Tomasetto C, Blanchette M, Le Hir H: **Human CWC22 escorts the helicase eIF4AIII to spliceosomes and promotes exon junction complex assembly**. *Nature Structural & Molecular Biology* 2012, **19**(10):983-990.
47. Buchwald G, Schussler S, Basquin C, Le Hir H, Conti E: **Crystal structure of the human eIF4AIII-CWC22 complex shows how a DEAD-box protein is inhibited by a MIF4G domain**. *Proceedings of the National Academy of Sciences* 2013, **110**(48):E4611-E4618.
48. Ideue T, Sasaki YTF, Hagiwara M, Hirose T: **Introns play an essential role in splicing-dependent formation of the exon junction complex**. *Genes & Development* 2007, **21**(16):1993-1998.
49. Linder P, Jankowsky E: **From unwinding to clamping - the DEAD box RNA helicase family**. *Nature Reviews Molecular Cell Biology* 2011, **12**(8):505-516.
50. Andersen Christian BF, Ballut L, Johansen Jesper S, Chamieh H, Nielsen Klaus H, Oliveira Cristiano LP, Pedersen Jan S, Seraphin B, Hir Herve L, Andersen Gregers R: **Structure of the Exon Junction Core Complex with a Trapped DEAD-Box ATPase Bound to RNA**. *Science* 2006, **313**(5795):1968-1972.
51. Bono F, Ebert J, Lorentzen E, Conti E: **The Crystal Structure of the Exon Junction Complex Reveals How It Maintains a Stable Grip on mRNA**. *Cell* 2006, **126**(4):713-725.
52. Gehring NH, Lamprinaki S, Kulozik AE, Hentze MW: **Disassembly of Exon Junction Complexes by PYM**. *Cell* 2009, **137**(3):536-548.
53. Bono F, Ebert J, Unterholzner L, Guttler T, Izaurralde E, Conti E: **Molecular insights into the interaction of PYM with the Mago-Y14 core of the exon junction complex**. *EMBO reports* 2004, **5**(3):304-310.
54. Bono F, Cook AG, Grunwald M, Ebert J, Conti E: **Nuclear import mechanism of the EJC component Mago-Y14 revealed by structural studies of importin 13**. *Molecular Cell* 2010, **37**(2):211-222.
55. Mingot J-M, Kostka S, Kraft R, Hartmann E, Gorlich D: **Importin 13: a novel mediator of nuclear import and export**. *The EMBO Journal* 2001, **20**(14):3685-3694.
56. Woodward LA, Mabin JW, Gangras P, Singh G: **The exon junction complex: a lifelong guardian of mRNA fate**. *WIREs RNA* 2017, **8**(3):e1411.

57. Mayeda A, Badolato J, Kobayashi R, Zhang MQ, Gardiner EM, Krainer AR: **Purification and characterization of human RNPS1: a general activator of pre-mRNA splicing**. *The EMBO Journal* 1999, **18**(16):4560-4570.
58. Murachelli AG, Ebert J, Basquin C, Le Hir H, Conti E: **The structure of the ASAP core complex reveals the existence of a Pinin-containing PSAP complex**. *Nat Struct Mol Biol* 2012, **19**(4):378-386.
59. Boehm V, Britto-Borges T, Steckelberg A-L, Singh KK, Gerbracht JV, Gueney E, Blazquez L, Altmuller J, Dieterich C, Gehring NH: **Exon Junction Complexes Suppress Spurious Splice Sites to Safeguard Transcriptome Integrity**. *Molecular Cell* 2018, **72**(3):482-495.e487.
60. Sahara S, Aoto M, Eguchi Y, Imamoto N, Yoneda Y, Tsujimoto Y: **Acinus is a caspase-3-activated protein required for apoptotic chromatin condensation**. *Nature* 1999, **401**(6749):168-173.
61. Aravind L, Koonin EV: **SAP – a putative DNA-binding motif involved in chromosomal organization**. *Trends in Biochemical Sciences* 2000, **25**(3):112-114.
62. Schwerk C, Prasad J, Degenhardt K, Erdjument-Bromage H, White E, Tempst P, Kidd VJ, Manley JL, Lahti JM, Reinberg D: **ASAP, a Novel Protein Complex Involved in RNA Processing and Apoptosis**. *Molecular and Cellular Biology* 2003, **23**(8):2981-2990.
63. McCallum SA, Bazan JF, Merchant M, Yin J, Pan B, de Sauvage FJ, Fairbrother WJ: **Structure of SAP18: A Ubiquitin Fold in Histone Deacetylase Complex Assembly**. *Biochemistry* 2006, **45**(39):11974-11982.
64. Kiel C, Serrano L: **The Ubiquitin Domain Superfold: Structure-based Sequence Alignments and Characterization of Binding Epitopes**. *Journal of Molecular Biology* 2006, **355**(4):821-844.
65. Fukumura K, Wakabayashi S, Kataoka N, Sakamoto H, Suzuki Y, Nakai K, Mayeda A, Inoue K: **The Exon Junction Complex Controls the Efficient and Faithful Splicing of a Subset of Transcripts Involved in Mitotic Cell-Cycle Progression**. *International Journal of Molecular Sciences* 2016, **17**(8):1153.
66. Trembley JH, Tatsumi S, Sakashita E, Loyer P, Slaughter CA, Suzuki H, Endo H, Kidd VJ, Mayeda A: **Activation of pre-mRNA splicing by human RNPS1 is regulated by CK2 phosphorylation**. *Molecular and Cellular Biology* 2005, **25**(4):1446-1457.
67. De Conti L, Baralle M, Buratti E: **Exon and intron definition in pre-mRNA splicing**. *Wiley Interdisciplinary Reviews: RNA* 2013, **4**(1):49-60.
68. Ashton-Beaucage D, Udell CM, Lavoie H, Baril C, Lefrançois M, Chagnon P, Gendron P, Caron-Lizotte O, Bonneil É, Thibault P *et al*: **The Exon Junction Complex Controls the Splicing of mapk and Other Long Intron-Containing Transcripts in Drosophila**. *Cell* 2010, **143**(2):251-262.
69. Roignant J-Y, Treisman JE: **Exon Junction Complex Subunits Are Required to Splice Drosophila MAP Kinase, a Large Heterochromatic Gene**. *Cell* 2010, **143**(2):238-250.
70. Malone CD, Mestdagh C, Akhtar J, Kreim N, Deinhard P, Sachidanandam R, Treisman J, Roignant J-Y: **The exon junction complex controls transposable element activity by ensuring faithful splicing of the piwi transcript**. *Genes & Development* 2014, **28**(16):1786-1799.
71. Hayashi R, Handler D, Ish-Horowicz D, Brennecke J: **The exon junction complex is required for definition and excision of neighboring introns in Drosophila**. *Genes & Development* 2014, **28**(16):1772-1785.
72. Malone CD, Brennecke J, Dus M, Stark A, McCombie WR, Sachidanandam R, Hannon GJ: **Specialized piRNA Pathways Act in Germline and Somatic Tissues of the Drosophila Ovary**. *Cell* 2009, **137**(3):522-535.

73. Liu M, Li Y, Liu A, Li R, Su Y, Du J, Li C, Zhu AJ: **The exon junction complex regulates the splicing of cell polarity gene dlg1 to control Wingless signaling in development**. *eLife* 2016, **5**:e17200.
74. Blazquez L, Emmett W, Faraway R, Pineda JMB, Bajew S, Gohr A, Haberman N, Sibley CR, Bradley RK, Irimia M *et al*: **Exon Junction Complex Shapes the Transcriptome by Repressing Recursive Splicing**. *Molecular Cell* 2018, **72**(3):496-509.e499.
75. Fukumura K, Inoue K, Mayeda A: **Splicing activator RNPS1 suppresses errors in pre-mRNA splicing: A key factor for mRNA quality control**. *Biochemical and Biophysical Research Communications* 2018, **496**(3):921-926.
76. Gonatopoulos-Pournatzis T, Wu M, Braunschweig U, Roth J, Han H, Best AJ, Raj B, Aregger M, O Hanlon D, Ellis JD *et al*: **Genome-wide CRISPR-Cas9 Interrogation of Splicing Networks Reveals a Mechanism for Recognition of Autism-Misregulated Neuronal Microexons**. *Molecular Cell* 2018, **72**(3):510-524.e512.
77. Irimia M, Weatheritt Robert J, Ellis JD, Parikshak Neelroop N, Gonatopoulos-Pournatzis T, Babor M, Quesnel-Vallieres M, Tapial J, Raj B, O'Hanlon D *et al*: **A Highly Conserved Program of Neuronal Microexons Is Misregulated in Autistic Brains**. *Cell* 2014, **159**(7):1511-1523.
78. Ustianenko D, Weyn-Vanhentenryck SM, Zhang C: **Microexons: discovery, regulation, and function**. *Wiley Interdisciplinary Reviews RNA* 2017, **8**(4):10.1002/wrna.1418.
79. Ohnishi T, Shirane M, Hashimoto Y, Saita S, Nakayama KI: **Identification and characterization of a neuron-specific isoform of protrudin**. *Genes to Cells* 2014, **19**(2):97-111.
80. Dergai M, Tsyba L, Dergai O, Zlatskii I, Skrypkina I, Kovalenko V, Rynditch A: **Microexon-based regulation of ITSN1 and Src SH3 domains specificity relies on introduction of charged amino acids into the interaction interface**. *Biochemical and Biophysical Research Communications* 2010, **399**(2):307-312.
81. Irimia M, Weatheritt Robert J, Ellis JD, Parikshak Neelroop N, Gonatopoulos-Pournatzis T, Babor M, Quesnel-Vallieres M, Tapial J, Raj B, O'Hanlon D *et al*: **A Highly Conserved Program of Neuronal Microexons Is Misregulated in Autistic Brains**. *Cell* 2014, **159**(7):1511-1523.
82. Nott A, Meislin SH, Moore MJ: **A quantitative analysis of intron effects on mammalian gene expression**. *RNA* 2003, **9**(5):607-617.
83. Nott A, Le Hir H, Moore MJ: **Splicing enhances translation in mammalian cells: an additional function of the exon junction complex**. *Genes & Development* 2004, **18**(2):210-222.
84. Wiegand HL, Lu S, Cullen BR: **Exon Junction Complexes Mediate the Enhancing Effect of Splicing on mRNA Expression**. *Proceedings of the National Academy of Sciences of the United States of America* 2003, **100**(20):11327-11332.
85. McCracken S, Lambermon M, Blencowe BJ: **SRm160 Splicing Coactivator Promotes Transcript 3'-End Cleavage**. *Molecular and Cellular Biology* 2002, **22**(1):148-160.
86. Murthy KG, Manley JL: **The 160-kD subunit of human cleavage-polyadenylation specificity factor coordinates pre-mRNA 3'-end formation**. *Genes & Development* 1995, **9**(21):2672-2683.
87. Kervestin S, Jacobson A: **NMD: a multifaceted response to premature translational termination**. *Nat Rev Mol Cell Biol* 2012, **13**(11):700-712.
88. Popp MW-L, Maquat LE: **Organizing Principles of Mammalian Nonsense-Mediated mRNA Decay**. *Annual review of genetics* 2013, **47**:139-165.

89. Metze S, Herzog VA, Ruepp M-D, Muhlemann O: **Comparison of EJC-enhanced and EJC-independent NMD in human cells reveals two partially redundant degradation pathways**. *RNA (New York, NY)* 2013, **19**(10):1432-1448.
90. Gehring NH, Kunz JB, Neu-Yilik G, Breit S, Viegas MH, Hentze MW, Kulozik AE: **Exon-Junction Complex Components Specify Distinct Routes of Nonsense-Mediated mRNA Decay with Differential Cofactor Requirements**. *Molecular Cell* 2005, **20**(1):65-75.
91. Lykke-Andersen J, Shu M-D, Steitz JA: **Human Upf Proteins Target an mRNA for Nonsense-Mediated Decay When Bound Downstream of a Termination Codon**. *Cell* 2000, **103**(7):1121-1131.
92. Schlautmann LP, Boehm V, Lackmann J-W, Altmuller J, Dieterich C, Gehring NH: **Exon junction complex-associated multi-adapter RNPS1 nucleates splicing regulatory complexes to maintain transcriptome surveillance**. *bioRxiv* 2021:2021.2008.2020.457088.
93. Venter JC, Adams Mark D, Myers Eugene W, Li Peter W, Mural Richard J, Sutton Granger G, Smith Hamilton O, Yandell M, Evans Cheryl A, Holt Robert A *et al*: **The Sequence of the Human Genome**. *Science* 2001, **291**(5507):1304-1351.
94. Conne B, Stutz A, Vassalli J-D: **The 3' untranslated region of messenger RNA: A molecular 'hotspot' for pathology?** *Nature Medicine* 2000, **6**:637.
95. Misquitta CM, Iyer VR, Werstiuk ES, Grover AK: **The role of 3'-untranslated region (3'-UTR) mediated mRNA stability in cardiovascular pathophysiology**. *Molecular and Cellular Biochemistry* 2001, **224**(1):53-67.
96. Timchenko LT: **Myotonic dystrophy: the role of RNA CUG triplet repeats**. *American Journal of Human Genetics* 1999, **64**(2):360-364.
97. Szostak E, Gebauer F: **Translational control by 3'-UTR-binding proteins**. *Briefings in functional genomics* 2013, **12**(1):58-65.
98. Matoulkova E, Michalova E, Vojtesek B, Hrstka R: **The role of the 3' untranslated region in post-transcriptional regulation of protein expression in mammalian cells**. *RNA Biology* 2012, **9**(5):563-576.
99. Kishore S, Luber S, Zavolan M: **Deciphering the role of RNA-binding proteins in the post-transcriptional control of gene expression**. *Briefings in functional genomics* 2010, **9**(5-6):391-404.
100. Keene JD: **Ribonucleoprotein infrastructure regulating the flow of genetic information between the genome and the proteome**. *Proceedings of the National Academy of Sciences* 2001, **98**(13):7018.
101. Anantharaman V, Koonin EV, Aravind L: **Comparative genomics and evolution of proteins involved in RNA metabolism**. *Nucleic Acids Research* 2002, **30**(7):1427-1464.
102. Sawicka K, Bushell M, Spriggs Keith A, Willis Anne E: **Polypyrimidine-tract-binding protein: a multifunctional RNA-binding protein**. *Biochemical Society Transactions* 2008, **36**(4):641-647.
103. Vlasova IA, Bohjanen PR: **Posttranscriptional regulation of gene networks by GU-rich elements and CELF proteins**. *RNA Biology* 2008, **5**(4):201-207.
104. Hui J, Reither G, Bindereif A: **Novel functional role of CA repeats and hnRNP L in RNA stability**. *RNA (New York, NY)* 2003, **9**(8):931-936.
105. Jonas S, Izaurralde E: **Towards a molecular understanding of microRNA-mediated gene silencing**. *Nature Reviews Genetics* 2015, **16**(7):421-433.
106. Bartel DP: **MicroRNAs: target recognition and regulatory functions**. *Cell* 2009, **136**(2):215-233.

107. Fang Z, Rajewsky N: **The Impact of miRNA Target Sites in Coding Sequences and in 3'UTRs**. *PLOS ONE* 2011, **6**(3):e18067.
108. Seok H, Ham J, Jang E-S, Chi SW: **MicroRNA Target Recognition: Insights from Transcriptome-Wide Non-Canonical Interactions**. *Molecules and cells* 2016, **39**(5):375-381.
109. Eichhorn Stephen W, Guo H, McGeary Sean E, Rodriguez-Mias Ricard A, Shin C, Baek D, Hsu S-h, Ghoshal K, Villen J, Bartel David P: **mRNA Destabilization Is the Dominant Effect of Mammalian MicroRNAs by the Time Substantial Repression Ensues**. *Molecular Cell* 2014, **56**(1):104-115.
110. Wahle E, Winkler GS: **RNA decay machines: Deadenylation by the Ccr4-Not and Pan2-Pan3 complexes**. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* 2013, **1829**(6):561-570.
111. Braun JE, Truffault V, Boland A, Huntzinger E, Chang C-T, Haas G, Weichenrieder O, Coles M, Izaurralde E: **A direct interaction between DCP1 and XRN1 couples mRNA decapping to 5' exonucleolytic degradation**. *Nature Structural & Molecular Biology* 2012, **19**(12):1324-1331.
112. Fukao A, Mishima Y, Takizawa N, Oka S, Imataka H, Pelletier J, Sonenberg N, Thoma C, Fujiwara T: **MicroRNAs Trigger Dissociation of eIF4AI and eIF4AII from Target mRNAs in Humans**. *Molecular Cell* 2014, **56**(1):79-89.
113. Fukaya T, Iwakawa H-o, Tomari Y: **MicroRNAs Block Assembly of eIF4F Translation Initiation Complex in Drosophila**. *Molecular Cell* 2014, **56**(1):67-78.
114. Vasudevan S, Steitz JA: **AU-rich-element-mediated upregulation of translation by FXR1 and Argonaute 2**. *Cell* 2007, **128**(6):1105-1118.
115. Valinezhad Orang A, Safaralizadeh R, Kazemzadeh-Bavili M: **Mechanisms of miRNA-Mediated Gene Regulation from Common Downregulation to mRNA-Specific Upregulation**. *International journal of genomics* 2014, **2014**:970607-970607.
116. Boguslawska J, Sokol E, Rybicka B, Czubaty A, Rodzik K, Piekielko-Witkowska A: **microRNAs target SRSF7 splicing factor to modulate the expression of osteopontin splice variants in renal cancer cells**. *Gene* 2016, **595**(2):142-149.
117. Sokol, E, Kedzierska H, Czubaty A, Rybicka B, Rodzik K, Tanski Z, Boguslawska J, Piekielko-Witkowska A: **microRNA-mediated regulation of splicing factors SRSF1, SRSF2 and hnRNP A1 in context of their alternatively spliced 3' UTRs**. *Experimental Cell Research* 2018, **363**(2):208-217.
118. Song L, Lin H-S, Gong J-N, Han H, Wang X-S, Su R, Chen M-T, Shen C, Ma Y-N, Yu J *et al*: **microRNA-451-modulated hnRNP A1 takes a part in granulocytic differentiation regulation and acute myeloid leukemia**. *Oncotarget* 2017, **8**(33):55453-55466.
119. Tazi J, Bakkour N, Stamm S: **Alternative splicing and disease**. *Biochimica et biophysica acta* 2009, **1792**(1):14-26.
120. Scotti MM, Swanson MS: **RNA mis-splicing in disease**. *Nature Reviews Genetics* 2016, **17**:19.
121. Biamonti G, Catillo M, Pignataro D, Montecucco A, Ghigna C: **The alternative splicing side of cancer**. *Seminars in Cell & Developmental Biology* 2014, **32**:30-36.
122. Pajares MJ, Ezponda T, Catena R, Calvo A, Pio R, Montuenga LM: **Alternative splicing: an emerging topic in molecular and clinical oncology**. *The Lancet Oncology* 2007, **8**(4):349-357.
123. Shuai S, Suzuki H, Diaz-Navarro A, Nadeu F, Kumar SA, Gutierrez-Fernandez A, Delgado J, Pinyol M, Lopez-Otin C, Puente XS *et al*: **The U1 spliceosomal RNA is recurrently mutated in multiple cancers**. *Nature* 2019, **574**(7780):712-716.

124. 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.
125. Zheng X, Zhan Z, Naren D, Li J, Yan T, Gong Y: **Prognostic value of SRSF2 mutations in patients with de novo myelodysplastic syndromes: A meta-analysis**. *PLOS ONE* 2017, **12**(9):e0185053.
126. Meggendorfer M, Roller A, Haferlach T, Eder C, Dicker F, Grossmann V, Kohlmann A, Alpermann T, Yoshida K, Ogawa S *et al*: **SRSF2 mutations in 275 cases with chronic myelomonocytic leukemia (CMML)**. *Blood* 2012, **120**(15):3080-3088.
127. El Marabti E, Younis I: **The Cancer Spliceome: Reprograming of Alternative Splicing in Cancer**. *Frontiers in Molecular Biosciences* 2018, **5**.
128. Anczukow O, Rosenberg AZ, Akerman M, Das S, Zhan L, Karni R, Muthuswamy SK, Krainer AR: **The splicing factor SRSF1 regulates apoptosis and proliferation to promote mammary epithelial cell transformation**. *Nature Structural & Molecular Biology* 2012, **19**(2):220-228.
129. Karni R, de Stanchina E, Lowe SW, Sinha R, Mu D, Krainer AR: **The gene encoding the splicing factor SF2/ASF is a proto-oncogene**. *Nat Struct Mol Biol* 2007, **14**.
130. Ghigna C, Giordano S, Shen H, Benvenuto F, Castiglioni F, Comoglio PM, Green MR, Riva S, Biamonti G: **Cell Motility Is Controlled by SF2/ASF through Alternative Splicing of the Ron Protooncogene**. *Molecular Cell* 2005, **20**(6):881-890.
131. Jiang L, Huang J, Higgs BW, Hu Z, Xiao Z, Yao X, Conley S, Zhong H, Liu Z, Brohawn P *et al*: **Genomic Landscape Survey Identifies SRSF1 as a Key Oncodriver in Small Cell Lung Cancer**. *PLoS genetics* 2016, **12**(4):e1005895-e1005895.
132. Olshavsky NA, Comstock CES, Schiewer MJ, Augello MA, Hyslop T, Sette C, Zhang J, Parysek LM, Knudsen KE: **Identification of ASF/SF2 as a critical, allele-specific effector of the cyclin D1b oncogene**. *Cancer research* 2010, **70**(10):3975-3984.
133. Elias AP, Dias S: **Microenvironment Changes (in pH) Affect VEGF Alternative Splicing**. *Cancer Microenvironment* 2008, **1**(1):131-139.
134. Jia R, Li C, McCoy JP, Deng C-X, Zheng Z-M: **SRp20 is a proto-oncogene critical for cell proliferation and tumor induction and maintenance**. *International journal of biological sciences* 2010, **6**(7):806-826.
135. Peiqi L, Zhaozhong G, Yaotian Y, Jun J, Jihua G, Rong J: **Expression of SRSF3 is Correlated with Carcinogenesis and Progression of Oral Squamous Cell Carcinoma**. *International journal of medical sciences* 2016, **13**(7):533-539.
136. Wang Z, Chatterjee D, Jeon HY, Akerman M, Vander Heiden MG, Cantley LC, Krainer AR: **Exon-centric regulation of pyruvate kinase M alternative splicing via mutually exclusive exons**. *Journal of Molecular Cell Biology* 2012, **4**(2):79-87.
137. Tang Y, Horikawa I, Ajiro M, Robles AI, Fujita K, Mondal AM, Stauffer JK, Zheng ZM, Harris CC: **Downregulation of splicing factor SRSF3 induces p53β, an alternatively spliced isoform of p53 that promotes cellular senescence**. *Oncogene* 2013, **32**(22):2792-2798.
138. Liu X, Ma H, Ma L, Li K, Kang Y: **RNA-binding protein with serine-rich domain 1 regulates microsatellite instability of uterine corpus endometrial adenocarcinoma**. *Clinics (Sao Paulo, Brazil)* 2021, **76**:e3318-e3318.
139. Arbyn M, Weiderpass E, Bruni L, de Sanjose S, Saraiya M, Ferlay J, Bray F: **Estimates of incidence and mortality of cervical cancer in 2018: a worldwide analysis**. *The Lancet Global health* 2020, **8**(2):e191-e203.

140. Johnson CA, James D, Marzan A, Armaos M: **Cervical Cancer: An Overview of Pathophysiology and Management**. *Seminars in Oncology Nursing* 2019, **35**(2):166-174.
141. Cohen PA, Jhingran A, Oaknin A, Denny L: **Cervical cancer**. *The Lancet* 2019, **393**(10167):169-182.
142. Francies FZ, Bassa S, Chatziioannou A, Kaufmann AM, Dlamini Z: **Splicing Genomics Events in Cervical Cancer: Insights for Phenotypic Stratification and Biomarker Potency**. *Genes* 2021, **12**(2):130.
143. Kahles A, Lehmann K-V, Toussaint NC, Huser M, Stark SG, Sachsenberg T, Stegle O, Kohlbacher O, Sander C, Caesar-Johnson SJ *et al*: **Comprehensive Analysis of Alternative Splicing Across Tumors from 8,705 Patients**. *Cancer Cell* 2018, **34**(2):211-224.e216.
144. Liu F, Dai M, Xu Q, Zhu X, Zhou Y, Jiang S, Wang Y, Ai Z, Ma L, Zhang Y *et al*: **SRSF10-mediated IL1RAP alternative splicing regulates cervical cancer oncogenesis via mIL1RAP-NF-κB-CD47 axis**. *Oncogene* 2018, **37**(18):2394-2409.
145. Cascino I, Fiucci G, Papoff G, Ruberti G: **Three functional soluble forms of the human apoptosis-inducing Fas molecule are produced by alternative splicing**. *The Journal of Immunology* 1995, **154**(6):2706.
146. Cheng J, Zhou T, Liu C, Shapiro John P, Brauer Matthew J, Kiefer Michael C, Barr Philip J, Mountz John D: **Protection from Fas-Mediated Apoptosis by a Soluble Form of the Fas Molecule**. *Science* 1994, **263**(5154):1759-1762.
147. Lawson CD, Ridley AJ: **Rho GTPase signaling complexes in cell migration and invasion**. *Journal of Cell Biology* 2017, **217**(2):447-457.
148. Schnelzer A, Prechtel D, Knaus U, Dehne K, Gerhard M, Graeff H, Harbeck N, Schmitt M, Lengyel E: **Rac1 in human breast cancer: overexpression, mutation analysis, and characterization of a new isoform, Rac1b**. *Oncogene* 2000, **19**(26):3013-3020.
149. Fiegen D, Haeusler L-C, Blumenstein L, Herbrand U, Dvorsky R, Vetter IR, Ahmadian MR: **Alternative Splicing of Rac1 Generates Rac1b, a Self-activating GTPase***. *Journal of Biological Chemistry* 2004, **279**(6):4743-4749.
150. Singh A, Karnoub AE, Palmby TR, Lengyel E, Sondek J, Der CJ: **Rac1b, a tumor associated, constitutively active Rac1 splice variant, promotes cellular transformation**. *Oncogene* 2004, **23**(58):9369-9380.
151. He X, Yuan C, Yang J: **Regulation and functional significance of CDC42 alternative splicing in ovarian cancer**. *Oncotarget 2015 Oct 6;6(30):29651-63 doi: 1018632/oncotarget4865* 2015.
152. Michelle L, Cloutier A, Toutant J, Shkreta L, Thibault P, Durand M, Garneau D, Gendron D, Lapointe E, Couture S *et al*: **Proteins Associated with the Exon Junction Complex Also Control the Alternative Splicing of Apoptotic Regulators**. *Molecular and Cellular Biology* 2012, **32**(5):954-967.
153. Piperigkou Z, Kyriakopoulou K, Koutsakis C, Mastronikolis S, Karamanos NK: **Key Matrix Remodeling Enzymes: Functions and Targeting in Cancer**. *Cancers* 2021, **13**(6):1441.
154. Melzer C, Hass R, Lehnert H, Ungefroren H: **RAC1B: A Rho GTPase with Versatile Functions in Malignant Transformation and Tumor Progression**. *Cells* 2019, **8**(1):21.
155. Ridley AJ: **Rho GTPases and cell migration**. *Journal of Cell Science* 2001, **114**(15):2713-2722.
156. Ono S: **Regulation of Actin Filament Dynamics by Actin Depolymerizing Factor/Cofilin and Actin-Interacting Protein 1: New Blades for Twisted Filaments**. *Biochemistry* 2003, **42**(46):13363-13370.

157. Talora C, Sgroi DC, Crum CP, Dotto GP: **Specific down-modulation of Notch1 signaling in cervical cancer cells is required for sustained HPV-E6/E7 expression and late steps of malignant transformation**. *Genes & Development* 2002, **16**(17):2252-2263.
158. Alfieri C, Zhang S, Barford D: **Visualizing the complex functions and mechanisms of the anaphase promoting complex/cyclosome (APC/C)**. *Open Biology* 2017, **7**(11):170204.
159. Quintero-Fabian S, Arreola R, Becerril-Villanueva E, Torres-Romero JC, Arana-Argaez V, Lara-Riegos J, Ramirez-Camacho MA, Alvarez-Sanchez ME: **Role of Matrix Metalloproteinases in Angiogenesis and Cancer**. *Frontiers in Oncology* 2019, **9**.
160. Morgan EL, Scarth JA, Patterson MR, Wasson CW, Hemingway GC, Barba-Moreno D, Macdonald A: **E6-mediated activation of JNK drives EGFR signalling to promote proliferation and viral oncoprotein expression in cervical cancer**. *Cell death and differentiation* 2021, **28**(5):1669-1687.
161. Hu L, Zou F, Grandis JR, Johnson DE: **Chapter 4 - The JNK Pathway in Drug Resistance**. In: *Targeting Cell Survival Pathways to Enhance Response to Chemotherapy.* vol. 3: Academic Press; 2019: 87-100.
162. Stallings-Mann ML, Waldmann J, Zhang Y, Miller E, Gauthier ML, Visscher DW, Downey GP, Radisky ES, Fields AP, Radisky DC: **Matrix metalloproteinase induction of Rac1b, a key effector of lung cancer progression**. *Science translational medicine* 2012, **4**(142):142ra195-142ra195.
163. Goka ET, Lippman ME, Cho CH, Hu T: **Chapter 9 - Rac1b: An emerging therapeutic target for chemoresistance in colorectal cancer**. In: *Drug Resistance in Colorectal Cancer: Molecular Mechanisms and Therapeutic Strategies.* vol. 8: Academic Press; 2020: 153-171.
164. Gembarska A, Luciani F, Fedele C, Russell EA, Dewaele M, Villar S, Zwolinska A, Haupt S, de Lange J, Yip D *et al*: **MDM4 is a key therapeutic target in cutaneous melanoma**. *Nature medicine* 2012, **18**(8):1239-1247.
165. Bardot B, Toledo F: **Targeting MDM4 Splicing in Cancers**. *Genes* 2017, **8**(2):82.
166. Dewaele M, Tabaglio T, Willekens K, Bezzi M, Teo SX, Low DHP, Koh CM, Rambow F, Fiers M, Rogiers A *et al*: **Antisense oligonucleotide-mediated MDM4 exon 6 skipping impairs tumor growth**. *The Journal of clinical investigation* 2016, **126**(1):68-84.
167. Wang J-W, Peng S-Y, Li J-T, Wang Y, Zhang Z-P, Cheng Y, Cheng D-Q, Weng W-H, Wu X-S, Fei X-Z *et al*: **Identification of metastasis-associated proteins involved in gallbladder carcinoma metastasis by proteomic analysis and functional exploration of chloride intracellular channel 1**. *Cancer Letters* 2009, **281**(1):71-81.
168. Lee JH, Kim JE, Kim BG, Han HH, Kang S, Cho NH: **STAT3-induced WDR1 overexpression promotes breast cancer cell migration**. *Cellular Signalling* 2016, **28**(11):1753-1760.
169. Yuan B, Zhang R, Hu J, Liu Z, Yang C, Zhang T, Zhang C: **WDR1 Promotes Cell Growth and Migration and Contributes to Malignant Phenotypes of Non-small Cell Lung Cancer through ADF/cofilin-mediated Actin Dynamics**. *International journal of biological sciences* 2018, **14**(9):1067-1080.
170. Kuchenbauer F, Mah SM, Heuser M, McPherson A, Ruschmann J, Rouhi A, Berg T, Bullinger L, Argiropoulos B, Morin RD *et al*: **Comprehensive analysis of mammalian miRNA* species and their role in myeloid cells**. *Blood* 2011, **118**(12):3350-3358.
171. Jagadeeswaran G, Zheng Y, Sumathipala N, Jiang H, Arrese EL, Soulages JL, Zhang W, Sunkar R: **Deep sequencing of small RNA libraries reveals dynamic regulation of conserved and novel microRNAs and microRNA-stars during silkworm development**. *BMC Genomics* 2010, **11**:52-52.

172. Okamura K, Phillips MD, Tyler DM, Duan H, Chou Y-t, Lai EC: **The regulatory activity of microRNA* species has substantial influence on microRNA and 3' UTR evolution**. *Nature Structural & Molecular Biology* 2008, **15**(4):354-363.
173. Yang J-S, Phillips MD, Betel D, Mu P, Ventura A, Siepel AC, Chen KC, Lai EC: **Widespread regulatory activity of vertebrate microRNA* species**. *RNA (New York, NY)* 2011, **17**(2):312-326.
174. Ro S, Park C, Young D, Sanders KM, Yan W: **Tissue-dependent paired expression of miRNAs**. *Nucleic Acids Research* 2007, **35**(17):5944-5953.
175. Calin GA, Cimmino A, Fabbri M, Ferracin M, Wojcik SE, Shimizu M, Taccioli C, Zanesi N, Garzon R, Aqeilan RI *et al*: **MiR-15a and miR-16-1 cluster functions in human leukemia**. *Proceedings of the National Academy of Sciences of the United States of America* 2008, **105**(13):5166-5171.
176. Cimmino A, Calin GA, Fabbri M, Iorio MV, Ferracin M, Shimizu M, Wojcik SE, Aqeilan RI, Zupo S, Dono M *et al*: **miR-15 and miR-16 induce apoptosis by targeting BCL2**. *Proceedings of the National Academy of Sciences of the United States of America* 2005, **102**(39):13944-13949.
177. Yanaihara N, Caplen N, Bowman E, Seike M, Kumamoto K, Yi M, Stephens RM, Okamoto A, Yokota J, Tanaka T *et al*: **Unique microRNA molecular profiles in lung cancer diagnosis and prognosis**. *Cancer Cell* 2006, **9**(3):189-198.
178. Murakami Y, Yasuda T, Saigo K, Urashima T, Toyoda H, Okanoue T, Shimotohno K: **Comprehensive analysis of microRNA expression patterns in hepatocellular carcinoma and non-tumorous tissues**. *Oncogene* 2006, **25**(17):2537-2545.
179. Iorio MV, Ferracin M, Liu C-G, Veronese A, Spizzo R, Sabbioni S, Magri E, Pedriali M, Fabbri M, Campiglio M *et al*: **MicroRNA Gene Expression Deregulation in Human Breast Cancer**. *Cancer research* 2005, **65**(16):7065-7070.
180. Lee EJ, Gusev Y, Jiang J, Nuovo GJ, Lerner MR, Frankel WL, Morgan DL, Postier RG, Brackett DJ, Schmittgen TD: **Expression profiling identifies microRNA signature in pancreatic cancer**. *International journal of cancer* 2007, **120**(5):1046-1054.
181. Wu F, Zhou J: **CircAGFG1 promotes cervical cancer progression via miR-370-3p/RAF1 signaling**. *BMC cancer* 2019, **19**(1):1067-1067.
182. Yungang W, Xiaoyu L, Pang T, Wenming L, Pan X: **miR-370 targeted FoxM1 functions as a tumor suppressor in laryngeal squamous cell carcinoma (LSCC)**. *Biomedicine & Pharmacotherapy* 2014, **68**(2):149-154.
183. Chen X-P, Chen Y-G, Lan J-Y, Shen Z-J: **MicroRNA-370 suppresses proliferation and promotes endometrioid ovarian cancer chemosensitivity to cDDP by negatively regulating ENG**. *Cancer Letters* 2014, **353**(2):201-210.
184. Shen X, Zuo X, Zhang W, Bai Y, Qin X, Hou N: **MiR-370 promotes apoptosis in colon cancer by directly targeting MDM4**. *Oncology letters* 2018, **15**(2):1673-1679.
185. Lo SS, Hung PS, Chen JH, Tu HF, Fang WL, Chen CY, Chen WT, Gong NR, Wu CW: **Overexpression of miR-370 and downregulation of its novel target TGFβ-RII contribute to the progression of gastric carcinoma**. *Oncogene* 2012, **31**(2):226-237.
186. Wu Z, Sun H, Zeng W, He J, Mao X: **Upregulation of MircoRNA-370 induces proliferation in human prostate cancer cells by downregulating the transcription factor FOXO1**. *PLOS ONE* 2012, **7**(9):e45825-e45825.
187. Kotagama K, Schorr AL, Steber HS, Mangone M: **ALG-1 Influences Accurate mRNA Splicing Patterns in the Caenorhabditis elegans Intestine and Body Muscle Tissues by Modulating Splicing Factor Activities**. *Genetics 2019 Jul;212(3):931-951 doi: 101534/genetics119302223 Epub 2019 May 9* 2019.

188. Kalsotra A, Wang K, Li PF, Cooper TA: **MicroRNAs coordinate an alternative splicing network during mouse postnatal heart development**. *Genes Dev 2010 Apr 1;24(7):653-8 doi: 101101/gad1894310 Epub 2010 Mar 18* 2010.
189. Balcells I, Cirera S, Busk PK: **Specific and sensitive quantitative RT-PCR of miRNAs with DNA primers**. *BMC Biotechnol 2011 Jun 25;11:70 doi: 101186/1472-6750-11-70* 2011.
190. Busk PK: **A tool for design of primers for microRNA-specific quantitative RT-qPCR**. *BMC Bioinformatics 2014 Jan 28;15:29 doi: 101186/1471-2105-15-29* 2014.
191. Chen C-Y, Ezzeddine N, Shyu A-B: **Chapter 17 Messenger RNA Half-Life Measurements in Mammalian Cells**. *Methods in Enzymology* 2008, **448**.
192. Qu B, Han X, Tang Y, Shen N: **A Novel Vector-Based Method for Exclusive Overexpression of Star-Form MicroRNAs**. *PLOS ONE* 2012, **7**(7):e41504.
193. Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, Tanzer A, Lagarde J, Lin W, Schlesinger F *et al*: **Landscape of transcription in human cells**. *Nature 2012 Sep 6;489(7414):101-8 doi: 101038/nature11233* 2012.
194. Chen M, Ai G, Zhou J, Mao W, Li H, Guo J: **circMTO1 promotes tumorigenesis and chemoresistance of cervical cancer via regulating miR-6893**. *Biomedicine & Pharmacotherapy* 2019, **117**:109064.
195. Scott AT, Craig CC, Patrick JL, Jack DK: **Identifying mRNA subsets in messenger ribonucleoprotein complexes by using cDNA arrays**. *Proceedings of the National Academy of Sciences* 2000, **97**(26):14085-14090.
196. Keene JD, Tenenbaum SA: **Eukaryotic mRNPs may represent posttranscriptional operons**. *Molecular Cell* 2002, **9**(6):1161-1167.
197. Hogan DJ, Riordan DP, Gerber AP, Herschlag D, Brown PO: **Diverse RNA-Binding Proteins Interact with Functionally Related Sets of RNAs, Suggesting an Extensive Regulatory System**. *PLOS Biology* 2008, **6**(10):e255.
198. Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, Rothballer A, Ascano Jr M, Jungkamp A-C, Munschauer M *et al*: **Transcriptome-wide Identification of RNA-Binding Protein and MicroRNA Target Sites by PAR-CLIP**. *Cell* 2010, **141**(1):129-141.
199. Durie D, Lewis SM, Liwak U, Kisilewicz M, Gorospe M, Holcik M: **RNA-binding protein HuR mediates cytoprotection through stimulation of XIAP translation**. *Oncogene* 2011, **30**(12):1460-1469.
200. Srisawat C, Engelke DR: **Streptavidin aptamers: affinity tags for the study of RNAs and ribonucleoproteins**. *RNA* 2001, **7**(4):632-641.
201. Srisawat C, Goldstein IJ, Engelke DR: **Sephadex-binding RNA ligands: rapid affinity purification of RNA from complex RNA mixtures**. *Nucleic Acids Research* 2001, **29**(2):e4-e4.
202. Leppek K, Stoecklin G: **An optimized streptavidin-binding RNA aptamer for purification of ribonucleoprotein complexes identifies novel ARE-binding proteins**. *Nucleic Acids Research* 2014, **42**(2):e13-e13.
203. Capitanio JS, Wozniak RW: **Host Cell Factors Necessary for Influenza A Infection: Meta-Analysis of Genome Wide Studies**. *arXiv: Cell Behavior* 2012.
204. Windbichler N, von Pelchrzim F, Mayer O, Csaszar E, Schroeder R: **Isolation of small RNA-binding proteins from E. coli: Evidence for frequent interaction of RNAs with RNA polymerase**. *RNA Biology* 2008, **5**(1):30-40.
205. Bardwell VJ, Wickens M: **Purification of RNA and RNA-protein complexes by an R17 coat protein affinity method**. *Nucleic Acids Research* 1990, **18**(22):6587-6594.

206. Bachler M, Schroeder R, von Ahsen U: **StreptoTag: a novel method for the isolation of RNA-binding proteins**. *RNA* 1999, **5**(11):1509-1516.

207. Guan D, Altan-Bonnet N, Parrott AM, Arrigo CJ, Li Q, Khaleduzzaman M, Li H, Lee CG, Pe'ery T, Mathews MB: **Nuclear factor 45 (NF45) is a regulatory subunit of complexes with NF90/110 involved in mitotic control**. *Mol Cell Biol 2008 Jul;28(14):4629-41 doi: 101128/MCB00120-08 Epub 2008 May 5* 2008.

208. Shamanna RA, Hoque M, Lewis-Antes A, Azzam EI, Lagunoff D, Pe'ery T, Mathews MB: **The NF90/NF45 complex participates in DNA break repair via nonhomologous end joining**. *Mol Cell Biol 2011 Dec;31(23):4832-43 doi: 101128/MCB05849-11 Epub 2011 Oct 3* 2011.

209. Rigo F, Hua Y, Chun SJ, Prakash TP, Krainer AR, Bennett CF: **Synthetic oligonucleotides recruit ILF2/3 to RNA transcripts to modulate splicing**. *Nat Chem Biol 2012 Apr 15;8(6):555-61 doi: 101038/nchembio939* 2012.

210. Higuchi T, Todaka H, Sugiyama Y, Ono M, Tamaki N, Hatano E, Takezaki Y, Hanazaki K, Miwa T, Lai S *et al*: **Suppression of MicroRNA-7 (miR-7) Biogenesis by Nuclear Factor 90-Nuclear Factor 45 Complex (NF90-NF45) Controls Cell Proliferation in Hepatocellular Carcinoma**. *J Biol Chem 2016 Sep 30;291(40):21074-21084 doi: 101074/jbcM116748210 Epub 2016 Aug 12* 2016.

211. Shin HJ, Kim SS, Cho YH, Lee SG, Rho HM: **Host cell proteins binding to the encapsidation signal ε in hepatitis B virus RNA**. *Archives of Virology* 2002, **147**(3):471-491.

212. Li Y, Belshan M: **NF45 and NF90 Bind HIV-1 RNA and Modulate HIV Gene Expression**. *Viruses* 2016, **8**(2):47.

213. Isken O, Baroth M, Grassmann CW, Weinlich S, Ostareck DH, Ostareck-Lederer A, Behrens S-E: **Nuclear factors are involved in hepatitis C virus RNA replication**. *RNA* 2007, **13**(10):1675-1692.

214. Rogan PK, Mucaki EJ, Shirley BC: **A proposed molecular mechanism for pathogenesis of severe RNA-viral pulmonary infections**. *F1000Research* 2021, **9**(943).

215. Sun N, Sun W, Li S, Yang J, Yang L, Quan G, Gao X, Wang Z, Cheng X, Li Z *et al*: **Proteomics Analysis of Cellular Proteins Co-Immunoprecipitated with Nucleoprotein of Influenza A Virus (H7N9)**. *International Journal of Molecular Sciences* 2015, **16**(11):25982-25998.

Appendix I

List of primers used in this study.

Oligo	Sequence (5' --3')	Purpose
ANAPC5 FP	GCCCAGTTATGGATGCTATGT	qPCR
ANAPC5 RP	GCTATTGAGAGCTGTGATTCCT	qPCR
ANAPC7 FP	TGACCCAAAGGCCAGATTAC	qPCR
ANAPC7 RP	CGTTCCTCAGCAAAGCAATTC	qPCR
APP FP	GCCAAAGAGACATGCAGTGA	qPCR
APP RP	AGTCATCCTCCTCCGCATC	qPCR
beta-actin FP	CTGTACGCCAACACAGTGCT	qPCR
beta-actin RP	GCTCAGGAGGAGCAATGATC	qPCR
Caspase 1 FP	CTGCTCTTCCACACCAGATAAT	qPCR
Caspase 1 RP	TTTCCTCCACATCACAGGAAC	qPCR
Caspase 4 FP	AAGAGAAGCAACGTATGGCAGGAC	qPCR
Caspase 4 RP	GGACAAAGCTTGAGGGCATCTGTA	qPCR
Cathepsin V FP	TGGAAGGCAACACACAGAAG	qPCR
Cathepsin V RP	GAAGCCATGTTTCCCTTGG	qPCR
CDKN2B FP	GAATGCGCGAGGAGAACAA	qPCR
CDKN2B RP	CATCATCATGACCTGGATCGC	qPCR
FGF2 FP	GCGACCCACACGTCAAACTA	qPCR
FGF2 RP	CTTAGAAGCCAGCAGCCG	qPCR
JAK3 FP	GAATCCCTCTCGGACAACATC	qPCR
JAK3 RP	TTTGTCGCAGTAGGTGAAGAG	qPCR
MAPK8 FP	GACGCCTTATGTAGTGACTCGC	qPCR
MAPK8 RP	TCCTGGAAAGAGGATTTTGTGG	qPCR
MDM4_FL FP	GATGCTGCTCAGACTCTCGC	qPCR

MDM4_FL RP	TGCACTTTGCTTCAGTTGGTC	qPCR
MDM4_S FP	GCCACTGCTACTACAGCAAAG	qPCR
MDM4_S RP	TCTGAGGTAGGCAGTGTGGG	qPCR
miR-6893-3p FP	CCTGCTGCCTTCACCT	qPCR
miR-6893-3p RP	GGTCCAGTTTTTTTTTTTTTTCTGGC	qPCR
NOTCH1 FP	GACCTCCCCAACACCTACAA	qPCR
NOTCH1 RP	TGACATCCCCCTCACAGC	qPCR
RAC1 FP	TGATGCAGGCCATCAAGTGT	qPCR
RAC1 RP	AGAACACATCTGTTTGCGGAT	qPCR
RAC1B FP	AAGACAAGCCGATTGCCGAT	qPCR
RAC1B RP	GACCCTGCGGATAGGTGATG	qPCR
RHOA FP	CATCGACAGCCCTGATAGTTTAG	qPCR
RHOA RP	TCTTATTCCCAACCAGGATGATG	qPCR
RNPS1 3'UTR FP	ACCCATGGTAGTTGCTGCTC	qPCR
RNPS1 3'UTR RP	AGCTGGCTCTCCACTCACTC	qPCR
RNPS1 FP	TCCCGCTCCAAATCCAAACC	qPCR
RNPS1 RP	CGGGTGAGTCTCCCAATGTG	qPCR
TIMP3 FP	CTGACAGGTCGCGTCTATGA	qPCR
TIMP3 RP	GGCGTAGTGTTTGGACTGGT	qPCR
Total MDM4 FP	AGGTGCGCAAGGTGAAATGT	qPCR
Total MDM4 RP	CCATATGCTGCTCCTGCTGAT	qPCR
WDR1 FP	TGCTTGGACTGAAGACAGTAAG	qPCR
WDR1 RP	GTAATCTCGCCCACAGAAGAG	qPCR
WDR-del35 FP	AAGGAACATCGACGACCACA	qPCR
WDR-del35 RP	GATGTATATCTGGCCGTCAGC	qPCR
MDM4 SPLICE VARIANT FP	TGTGGTGGAGATCTTTTGGG	RT-PCR
MDM4 SPLICE VARIANT	GCAGTGTGGGGATATCGT	RT-PCR

RP		
RAC1 SPLICE VARIANT FP	TGCCAATGTTATGGTAGATGG	RT-PCR
RAC1 SPLICE VARIANT RP	TGGGAGTCAGCTTCTTCTCC	RT-PCR
tdTomato FP	CCAAGCTGAAGGTGACCAAG	RT-PCR
tdTomato RP	CGCGCATCTTCACCTTGTAG	RT-PCR
pMirGlo mir-6893 mutant FP	ATTTGCAGGCGATAGGCAGCACCC	Site-directed Mutagenesis
pMirGlo mir-6893 mutant RP	CCTGCCAGAGTCAAGGGT	Site-directed Mutagenesis

Appendix II

List of antibodies used in this study.

Antibodies	Company	Identifier
Caspase 4	Santa cruz	sc-56056
GAPDH	CST	3683S
MMP9	CST	3852
N-Cadherin	CST	CST-13116T
P-STAT3	Invitrogen	44-384G
Rac1b	Sigma	09-271
RhoA	CST	CST-2117T
STAT3	CST	9132
Tubulin	Sigma	T6074

Appendix III

List of proteins identified by mass spectrometry from induced Rosetta lysate transformed with empty pHis_TEV parent vector.

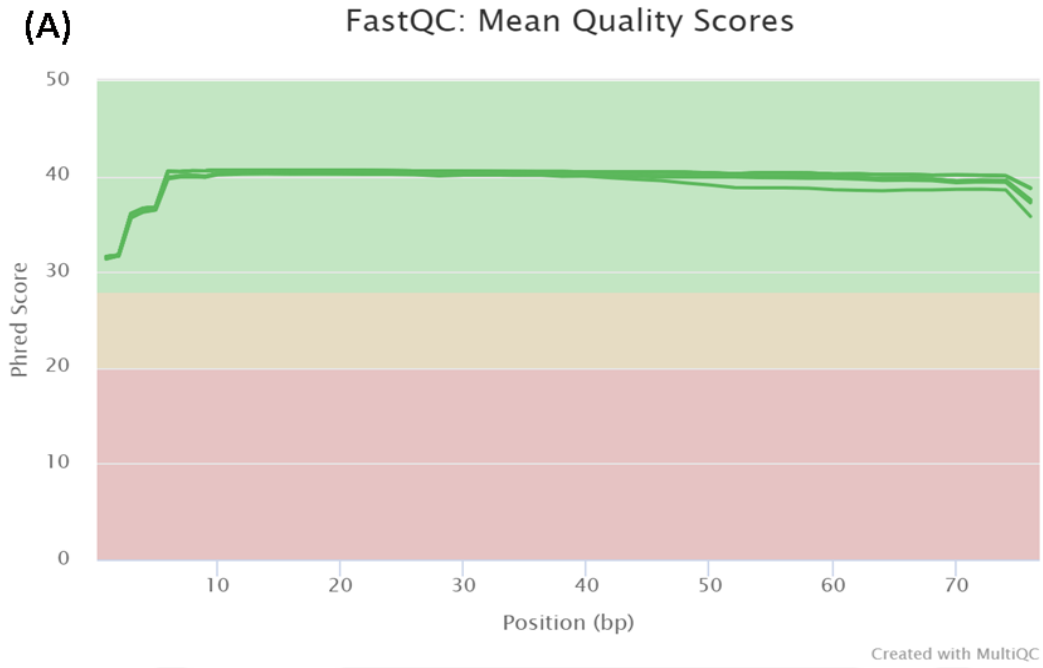
Gene Name	Description	Score	Queries Matched
Rho	Transcription termination factor Rho OS=Escherichia coli	429.98	42
ENO	Enolase OS=Escherichia coli	375.78	34
SHMT	Serine hydroxymethyltransferase OS=Escherichia coli	364.15	31
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase A OS=Escherichia coli	306.4	30
TALDO1	Transaldolase B OS=Escherichia coli	237.9	28
IscS	Cysteine desulfurase IscS OS=Escherichia coli	92.44	15
RpoA	DNA-directed RNA polymerase subunit alpha OS=Escherichia coli	227.69	19

List of proteins identified by mass spectrometry from induced Rosetta lysate transformed with pHis_TEV_RNPS1 vector.

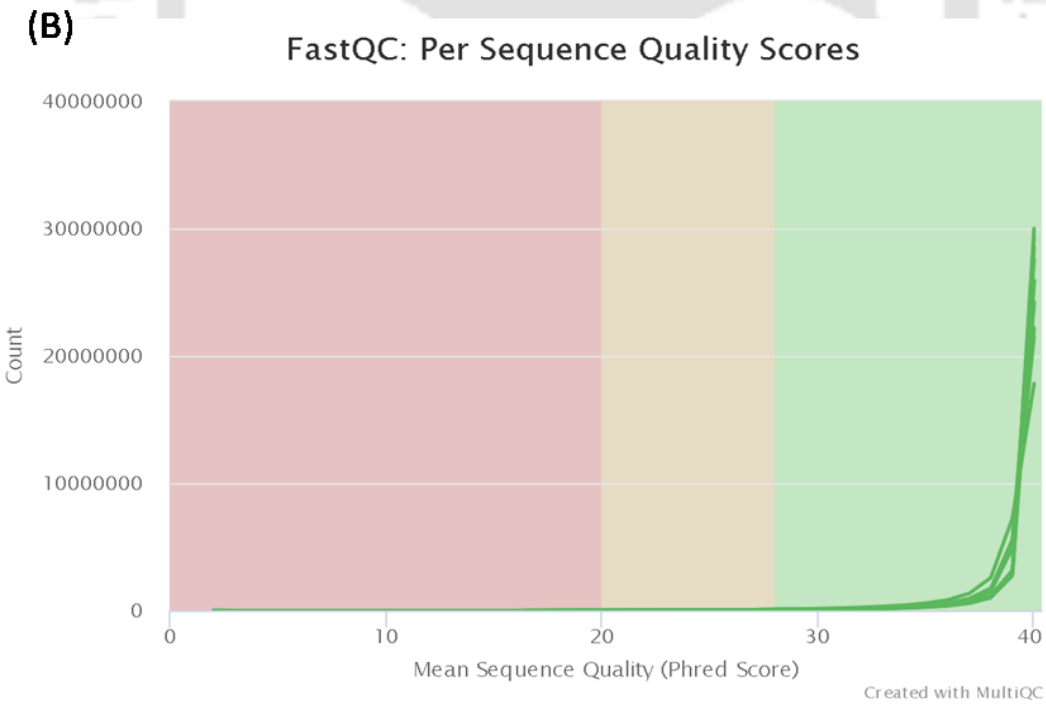
Gene Name	Description	Score	Queries Matched
Rho	Transcription termination factor Rho OS=Escherichia coli	403.96	34
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase A OS=Escherichia coli	285.22	34
RNPS1	RNA-binding protein with serine-rich domain 1 OS=Homo sapiens	246.77	27
ENO	Enolase OS=Escherichia coli	213.2	23
IscS	Cysteine desulfurase IscS OS=Escherichia coli	84.08	23
SHMT	Serine hydroxymethyltransferase OS=Escherichia coli	163.78	22
RpoA	DNA-directed RNA polymerase subunit alpha OS=Escherichia coli	200.31	20
TALDO1	Transaldolase B OS=Escherichia coli	119.35	20

Appendix IV

(A)



(B)



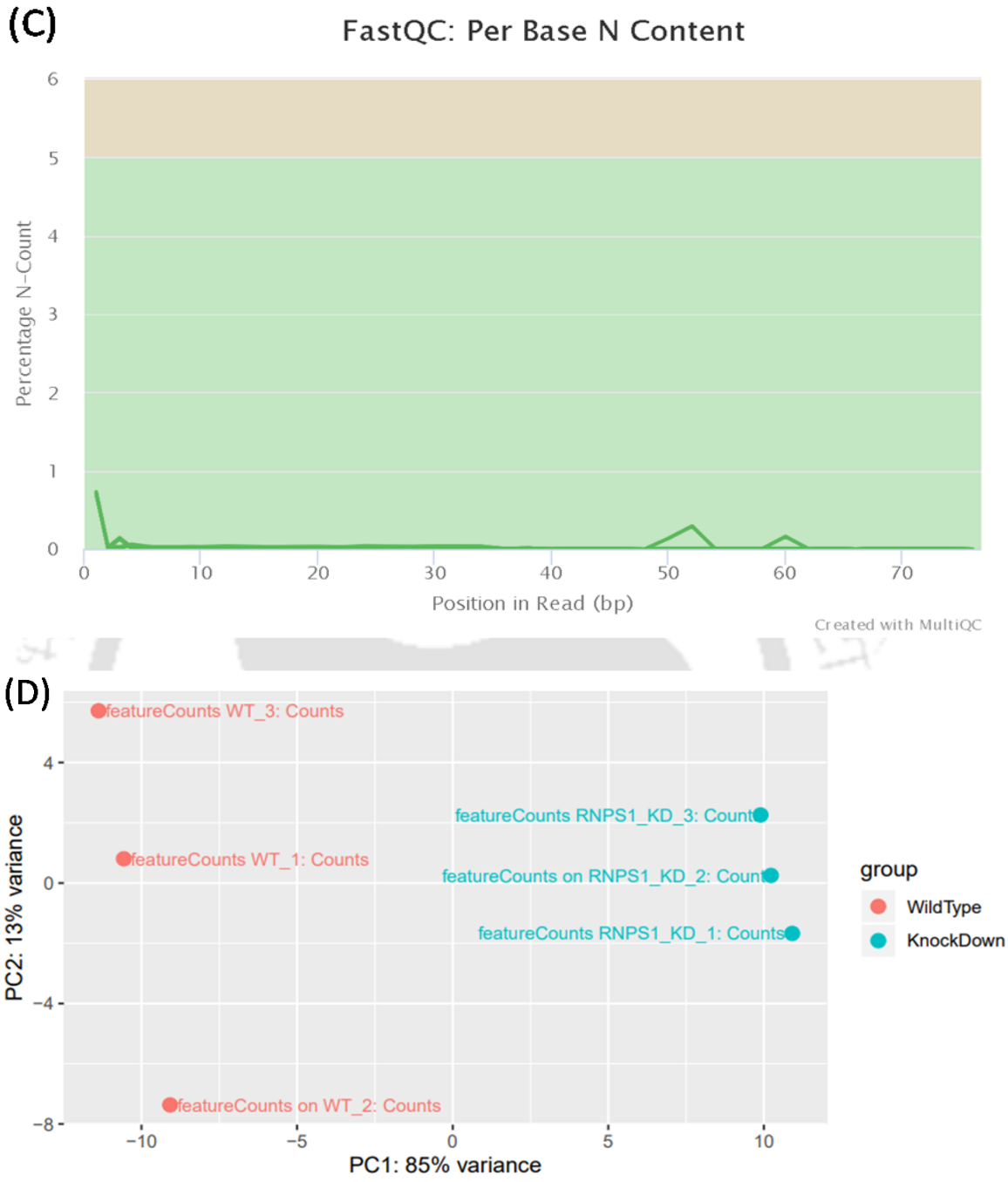


Figure A4.1. RNA-Seq data analysis and statistics. (A) The graph shows the mean quality score of RNA-Seq data. In high quality sequence data this is expected to be above 30. This shows the quality of the data is good. (B) The graph shows the distribution of average read quality score of RNA-Seq data. (C) The graph shows the percentage of base calls at each position for which an indistinguishable base was called in RNA-Seq data. In high quality sequence data this is expected to be close to zero. (D) Principal-component analysis showing distinct overall gene expression profiles between wild type and RNPS1 knockdown samples.

Appendix V

List of Publications

1. **Deka B**, Chandra P, Yadav P, Rehman A, Kumari S, Kunnumakkara AB, Singh KK. RNPS1 functions as an oncogenic splicing factor in cervical cancer cells. IUBMB Life. 2022; doi: 10.1002/iub.2686.
2. **Deka B**, Singh KK. Molecular cloning, expression and generation of a polyclonal antibody specific for RNPS1. Molecular Biology Reports. 2022; 9095–9100. https://doi.org/10.1007/s11033-022-07676-8.
3. **Deka B,** Singh KK. Identification of candidate RNA binding proteins associated with RNPS1 3'UTR. (Book Chapter) Healthcare Research and Related Technologies - Proceedings of NERC 2022 (Springer Nature).
4. **Deka B**, Chandra P, Singh KK. Functional roles of human Up-frameshift suppressor 3 (UPF3) proteins: From nonsense-mediated mRNA decay to neurodevelopmental disorders. Biochimie 2020; 180:10-22. doi.org/10.1016/j.biochi.2020.10.011.
5. **Deka B** and Singh KK. The arginine and serine (RS)-rich domains of Acinus modulate splicing. Cell. Bio. International. 2019; 43(8):954-959. doi: 10.1002/cbin.11163.
6. **Deka B**, Singh KK. Multifaceted Regulation of Gene Expression by the Apoptosis- and Splicing-Associated Protein Complex and Its Components. International Journal of Biological Sciences. 2017; 13(5):545-560. doi:10.7150/ijbs.18649.
7. **Deka B**, Rehman A, Singh KK. miR-6893-3p is a bonafide negative regulator of splicing activator, RNPS1. 2022. Cell. Bio. International. (Manuscript Communicated).
8. Kumari S, **DeKa B**, Singh KK. Serine-rich domain of RNPS1 functions in activation of alternative splicing. 2022. Genes to cells (Under Peer Review).
9. Chandra P, **Deka B**, Kumari S, Rehman A, Yadav P, Singh KK. UPF3B-knockout stabilizes GADD45G transcripts and delays the cell proliferation rate in HEK-293 cells. 2022. (To be Communicated).

Conferences and Workshops attended

Presentations (poster and oral) in conferences

1. **Deka B**, Singh KK. Elucidating the role of microRNA and RBP-mediated modulation of RNPS1 gene expression. Research Conclave'19, Indian Institute of Technology Guwahati. (Poster presentation).
2. **Deka B**, Singh KK. Understanding the Role of RNA-binding proteins in the Post-transcriptional regulation of RNPS1 gene. 1st Departmental Retreat (Biotech Express) organized by Department of Biosciences and Bioengineering, Indian Institute of Technology Guwahati, 2019 (Oral presentation).
3. **Deka B**, Chandra P, Singh KK. RNPS1 functions as an oncogenic splicing factor in cervical carcinoma. 41st Annual conference of the Indian Association for Cancer Research 2022, Amity University. (Poster presentation).
4. **Deka B**, Rehman A, Singh KK. Understanding the Role of microRNAs in the Post-Transcriptional Regulation of RNPS1 gene. Cold Spring Harbor meeting: Regulatory & Non-Coding RNAs 2022. (Poster presentation).
5. **Deka B**, Rehman A, Singh KK. Understanding the Role of microRNAs in the Post-Transcriptional Regulation of RNPS1 gene. North-East Research Conclave, IIT Guwahati 2022. (Poster presentation).
6. **Deka B**, Chandra P, Singh KK. RNPS1 functions as an oncogenic splicing factor in cervical cancer cells. 11th RNA Group Meeting 2022, NCCS Pune. (Poster presentation).

Workshops attended

1. Deka B. Participated in three-day Flow Cytometry workshop on "Flow Applications in Basics, Applied and Clinical Biology", organized by Indian Institute of Technology Guwahati and Dr. B. Borooah Cancer Institute, Guwahati. 2016.

2. Deka B. Participated in three-day workshop on “Genome/Transcriptome Sequence Analysis”, organized by ADBS and IBAB, Bangalore, 2019.