UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE QUÍMICA E BIOQUÍMICA

Regulation of the alternative splicing of RAC1b in tumour cells

Cláudia Alexandra da Silva Rodrigues

Mestrado em Bioquímica
Especialização em Bioquímica Médica

Dissertação orientada por:
Doutor Peter Jordan e Doutora Vânia Gonçalves

2018

Acknowledgments

I would like to use this chapter to express how grateful I am for having had this amazing internship opportunity at the Oncobiology and Signalling Pathways Lab from the Department of Human Genetics at the National Health Institute Dr. Ricardo Jorge, and for having had the chance to meet so many wonderful people and professionals who led me through this internship period.

I will start by expressing my deepest gratitude and special thanks to National Health Institute Dr. Ricardo Jorge and to my supervisors, doctors Peter Jordan and Vânia Gonçalves, for having given me this opportunity to learn and grow professionally, in such a great work environment. Vânia, thank you for having had the patience and care to teach and guide me through all the experimental work. I am so grateful for having had such a nice and awesome person as my supervisor, I truly learned a lot from you, and for that, I will always be thankful. I would also like to thank Dr. Peter Jordan and Dr. Paulo Matos whom, in spite of being extraordinarily busy with their duties, always took time out to hear me out, to guide and keep me on the correct path. I feel very lucky for having had such kind, nice and understanding people to guide me through this work. I would also like to express my gratitude towards Patrícia, you are an amazing person who gives out great professional and personal advice. Thank you for always being so willing to help, for caring and for being such a great person! A huge thank you to the rest of girls at the lab, Andreia, Joana, Márcia, Ana and Cláudia, for always being so kind, welcoming and willing to help. All of you showed me that working can also be fun, in a professional way of course!

I also want to express my deepest gratitude towards the most important people in my life, my parents, for making this opportunity possible and for always believing in me even when I didn't. I wouldn't be here without all of your hard work, support and encouragement! Thank you for being such amazing parents! I also need to acknowledge the crucial role of my closest family and friends during this stage of my live. I appreciate how my family continuously supported my choices and how they were always there for me when I needed it. From my friends, I am especially grateful for Joana, Raquel, Madalena and Catarina, whom have been putting up with me for a while now! Thank you for being such great friends, without you my academic experience wouldn't have been the same!

I perceived this opportunity as a big milestone in my career development. I will try to use my gained skills and knowledge in the best possible way, and I will continue to work on their improvement.

Resumo

A expressão génica é o mecanismo pelo qual a informação codificada num gene é convertida num produto funcional. A regulação deste processo permite que as células expressem genes diferencialmente consoante o seu tipo, fase de desenvolvimento ou mesmo em resposta a estímulos externos. Portanto, através da regulação da expressão génica, as células conseguem ativar genes seletivamente dependendo das suas necessidades e funções. Um dos processos cruciais envolvidos neste processo regulatório é o *splicing* do pré-mRNA, que consiste na remoção dos intrões e junção dos exões numa sequência codificante contígua. Esta reação de *splicing* é catalisada pelo spliceossoma, um complexo ribonucleoproteico que reconhece e interage com as sequências de consenso nos limites dos exões e dos intrões, de forma a direcionar a excisão dos intrões e a ligação dos exões do RNA. Na maioria dos genes humanos a inclusão ou exclusão dos exões no transcrito final ocorre de forma diferencial, tornando-se assim possível a produção de múltiplos mRNAs diferentes a partir de um único gene. Este processo é denominado de *splicing* alternativo e é um dos mecanismos que modula a regulação da expressão proteica e a produção de um proteoma complexo e diversificado em eucariotas mais complexos. Durante o processo de *splicing* alternativo a decisão de qual exão é removido e qual é incluído no mRNA final, para além das sequencias de consenso é também fortemente influenciada pela interação entre elementos regulatórios *cis* e os fatores *trans*. Os elementos *cis* ocorrem tanto nas regiões exónicas como nas regiões intrónicas e podem promover a inclusão do exão através do recrutamento do spliceossoma (sequências promotoras) ou promover a sua exclusão por interferência com a ligação do spliceossoma às sequências de consenso (sequências silenciadoras). As sequências promotoras são, geralmente, ligadas por fatores que atuam em *trans* positivos, como é o caso das proteínas SR, enquanto que as sequências silenciadoras são normalmente ligadas por fatores que atuam em *trans* negativos, como é o caso das proteínas hnRNPs. Para além da presença de elementos regulatórios *cis* e fatores *trans*, uma variedade de outros fatores podem influenciar o *splicing* alternativo, como é o caso da presença de estruturas secundárias no mRNA, da presença de microRNAs, da arquitetura dos exões e intrões, da força relativa das sequências de consenso no local de *splicing* e ainda da velocidade de elongação durante a síntese do pré-mRNA pela RNA polimerase II. A soma das contribuições de cada um destes elementos define o potencial de reconhecimento de um exão e a sua respetiva afinidade pelo spliceossoma. Devido ao seu papel central na expressão e função de proteínas, é expectável que problemas ao nível da regulação do *splicing* alternativo possam resultar no desenvolvimento de doenças. Um exemplo disso é a sobreexpressão da variante de *splicing* hiperativa do gene RAC1, RAC1b, em diversos tumores malignos. A variante RAC1b é caracterizada pela inserção de um exão (exão 3b) extra entre os exões 3 e 4 de RAC1. Estes 19 aminoácidos extra codificados pelo exão 3b, para além de uma regulação diferente também conferem uma seletividade na sinalização a jusante de RAC1b. Esta variante promove a produção de ROS e a via de sinalização do NF-kB em detrimento de outras vias clássicas ativadas por RAC1, promovendo a progressão do ciclo celular e a sobrevivência das células. Em cancro colorretal, a variante RAC1b encontra-se sobre expressa num subgrupo específico de tumores que também apresentam uma mutação oncogénica em BRAF ($BRAF^{V600E}$), tendo sido demonstrado a existência de uma cooperação entre estes eventos no sentido de promover a sobrevivência das células tumorais. Até agora, sabe-se que o *splicing* alternativo de RAC1 é regulado por duas proteínas SR, SRSF1 e SRSF3. SRSF1 promove a inclusão do exão alternativo 3b, ao contrário de SRSF3 que promove a sua exclusão. Ambos os fatores são regulados por vias de sinalização a montante, nomeadamente, os níveis proteicos de SRSF1 aumentam quando a via de sinalização PI3K é inibida, enquanto que a via β-catenin/TCF4 estimula a expressão de SRSF3. É provável que existam outros elementos que regulem o *splicing* alternativo de RAC1, e recentemente, as proteínas PTBP1 e ESRP1 foram descritas como possíveis modeladoras deste evento em diferentes tipos de células, enquanto que a nucleoporina RANBP2 foi relacionada com a distribuição de proteínas SR fosforiladas, as quais são responsáveis

pela regulação da expressão de RAC1b. Neste trabalho experimental foram estudados estes três possíveis modeladores do *splicing* alternativo de RAC1 em células colorretais.

Para estudar estes possíveis novos mecanismos de regulação da expressão de RAC1b em células colorretais, começámos por construir vetores de expressão para PTBP1 e ESRP1, sendo que o vetor de expressão para RANBP2 já se encontrava disponível no laboratório de acolhimento. Posteriormente, os possíveis efeitos da sobreexpressão de PTBP1, de ESRP1 e de RANBP2 no *splicing* alternativo de RAC1 foram determinados através do uso de um minigene RAC1. Basicamente, cada plasmídeo que codificava as proteínas em estudo foi co-transfetado em células com o minigene RAC1. Os resultados foram observados através de um RT-PCR semi-quantitativo, com *primers* específicos para os transcritos RAC1 e RAC1b derivados do minigene RAC1. A expressão endógena de RAC1b foi também avaliada por *Western blot* (WB) através do uso de um anticorpo específico contra RAC1b. De acordo com os resultados, os efeitos significativos que foram observados para ESRP1 e RANBP2 foram confirmados ao nível endógeno, através do uso de siRNAs comercialmente disponíveis de forma a silenciar a sua expressão. Os resultados foram mais uma vez obtidos através de um RT-PCR semi-quantitativo, mas desta vez os *primers* utilizados foram específicos para os transcritos endógenos. Os níveis proteicos de RAC1b endógeno foram também avaliados através de WB. As experiências foram realizadas principalmente em células epiteliais normais do cólon, NCM460, mas confirmadas com células HeLa e HT29 para determinar se os resultados observados eram dependentes da linha celular. A localização celular das proteínas transfetadas foi visualizada através de ensaios de imunofluorescência em células NCM460 e HeLa.

A sobreexpressão de PTBP1, nas células NCM460, não afetou significativamente a inclusão do exão 3b no transcrito final. No entanto, a sobreexpressão de ESRP1 e de RANBP2 promoveu a exclusão do exão 3b. As experiências de imunofluorescência mostraram que a expressão de PTBP1 e ESRP1 ocorre tanto no núcleo como no citoplasma, enquanto que a localização de RANBP2 se restringe essencialmente à membrana nuclear. Esta observação vai de encontro com o esperado, dado que PTBP1 e ESRP1 são ambos fatores de *splicing*, enquanto que RANBP2 faz parte do complexo do poro nuclear. Nas células NCM460, o silenciamento de ESRP1 diminuiu a expressão endógena de RAC1b, enquanto que a depleção de RANBP2 levou ao aumento de RAC1b. O efeito da depleção de ESRP1 no *splicing* de RAC1 não corroborou os resultados obtidos nas experiências de sobreexpressão, dando assim a indicação de que este fator de *splicing*, devido ao seu papel na manutenção do fenótipo epitelial, pode ser altamente regulado por um mecanismo de *feedback* negativo, no qual regula a sua própria expressão. Os resultados obtidos para as células HeLa e HT29 seguiram a mesma tendência observada nas células NCM460. Assim, no geral, o silenciamento de ESRP1 diminuiu a expressão endógena de RAC1b, enquanto que a depleção de RANBP2 levou ao seu aumento, ambos independentemente da linha celular utilizada, sugerindo que tanto ESRP1 como RANBP2 são reguladores gerais da expressão de RAC1b.

Em conclusão, esta tese forneceu fortes evidências de que ESRP1 e RANBP2 estão envolvidos na regulação da expressão de RAC1b e identificou pela primeira vez outros fatores, para além de SRSF1 e de SRSF3, envolvidos na regulação da expressão de RAC1b em células colorretais. Mais experiências são necessárias para esclarecer como é que estas proteínas estão a regular o *splicing* alternativo de RAC1. Este conhecimento será útil na caracterização dos mecanismos de regulação de *splicing* alternativo de RAC1 e, eventualmente, para desenvolver moduladores farmacológicos eficazes que possam restaurar a sinalização normal de RAC1 em células tumorais.

Palavras chave: *Splicing* alternativo; RAC1b; ESRP1; RANBP2; Cancro colorretal

Abstract

Regulation of gene expression allows cells to differentially express genes in different cell types, developmental stages or even in response to external conditions. Alternative splicing is a crucial regulatory process in the pathway of gene expression and the mechanism by which multiple protein isoforms can be generated from a single gene. This way, complex organisms can regulate protein expression and generate a more diverse proteome from a given gene number within the genome. Due to its central role in modulating gene expression, it is not surprising that aberrant regulation of alternative splicing is associated with human disease. On example is the selective overexpression of a hyperactive splice variant of the RAC1 gene, RAC1b, in several malignant tumours. RAC1b promotes reactive oxygen species production and the NF-kB pathway activation, but not other classical RAC1 signalling pathways, and stimulates cell cycle progression and cell survival. In colorectal cancer, RAC1b was found to be overexpressed in a specific subgroup, namely in 80% of tumours with mutation in the oncogene BRAF, suggesting that both events cooperate to promote the survival of colorectal cells. Previous studies in colorectal cells showed that the splicing factor SRSF1 acts as an enhancer of RAC1 alternative splicing by promoting the inclusion of alternative exon 3b, while SRSF3 acts as a silencer by promoting the skipping of the exon 3b. Besides SRSF1 and SRSF3 it is likely that RAC1 alternative splicing can be regulated by additional factors, and recently, PTBP1 and ESRP1 were described as possible modulators of the alternative splicing of RAC1b in different cell types while RANBP2 was shown to be a modulator of the distribution of phosphorylated SR proteins, which are known to regulate RAC1b expression.

To study whether these factors regulate RAC1b expression in colorectal cells, we started by analysing the effects of PTBP1, ESRP1 and RANBP2 overexpression on RAC1 alternative splicing. For that, each expression vector encoding the proteins in study was co-transfected with the RAC1 minigene into cells. The results were then assessed through a semi-quantitative RT-PCR, with specific primers for RAC1 and RAC1b transcripts derived from the RAC1 minigene. Effects on endogenous RAC1b expression was also assayed by Western blot and cellular localization of the transfected proteins assessed by immunofluorescence. Positive effects were then confirmed at the endogenous protein level through the use of commercially available siRNAs to deplete the regulators. The experiments were mainly performed in NCM460 colon cells but confirmed using HeLa and HT29 cells to determine if the observed results were cell line dependent. As expected, immunofluorescence experiments showed that both PTBP1 and ESRP1 can be found in the nucleus and cytoplasm, while RANBP2 is found at the nuclear membrane and also at the cytoplasm. RANBP2 overexpression was found to promote the skipping of the alternative exon 3b, while its depletion promoted exon 3b inclusion. Thus, RANBP2 emerged as a candidate negative regulator of RAC1 alternative splicing that promotes the skipping of exon 3b. In the case of ESRP1, both overexpression and depletion promoted the skipping of the alternative exon 3b. ESRP1 overexpression might interfere with a negative feedback mechanism, in which ESRP1 regulates its own expression, due to its role in maintaining the epithelial phenotype. Based on the depletion experiment, ESRP1 can also be considered a candidate positive regulator of RAC1 alternative splicing, which promotes the inclusion of exon 3b, unlike RANBP2. No significant effect of PTBP1 on RAC1 alternative splicing was observed.

In conclusion, this thesis provided strong evidence that ESRP1 and RANBP2 are involved in RAC1b expression regulation and identified for the first time other factors besides SRSF1 and SRSF3 that are involved in the regulation of RAC1b expression in colorectal cells. Further experiments are needed to clarify how these proteins are regulating RAC1 alternative splicing. This knowledge will be useful to characterize RAC1 alternative splicing regulation mechanisms and eventually to develop effective pharmacological modulators that can restore normal RAC1 signalling in tumour cells.

Keywords: Alternative Splicing; RAC1b; ESRP1; RANBP2; Colorectal cancer

Table of contents

Acknowledgments ... iii
Resumo ... iv
Abstract ... vi
List of tables and figures ... viii
List of abbreviations ... ix
1. Introduction ... 1
1.1. Molecular mechanism of pre-mRNA Splicing ... 1
1.2. Regulation of pre-mRNA splicing ... 4
1.3. Effects of alternative splicing ... 6
1.4. The small GTPase RAC1 and its splice variant RAC1b ... 7
1.5. Colorectal cancer and RAC1b alternative splicing ... 9
1.6. PTBP1, ESRP1 and RANBP2 as possible modulators of RAC1b alternative splicing ... 10
2. Objectives ... 11
3. Experimental Procedures ... 12
3.1. Polymerase Chain Reaction (PCR) ... 12
3.1.1. PTBP1 and ESRP1 amplification ... 13
3.1.2. Agarose gel electrophoresis ... 13
3.2. Cloning ... 14
3.2.1. PTBP1 and ESRP1 cloning ... 14
3.2.4. Sequencing ... 17
3.3. Cell Culture ... 18
3.4. Cell Transfection ... 20
3.5. Cell lysis ... 24
3.6. SDS-PAGE and Western blot ... 25
3.7. Semi-quantitative RT-PCR ... 28
3.8. Immunofluorescence assay ... 29
3.9. Data Treatment ... 30
4. Results ... 32
4.1. Production of expression vectors for PTBP1 and ESRP1 ... 32
4.2. The effect of PTBP1, ESRP1 and RANBP2 overexpression in the alternative splicing of RAC1 ... 33
4.2.1. Subcellular localization of PTBP1, ESRP1 and RANBP2 ... 36
4.3. The effect of ESRP1 and RANBP2 depletion in the alternative splicing of endogenous RAC1 transcripts ... 37
5. Discussion, Conclusions and Future Perspectives ... 42
6. References ... 46

List of tables and figures

Table 3.1 - Primers sequence used for the amplification of ESRP1 and PTBP1 cDNA sequences. 13
Table 3.2 - Composition of the reaction mixes for the digestion of the TOPO vectors, TOPO_PTBP1 and TOPO_ESRP1, and the pcDNA3_myc vector. 15
Table 3.3 - Primers and PCR programs used for the screening of the integration of PTBP1 and ESRP1 into the pCR™2.1-TOPO® vector and the pcDNA3_myc expression vector. 17
Table 3.4 - Primers used to sequence the vectors cloned with ESRP1, PTBP1 and RANBP2. 18
Table 3.5 - Information on every cell line used on this experimental work, including their origin, type and culture conditions. 20
Table 3.6 - Co-transfection conditions of pcDNA3_EGFP, pcDNA3_myc_PTBP1, pcDNA3_myc_ESRP1 or pkTol2Chy_RANBP2_1-3224 with RAC1 minigene (MG) in NCM460, HeLa and HT29 cell lines. Transfection conditions of pcDNA3_EGFP or pkTol2Chy_RANBP2_1-3224 in NCM460, Caco2, HT29 and HEK 293 cell lines. All transfections and co-transfections were performed in 6-well plates..... 23
Table 3.7 - Transfection conditions of siGFP (Eurofins Genomics), siESRP1 (Santa Cruz Biotechnology) and siRANBP2 (Santa Cruz Biotechnology) for NCM460, HeLa and HT29 cell lines, all performed in 24-well plates..... 24
Table 3.8 - Cell lysis conditions for every transfection preformed in this experimental work..... 25
Table 3.9 - Recipes of the polyacrylamide gels..... 26
Table 3.10 - Antibodies used in Western blot and respective dilutions..... 27
Table 3.11 - Primers and PCR programs used for the semi-quantitative RT-PCRs of RAC1b from the minigene (MG), endogenous RAC1b, ESRP1 and RANBP2. 29

Figure 1.1 – Consensus sequences that define exon/intron boundaries..... 1
Figure 1.2 – Schematic representation of the transesterification steps of RNA splicing 2
Figure 1.3 – Spliceosome assembly and catalysis..... 2
Figure 1.4 – Schematic representation of the five main types of alternative splicing events..... 3
Figure 1.5 – Schematic representation of regulatory cis-elements..... 5
Figure 1.6 – Schematic representation of RAC1 activation and regulation 8
Figure 1.7 – Schematic representation of the human RAC1 gene..... 8
Figure 3.1 – Western blot transference..... 27
Figure 3.2 – Antigen detection methods..... 30
Figure 4.1 – PTBP1 and ESRP1 cDNA amplification 32
Figure 4.2 – PCR screenings. 33
Figure 4.3 – RAC1 minigene..... 34
Figure 4.4 – Western blot analysis of the transfected tagged-proteins expression..... 34
Figure 4.5 – Analysis of RAC1 exon 3b inclusion after PTBP1, ESRP1 and RANBP2 overexpression in NCM460 cells 35
Figure 4.6 – Analysis of RAC1 exon 3b inclusion after PTBP1, ESRP1 and RANBP2 overexpression in HeLa cells..... 36
Figure 4.7 – Subcellular localization of transfected PTBP1, ESRP1 and RANBP2..... 37
Figure 4.8 – Time course experiment of siRNA-mediated depletion of ESRP1 and RANBP2..... 38
Figure 4.9 – Analysis of endogenous RAC1b expression after depletion of ESRP1 and RANBP2 in NCM460 cells 39
Figure 4.10 – ESRP1 and RANBP2 depletion in NCM460, HeLa and HT29 after 72 h incubation ... 40
Figure 4.11 – Analysis of endogenous RAC1b expression in NCM460, HeLa and HT29 cells..... 41

List of abbreviations

A	Adenosine
APS	Ammonium persulfate
ATP	Adenosine triphosphate
BAP	Bacterial alkaline phosphatase
bp	Base pairs
BPS	Branch-point site
BSA	Bovine serum albumin
C	Cytosine
cDNA	mRNA-complementary deoxyribonucleic acid
CO_2	Carbon dioxide
CRC	Colorectal cancer
CTD	Carboxy terminal domain
DAPI	4',6-diamidino-2-phenylindole
ddNTP	Chain-terminating dideoxy nucleotide
DMEM	Dulbecco's Modified Eagle's Medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide
dsDNA	Double-stranded DNA
DTT	Dithiothreitol
ECL	Enhanced Chemiluminescence
EDTA	Ethylenediamine tetraacetic acid
EMT	Epithelial-to-mesenchymal transition
ERK	Extracellular signal-regulated kinase
ESE	Exonic splicing enhancer
ESRP1	Epithelial Splicing Regulatory Protein 1
ESS	Exonic splicing silencer
EtBr	Ethidium bromide
FBS	Foetal bovine serum
G	Guanosine
GAPs	GTPase-activating protein
GDP	Guanosine diphosphate
GEF	Guanine nucleotide exchange factor
GFP	Green fluorescent protein
GSK3β	Glycogen synthase kinase-3 beta
GTP	Guanosine triphosphate
GTPases	Guanosine triphosphate phosphohydrolases
H_2O_2	Hydrogen peroxide
HCl	Hydrochloric acid
hnRNP	Heterogeneous nuclear ribonucleoprotein
HRP	Horseradish peroxidase
IF	Immunofluorescence assay
IRES	Internal ribosome entry site
ISE	Intronic splicing enhancer
ISS	Intronic splicing silencer
JNK	c-Jun N-terminal kinase

kb	Kilobase
KCl	Potassium chloride
kDa	Kilodalton
KH_2PO_4	Monopotassium phosphate
LB	Luria Bertani
mA	Milliampere
$MgCl_2$	Magnesium chloride
mL	Millilitre
mM	Millimolar
mRNA	Messenger ribonucleic acid
Myc	Epitope tag derived from the c-myc gene product
$Na_2HPO_4.2H_2O$	Sodium phosphate dibasic dihydrate
NaCl	Sodium Chloride
ncRNA	Non-coding RNAs
NF-kB	Nuclear factor kappa-light-chain-enhancer of activated B cells
ng	Nanogram
nm	Nanometre
NMD	Nonsense-mediated mRNA-decay
NP-40	Nonidet P-40
NPC	Nuclear pore complexes
^oC	Degree Celsius
OIS	Oncogene-induce senescence
PAGE	Polyacrylamide gel electrophoresis
PAK	p21 activating kinase
PBS	Phosphate-buffered saline
PCR	Polymerase Chain Reaction
PFA	Formaldehyde
PI3K	Phosphatidylinositol 3 kinase
Pol II	RNA polymerase II
pre-mRNA	Precursor mRNA
PTBP1	Polypyrimidine tract binding protein 1
PTC	Premature termination codons
PVDF	Polyvinylidene fluoride membrane
RAC1	Ras-related C3 botulinum toxin substrate 1
RAC1MG	RAC1 minigene
RANBP2	RAN binding protein 2
Rho-GDI	Rho-GDP dissociation inhibitor
RISC	RNA-induced silencing complex
RNA	Ribonucleic acid
ROS	Reactive oxygen species
RRM	RNA recognition motif
RT-PCR	Reverse transcription polymerase chain reaction
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SF1	Splicing factor 1
siRNA	Short/small interfering ribonucleic acid
snRNPs	Small nuclear ribonucleoproteins
SR proteins	Serine/arginine-rich protein

SRPK1	Serine/arginine-rich protein-specific kinase 1
SRSF1	Serine/arginine-rich splicing factor 1
SRSF3	Serine/arginine-rich splicing factor 3
ssDNA	Single-stranded deoxyribonucleic acid
T	Thymidine
TBE	Tris-Borate-EDTA
TBS	Tris-buffered saline
TCF4	Transcription factor 4
TEMED	Tetramethylethylenediamine
Tm	Melting temperature
Tris	Tris(hydroxymethyl)aminomethane
U	Uridine
U2AF	U2 auxiliary factor
UTRs	Untranslated regions
UV	Ultraviolet
V	Volt
v/v	Volume per volume
WB	Western blot
Wnt	Wingless
X-Gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside
µg	Microgram
µL	Microliter

1. Introduction

Gene expression is a complex mechanism by which the information encoded in a gene is converted into a functional product. In the case of protein coding genes, this process consists in the transcription of the gene into a precursor mRNA (pre-mRNA) that is processed (originating a functional mRNA) and transported to the cytoplasm, where it is translated into protein (Lodish 2016). Regulation of this multistep process allows cells to differentially express genes in different cell types, developmental stages or even in response to external conditions (Lodish 2016). Thus, by regulating gene expression cells can activate genes selectively depending on their needs and purpose. Pre-mRNA splicing, one of the RNA processing steps, is a crucial regulatory point in the pathway of gene expression and will be described in more detail in this work.

1.1. Molecular mechanism of pre-mRNA Splicing

In most eukaryotic genes, coding regions (exons) are interrupted by noncoding regions (introns), and during transcription both are copied into the pre-mRNA (Matlin et al. 2005; Lodish 2016). One of the steps of RNA processing is pre-mRNA splicing, where introns are removed, and exons are joined to form a contiguous coding sequence (Matlin et al. 2005; Lodish 2016). The pre-mRNA splicing reaction is directed by four conserved sequences that define the exon/intron boundaries (**Figure 1.1**). These four consensus sequences include: the exon–intron junction at the intron's 5′ end (**GU**, 5′ splice site or splice donor site); the exon–intron junction at the intron's 3′ end (**AG**, 3′ splice site or splice acceptor site); the branch-point site (BPS) sequence; and the polypyrimidine tract (**Figure 1.1**) (Matlin et al. 2005).

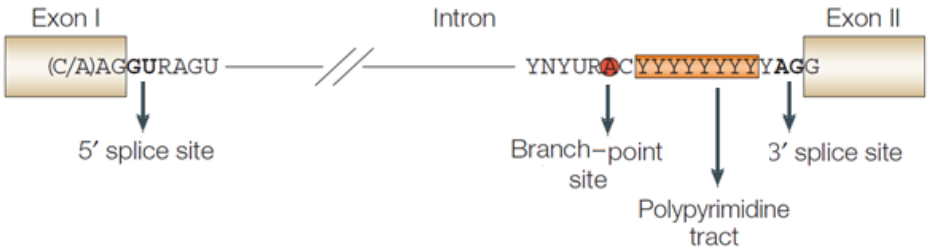

Figure 1.1 – Consensus sequences that define exon/intron boundaries. Y=U or C; R=G or A. (adapted from (Matlin et al. 2005))

Succinctly, the splicing reaction comprises two consecutive steps of transesterification (Papasaikas and Valcárcel 2016). In the first step, a nucleophilic attack on the phosphate group between the 5' exon–intron junction is carried out by the 2' hydroxyl group of an adenosine residue from the BPS, generating a 2'-5' phosphodiester bond and, consequently, a lariat intermediate. In the second step, the free 3' hydroxyl of the 5' exon attacks the phosphate group between the intron and the 3' exon, splicing the two exons together and releasing the intron lariat (**Figure 1.2**) (Papasaikas and Valcárcel 2016). This reaction is catalysed by the spliceosome, a ribonucleoprotein complex composed of five small nuclear ribonucleoprotein complexes (snRNPs) - U1, U2, U4, U5 and U6 - and several other non-snRNPs associated factors (House and Lynch 2008; Chen and Manley 2009; Wang et al. 2015). The five snRNPs are assembled around one small nuclear RNA, U1 - U6 snRNAs, that use RNA-RNA base pairing to direct the snRNPs to the consensus sequences at the exon/intron boundaries (**Figure 1.1**) and direct RNA excision and ligation. Several snRNPs also interact with each other to ensure that the distant regions of the substrate required for splicing catalysis are correctly juxtaposed (House and Lynch 2008; Chen and Manley 2009; Wang et al. 2015). The process of pre-mRNA splicing begins with the assembly of complex E, which is defined by the base pairing of U1 snRNA to the 5′ splice site, the binding of splicing factor 1 (SF1) to the branch site, in an ATP-

independent manner, and the recruitment of U2 auxiliary factor (U2AF) heterodimer (comprising U2AF65 and U2AF35) to the polypyrimidine tract and 3′ splice site. This complex is converted into the ATP-dependent pre-spliceosome A complex by the replacement of SF1 by U2 snRNP at the branch point. Further recruitment of the U4/U6–U5 tri-snRNP leads to the formation of the B complex, which contains all spliceosomal subunits that carry out pre-mRNA splicing. Finally, the C complex assembles by extensive conformational changes and remodelling of both the snRNPs and the protein components that are present in the B complex, including the loss of U1 and U4 snRNPs, to produce an active site that is capable of catalysing the transesterification chemistry required for exon ligation and lariat release (**Figure 1.3**) (Chen and Manley 2009; Wang et al. 2015; Wang et al. 2015).

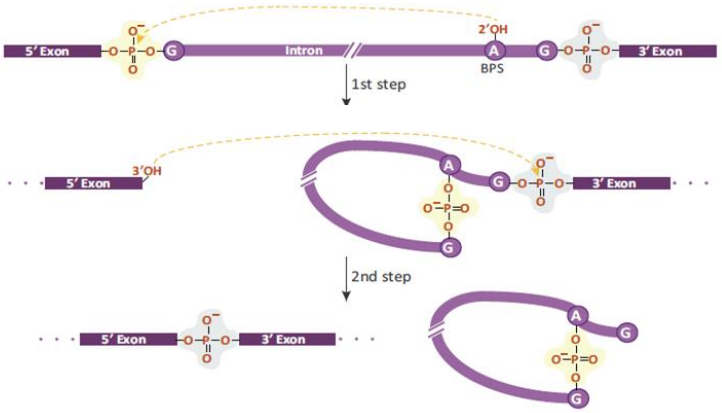

Figure 1.2 – Schematic representation of the transesterification steps of RNA splicing. Each step involves a nucleophilic attack on the terminal phosphodiester bonds of the intron. In the first step this is carried out by the 2′ hydroxyl of the branch point site (BPS) and in the second step by the 3′ hydroxyl of the 5′ exon. (adapted from (Papasaikas and Valcárcel 2016))

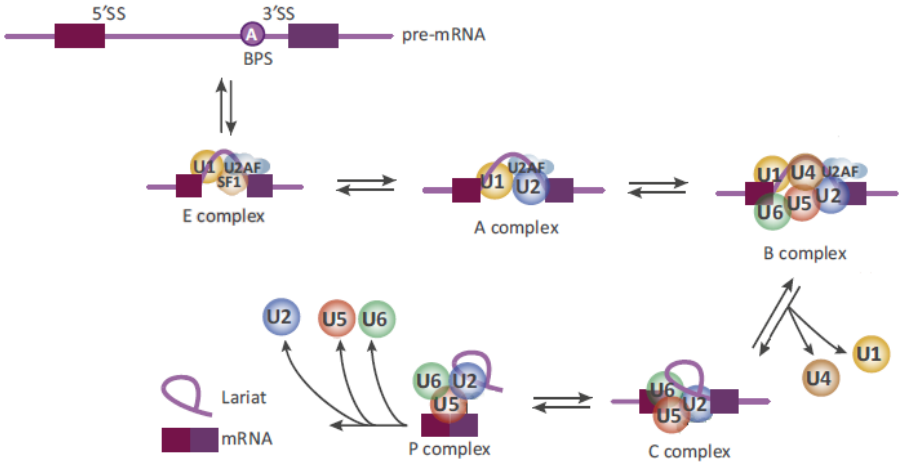

Figure 1.3 – Spliceosome assembly and catalysis. Canonical stepwise assembly of the spliceosome on pre-mRNA substrates indicating the transitions from the E through P complexes. 5′SS: 5′ Splice site; BPS: Branch point site; 3′SS: 3′ Splice site; snRNPs: U1, U2, U4, U5 and U6; U2AF: U2 auxiliary factor; SF1: Splicing factor 1. (adapted from (Papasaikas and Valcárcel 2016))

The process described above is referred to as constitutive splicing (**Figure 1.4A**). With the recent technologic advances allowing genome-wide transcript analysis, it has become clear that more than 90% of human genes can generate mRNAs with a differential inclusion or exclusion of exons in the final mRNA product (Poulos et al. 2011). This process is called alternative splicing (House and Lynch 2008) and is a mechanism by which multiple protein isoforms can be generated from a single gene. This process is predicted to occur in most mammalian genes, being an important mechanism by which complex organisms can regulate protein expression and generate a more diverse proteome from a given gene number within the genome (House and Lynch 2008; Wang et al. 2015).

There are 7 main types of alternative splicing that can lead to the production of different isoform transcripts: (1) cassette exon (skipped exon); (2) intron retention; (3) mutually exclusive exons; (4) alternative 3' splice site; (5) alternative 5' splice site; (6) alternative first exon and (7) alternative last exon (**Figure 1.4 B-H**) (Wang et al. 2015; Iñiguez and Hernández 2017). The most common alternative splicing pattern type is the cassette exon (skipped exon), with a prevalence of approximately 30% (Wang et al. 2015). This splice event occurs when one exon is spliced out of the primary transcript together with its flanking introns (Koscielny et al. 2009). In human transcripts intron retention occurs mainly in the untranslated regions (UTRs) and has been associated with weaker splice sites and short intron length (Wang et al. 2015). This pattern happens when a reported intron, or a part of it, is not removed and retained as part of the mature mRNA (Iñiguez and Hernández 2017). The splicing event known as mutually exclusive exons refers to the case in which two consecutive exons are never simultaneously included in the mature mRNA transcript (Koscielny et al. 2009). The alternative 3' splice site pattern, also known as alternative acceptor sites, corresponds to the case in which two or more splice sites are available near the 5' end of an exon, resulting in the use of an alternative 3' splice junction (acceptor site), therefore changing the 5' boundary of the downstream exon (Koscielny et al. 2009). The alternative 5' splice site pattern also called alternative donor sites, occurs when two or more splice sites are recognized near the 3' end of an exon, resulting in the use of an alternative 5' splice junction (donor site), therefore changing the 3' boundary of the upstream exon (Koscielny et al. 2009). In alternative first exon, two or more first exons exist that are mutually exclusive and spliced to the same second exons in each variant. This might be explained by the usage of an alternative promoter; however, additional biological evidence is needed to support this theory (Koscielny et al. 2009). In alternative last exon, the penultimate exon of each splice variant is identical, but one of mutually exclusive last exons is chosen. This pattern may result in an alternative polyadenylation usage or the differential presence of miRNA binding sites in the resulting mature mRNA (Koscielny et al. 2009). It is important to mention that each of the types of alternative splicing summarized above can occur within both translated and untranslated regions of transcripts. This means that alternative exons within 5’ and 3’ UTRs can either add or remove RNA regulatory motifs and, thereby, modulate the stability and translation of transcripts (Iñiguez and Hernández 2017).

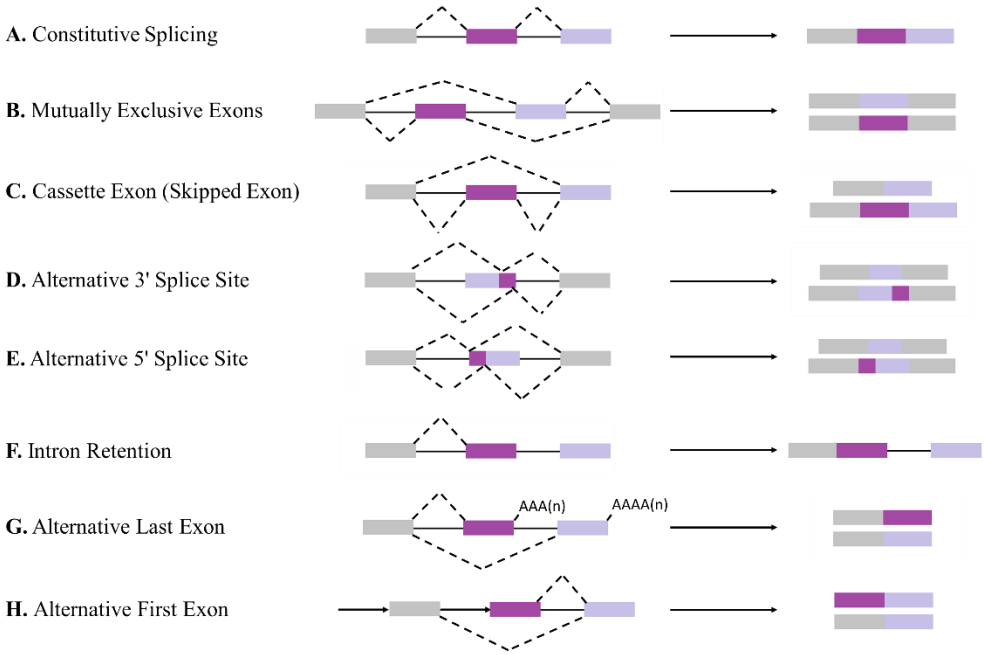

Figure 1.4 – Schematic representation of the five main types of alternative splicing events. (A) Constitutive splicing; (B) mutually exclusive exons; (C) cassette alternative exon; (D) alternative 3' splice site; (E) alternative 5' splice site; and (F) intron retention.

1.2. Regulation of pre-mRNA splicing

As described above, consensus sequences direct the spliceosome assembly on the pre-mRNA. However, in higher eukaryotes these sequence elements are highly degenerate, which indicates that additional marks are required for the spliceosome to identify the "real" splice sites among the numerous pseudo sites found in any pre-mRNA transcript (De Conti et al. 2013). At the most basic level, the relative strength of the splice sites (i.e., how near to consensus is their sequence) plays a major role in determining the inclusion or exclusion of exons (De Conti et al. 2013). Usually, the stronger the splice site, the more often it is used. However, a pair of 'strong' splice sites is not sufficient to define an exon; many 'pseudo-exons' that are flanked by predicted splice sites are not spliced or remain cryptic until specifically activated, e.g. in the above mentioned modes of 'alternative 3' splice site' or 'alternative 5' splice site'(De Conti et al. 2013; Kelemen et al. 2013). So, in order to efficiently recognize the "real" splicing sites, the spliceosome is further guided by supplementary cis-acting elements that recruit both positive and negative splicing regulatory factors (trans-acting factors), which either strengthen or weaken spliceosome interaction, in a process called exon definition (De Conti et al. 2013). These auxiliary splicing elements are highly variable in sequence, but they are important in defining both constitutive and alternative exons (Kelemen et al. 2013). Thus, the decision of which exon is removed, and which exon is included in the mature mRNA during the process of alternative splicing, is strongly influenced by the interaction between cis-acting elements and trans-acting factors (Hertel 2008; House and Lynch 2008; Kelemen et al. 2013; Wang et al. 2015). Cis-acting auxiliary sequences occur within both exonic and intronic regions and can either promote recruitment of the spliceosome and exon inclusion (splicing enhancers) or disrupt assembly of the splicing machinery and cause exon skipping (splicing silencers) (House and Lynch 2008; Wang et al. 2015). Depending on the position and function of the cis-regulatory elements, they are divided into four categories (**Figure 1.5**): exonic splicing enhancers (ESEs), exonic splicing silencers (ESSs), intronic splicing enhancers (ISEs) and intronic splicing silencers (ISSs) (Wang et al. 2015). Splicing enhancers (ESE and ISE) are, usually, bound by positive trans-acting factors, such as members of the SR (serine/arginine-rich) family of nuclear phosphoproteins, whereas splicing silencers (ESS and ISS) are, generally, bound by negative acting factors, such as members of the family of heterogeneous nuclear ribonucleoproteins (hnRNPs) (House and Lynch 2008; Kelemen et al. 2013; Wang et al. 2015). SR proteins promote exon inclusion by recruiting the splicing machinery to the adjacent intron, while hnRNPs repress spliceosomal assembly, for example, by multimerization along exons, looping out exons or by blocking the recruitment of snRNPs (House and Lynch 2008; Kelemen et al. 2013; Wang et al. 2015). SR proteins binding sites are present not only within alternatively spliced exons, but also within constitutively spliced exons, defining a crucial role in productive spliceosome assembly (House and Lynch 2008). Even though it is a general rule that SR proteins and hnRNPs promote or antagonize exon inclusion, respectively, there are numerous exceptions to this rule. For example, the splicing of the transcript encoding GTPase RAC1, is regulated by the antagonic effect of two SR-proteins, SRSF3 and SRSF1 previously known as SRp20 and ASF/SF2, respectively (Manley and Krainer 2010), whereas the myosin phosphatase targeting subunit-1 (MYPT1), is regulated by an antagonism between the two hnRNPs, PTB and TIA-1 (Shukla 2005; Gonçalves et al. 2009). In general, the cis-acting elements function additively and enhancing elements tend to play dominant roles in constitutive splicing, while the silencers are relatively more important in the control of alternative splicing (House and Lynch 2008; Kelemen et al. 2013; Wang et al. 2015). Typically, silencers and enhancers are present within the vicinity of potential exon/intron junctions, suggesting that the interplay between activating and repressing cis-acting elements modulates the probability of exon inclusion (House and Lynch 2008; Kelemen et al. 2013).

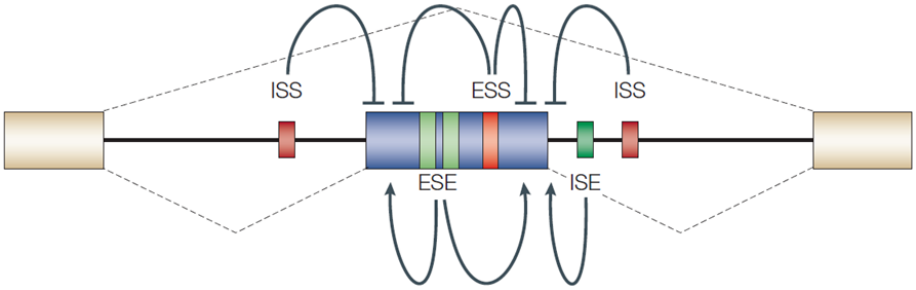

Figure 1.5 – Schematic representation of regulatory cis-elements. ESE: exonic splicing enhancer; ESS: exonic splicing silencer; ISE: intronic splicing enhancer; ISS: intronic splicing silencers. Enhancers can activate adjacent splice sites or antagonize silencers, whereas silencers can repress splice sites or enhancers. Exon inclusion or skipping is determined by the balance of these competing influences, which in turn might be determined by relative concentrations of the activator and repressor proteins. Dashed lines: Two alternative splicing pathways, with the middle exon either included or excluded. (from (Matlin et al. 2005))

Beside the presence of regulatory cis-elements, a variety of other factors can influence alternative splicing. First, the assembly of local RNA secondary structures has been shown to interfere with the recognition of splice sites and cis-acting sequences by protein factors (Hertel 2008; Kelemen et al. 2013). This is because the recognition of these elements by RNA-binding proteins depends on the single-stranded structure of the pre-mRNA. If the structure is not single-stranded, the sequence elements might be concealed within the stable helices of the secondary RNA structure, becoming unavailable to the protein factors, influencing, therefore, pre-mRNA splicing (Hertel 2008; Kelemen et al. 2013). Second, the expression level of key splicing factors in a given cell will affect the interplay between activating and repressing effects on the spliceosome. Besides transcriptional regulation of splicing factor genes, non-coding RNAs (ncRNAs), including microRNA and small interfering RNA, have also been shown to regulate alternative splicing, generally through the modulation of the expression of key splicing factors during development and differentiation (Luco and Misteli 2011).

Third, exon/ intron architecture of a gene also influences splice site recognition, which has been shown to be more efficient when introns or exons are small (Hertel 2008; De Conti et al. 2013). Exon skipping is promoted by big exons or large flanking introns. These observations suggested that splice sites are recognized across an optimal nucleotide length and predicted that intron length significantly influences the efficiency of pre-mRNA splicing and alternative splice site choice (Hertel 2008; De Conti et al. 2013).

Fourth, the type of promoter used to drive transcription by RNA polymerase II can also impact the level of alternative splicing of a downstream exon. Two non-exclusive models were proposed to explain this effect: the recruitment model and the kinetic model (Hertel 2008; Kelemen et al. 2013; Braunschweig et al. 2013). The recruitment model assumes splicing factors assemble at the carboxy terminal domain (CTD) of RNA polymerase II and are released onto the nascent pre-mRNA during transcription. As these factors influence splice sites in a concentration dependent manner, the pre-loading of the CTD influences alternative exon usage (Hertel 2008; Kelemen et al. 2013; Braunschweig et al. 2013). The kinetic model postulates that protein complexes need time to assemble on an exon during its recognition. Everything that slows down a polymerase would give more time for the recruitment of the regulatory complexes and would favour alternative exon usage, as these exons usually depend more strongly on auxiliary factors (Hertel 2008; Kelemen et al. 2013; Braunschweig et al. 2013). Fundamentally, these mechanisms influence patterns of alternative splicing via the variations in RNA polymerase II elongation and recruitment of splicing factors. These mechanisms also offer an attractive explanation for how epigenetic marks and chromatin structure can change alternative splicing. (Naftelberg et al. 2015) A very interesting case is the one of Hu proteins that can induce local histone hyperacetylation by association with their target sequences on the pre-mRNA

surrounding alternative exons of Nf1 and Fas genes. This hyperacetylation favours higher elongation rates, which in turn decreases exon 23a inclusion in the Nf1 transcript, generating a chromatin-mediated reinforcement of the primary splicing decision (Zhou et al. 2011).

Overall, every exon has a specific set of identity elements, including the strength of the splice sites, the presence or absence of splicing enhancers or silencers, the presence or absence of local RNA secondary structures, the exon/intron architecture and the process of pre-mRNA synthesis by RNA polymerase II, that permit its recognition by the spliceosome. The sum of contributions from each of these identity elements defines the overall recognition potential of an exon or the overall binding affinity for the spliceosome.

Different cell types or distinct biological processes such as the cell cycle, tissue differentiation or developmental stages, will exhibit fluctuations in the expression level, of spliceosomal components and splicing activator/repressors (Hertel 2008). A further layer is added by signal transduction pathways in response to extracellular signals, in which cells can regulate gene expression through the use of alternative splicing (Shin and Manley 2004; Blaustein et al. 2007; Gonçalves et al. 2017). This signal-mediated splicing regulation is operated through the activation of intricate networks of signal transduction pathways (Kalyna et al. 2012) that influence trans-acting splicing regulatory factors through post-translational modification, including protein phosphorylation. This can result in the alteration of their activity or cellular localization (Shin and Manley 2004; Blaustein et al. 2007; Gonçalves et al. 2017). Such alterations may influence splicing efficiency or induce alternative splicing (as described in 1.5). As a result, the same exon in these different scenarios has the same structural properties but the cellular recognition potential can vary, meaning that exons that are alternatively included in one cell type or biological process can be alternatively excluded in another (Hertel 2008). Essentially, regulation of alternative splicing can be achieved through modulating any one of the exon recognition components. The interplay between the several regulatory sequences, complexes, processes, pathways, and factors mention above establishes the presence of a combinatorial control in alternative splicing (Smith and Valcárcel 2000).

1.3. Effects of alternative splicing

An important consequence of alternative splicing is phenotypic complexity by increasing transcriptomic and proteomic diversity (Stamm et al. 2005; Hertel 2008; Cieply and Carstens 2015). The magnitude of these effects ranges from a complete loss of function or acquisition of a new function to very subtle modulations, due to variation in various aspects of protein function, including, their binding properties, intracellular localization, enzymatic activity, stability and regulation by posttranslational modifications (Stamm et al. 2005; Hertel 2008; Cieply and Carstens 2015). In addition to proteome modulation, alternative splicing is also a crucial regulatory stage in the pathway of gene expression. For example, as mentioned above, alternative exons within UTRs can modulate the stability and translation of transcripts. Up to one third of human alternative splicing events introduce premature termination codons (PTC), which are recognized and lead to the degradation of transcripts by the nonsense-mediated mRNA-decay (NMD) pathway (Lewis et al. 2003; Lareau et al. 2007a). Thus, the sensitivity of mRNA transcripts to NMD is modulated by alternative splicing events (Kalyna et al. 2012). One interesting example are the ‘poison exons’ found in transcripts from splicing factors. The inclusion of these exons introduces a PTC in the mRNA sequence, resulting in the transcript degradation by NMD. The inclusion or exclusion of this exon will indirectly influence the alternative splicing of the splicing factors targets (Lareau et al. 2007b; Rossbach et al. 2009).

Due to its central role in protein expression and function, alternative splicing seems to be an important mechanism in defining biological function. As so, it is not surprising that aberrant regulation of alternative splicing leads to human disease (Zhang and Manley 2013). Mutations in the consensus

splice site sequences, in the splicing regulatory sequences, in the splicing machinery and in the regulatory splicing factors genes have been suggested to cause aberrant splicing (Daguenet et al. 2015). In genetic diseases these types of mutation will interfere with the splice site recognition efficiency which can lead to exon skipping, intron retention or the introduction of a new splice site within an exon or intron (Daguenet et al. 2015). This deregulation in splicing has the potential to originate protein isoforms that ultimately contribute to human disease (Wang et al. 2015). Alterations in regulatory splicing factors' cellular concentration, composition, localization and activity have been suggested to be the cause of aberrant splicing, in multifactorial diseases, such as cancer (Wang et al. 2015). In genetic disease, an immediate cause for aberrant splicing is the alteration of the splice site recognition efficiency, while irregularities in protein isoforms in different systems ultimately contribute to multifactorial diseases (Wang et al. 2015). The potential roles for splicing in cancer are well documented and include changes in genes associated with cell migration, regulation of cell growth, hormone responsiveness, apoptosis and response to chemotherapy (Shkreta et al. 2013). Alternative splicing has been implicated in nearly all aspects of cancer development, thus, understanding the basic mechanisms and patterns of splicing in tumour progress will shed light on the biology of cancer and lay the foundation for diagnostic, prognostic and therapeutic(Kim and Kim 2012).

1.4. The small GTPase RAC1 and its splice variant RAC1b

Ras-related C3 botulinum toxin substrate 1 (RAC1) is a member of the Rac family of guanosine triphosphate phosphohydrolases (GTPases), a subfamily of the Rho family of small GTPases, which are best known for their role in regulating the actin cytoskeleton and gene expression (Marei and Malliri 2017). RAC1 exist in two different conformational states, an inactive GDP-bound form and an active GTP-bound form (Jordan et al. 1999; Matos et al. 2000). The interconversion between the two states occurs through a cycle of guanine exchange and GTP hydrolysis. The conformational changes induced upon GTP binding involve two important regulatory protein regions of RAC1, known as Switch I and Switch II. Consequently, the switch regions provide binding domains for both regulatory and effector proteins. In the active state, these regulatory regions enable their interaction with downstream effectors, allowing these GTPases to function as molecular switches (Wennerberg 2005). The transition between the active and inactive states of RAC1 occurs at the plasma membrane following appropriate cellular signals and is tightly controlled and spatially regulated by guanine nucleotide exchange factors (GEFs) which convert RAC1 to its active form, GTPase-activating proteins (GAPs), which inactivate RAC1, and Rho-GDP dissociation inhibitors (Rho-GDIs) that bind to and remove RAC1 from the plasma membrane, keeping it inactive in the cytoplasm and blocking its activation by GEFs (**Figure 1.6**) (Symons and Settleman 2000). Following its activation, RAC1 interacts with downstream effector proteins and activates signalling cascades that trigger various cellular responses such as secretory processes, phagocytosis of apoptotic cells, epithelial cell polarization, neuron adhesion, migration and differentiation, and growth-factor induced formation of membrane ruffles.

The mammalian RAC1 gene is composed of seven coding exons that after undergoing alternative splicing can originate two different transcripts, RAC1 and RAC1b (**Figure 1.7**) (Gonçalves et al. 2009). RAC1b is characterized by the insertion of an additional exon (exon 3b) between exons 3 and 4 of RAC1. Consequently, RAC1b contains 57 additional nucleotides that result in an in-frame insertion of 19 amino acid residues between codons 75 and 76 of RAC1, near an important regulatory region of the GTPase, the switch II domain, changing the regulation and signalling properties of the protein (Jordan et al. 1999).

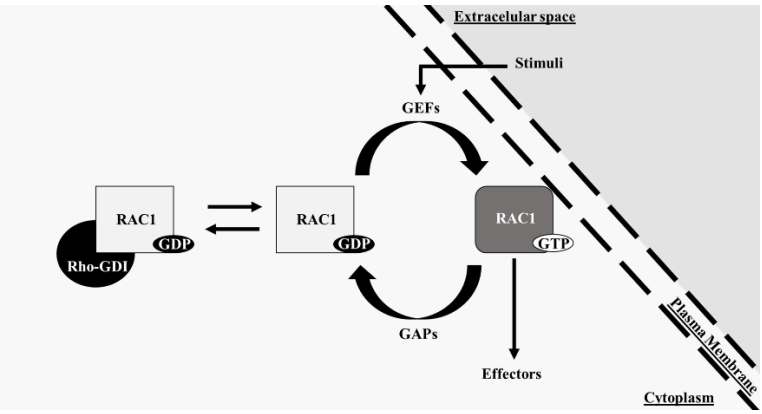

Figure 1.6 – Schematic representation of RAC1 activation and regulation. RAC1 GTPase cycles between inactive GDP-bound and active GTP-bound states. RAC1 activation is facilitated by the action of GEFs, which promotes GDP dissociation from RAC1 and allows GTP to bind instead. Through the association with GAPs, the intrinsic GTPase activity of RAC1 is accelerated, thereby inactivating RAC1. Through association with RhoGDIs RAC1 can be sequestered in its inactive state.

RAC1b is a highly activated variant of RAC1 because, despite the lower levels of expression compared to RAC1, RAC1b exists predominantly in the active GTP-bound state. This is essentially due to RAC1b disability to interact with Rho-GDI, which keeps this GTPase constitutively membrane-bound, a location that favours the interaction with activators, and consequently promotes the active GTP-bound state (Matos et al. 2003). Moreover, RAC1b shows impaired intrinsic GTPase activity and increased GDP to GTP exchange rates, although this variant can still be down regulated by activated GAPs and it is influenced by GEFs action (Matos et al. 2003; Fiegen et al. 2004; Singh et al. 2004).

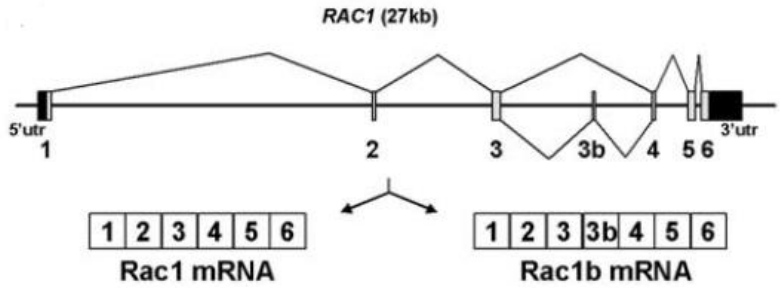

Figure 1.7 – Schematic representation of the human RAC1 gene. RAC1 gene has seven coding exons, including the alternative exon 3b, that after undergoing the alternative splicing event can originate two alternative transcripts, RAC1 and RAC1b. (from (Gonçalves et al. 2009))

In addition, the 19 extra amino acids of RAC1b seem to confer a selective downstream signalling to this variant, since several pathways activated by RAC1 are not activated by RAC1b (Matos et al. 2003). Unlike RAC1, GTP-bound RAC1b is unable to induce lamellipodia formation, which means that this isoform doesn't have the ability to induce actin cytoskeleton reorganization. Additionally, RAC1b is incapable to activate PAK1 effector and stimulate the JNK cascade, two well-established RAC1 signalling pathways (Matos et al. 2003; Fiegen et al. 2004; Singh et al. 2004). On the other hand, RAC1b was shown to bind more effectively than RAC1 to proteins that can promote loss of epithelial cell structure and increased cell proliferation, such as p120 catenin and RACK1 (Orlichenko et al. 2010). RAC1b also favours specific pathways conducting to the production of reactive oxygen species (ROS) and NF-kB canonical pathway activation (Matos et al. 2003; Matos and Jordan 2005; Radisky et al. 2005). This RAC1 isoform was found to be overexpressed in several malignant tumours including colorectal, breast, lung, thyroid and pancreas (Schnelzer et al. 2000; Matos and Jordan 2008; Stallings-Mann et al. 2012; Silva et al. 2013; Mehner et al. 2014).

1.5. Colorectal cancer and RAC1b alternative splicing

Colorectal cancer (CRC) is one of the most common types of cancer worldwide, representing in 2012 the third highest incidence, with 1.4 million cases, after lung and breast cancer (Ferlay et al. 2015). The main risk factors identified for the development of CRC include family histories of either colorectal cancer or inflammatory bowel disease, but the disease burden at the population-level is mainly accounted for by modifiable life-style factors such as smoking, excessive alcohol consumption, high consumption of red and processed meat, obesity, and diabetes (Brenner et al. 2014).

Normal gastrointestinal epithelium is organized along a crypt-villus axis. A pool of colon stem and progenitor cells, the most undifferentiated cell types that are able of self-renewal and pluripotency, are located at the bottom of the crypt. These cells differentiate along the crypt-villus axis, into all epithelial colon lineages. In about 14 days they arrive at the top of the villus and undergo programmed cell death (apoptosis) (Peifer 2002; Kosinski et al. 2007). During colorectal adenocarcinoma development, some progenitor cells acquire sequential genetic and epigenetic mutations in specific oncogenes and/or tumour suppressor genes, conferring them a selective advantage on proliferation and self-renewal (Vogelstein et al. 1988; Ewing et al. 2014). So, the normal epithelium becomes a hyperproliferative mucosa and subsequently gives rise to a benign adenoma that evolves into carcinoma and metastasis in about 10 years (Vogelstein et al. 1988).

In colorectal cancer, RAC1b is overexpressed in a specific subgroup, namely in 80% of tumours with mutation in the oncogene BRAF ($BRAF^{V600E}$), suggesting that both events cooperate to promote the survival of colorectal cells (Matos et al. 2008; Matos et al. 2016). Activating mutations in the BRAF gene have been found to induce oncogene-induce senescence (OIS), an important tumour suppressing mechanism at early stages of cancer. In colorectal cancer, RAC1b overexpression was found to antagonize OIS, suggesting that this hyperactive splice variant is selected in early stages of tumour development. Knowing that RAC1b expression was found to be increased in patients with inflammatory bowel diseases or in mouse models of acute colitis and that chronic inflammation is a known risk factor for colorectal cancer, the authors suggested that RAC1b overexpression in $BRAF^{V600E}$-initiated tumour cells could be triggering escape from OIS, leading to cancer progression (Henriques et al. 2015). In another study, RAC1b was reported to be a putative accelerator of tumour progression by positively regulating the expression of proliferation-promoting genes through Wnt pathway activation and decreasing the adhesive properties of colorectal cancer cells by negatively regulation of E-cadherin expression (Esufali et al. 2007). RAC1b overexpression was also associated with a poor outcome of patients with wild-type KRAS/BRAF metastatic colorectal cancer treated with FOLFOX/XELOX chemotherapy (Alonso-Espinaco et al. 2014). As a result of its hyperactive properties and selective overexpression in cancerous tissue, RAC1b has been highlighted as a promising therapeutic target, being, therefore, important to understand the basic mechanisms underlining the regulation of its expression.

Previous studies in this sense showed that in colorectal cells, RAC1 alternative splicing event is regulated by two SR proteins with antagonistic roles, SRSF1 and SRSF3 (Gonçalves et al. 2009). SRSF1 acts as an enhancer by promoting the inclusion of alternative exon 3b, while SRSF3 acts as a silencer by promoting the skipping of the exon 3b (Gonçalves et al. 2009). Both splicing factors were found to be regulated by upstream signalling pathways: the inhibition of the phosphatidylinositol 3-kinase pathway increased protein levels of SRSF1 and promoted RAC1b, whereas activation of β-catenin/TCF4 increased expression of SRSF3 and inhibited that of RAC1b (Gonçalves et al. 2009). Thus, extracellular stimuli might induce or sustain RAC1b overexpression in tumour cells through signal transduction pathways. For example, the protein kinases SRPK1 and GSK3β were also found required to sustain RAC1b levels and both were shown to act upon the phosphorylation of splicing factor SRSF1 (Goncalves et al. 2014). However, besides SRSF1 and SRSF3 it is likely that RAC1

alternative splicing can be regulated by additional factors, and recently, PTBP1 and ESRP1 were described as possible modulators of the alternative splicing of RAC1b in different cell types while RANBP2 was shown to be a modulator of the distribution of phosphorylated SR proteins, which are known to regulate RAC1b expression (Saitoh et al. 2012; Ishii et al. 2014; Hollander et al. 2016; Vecchione et al. 2016).

1.6. PTBP1, ESRP1 and RANBP2 as possible modulators of RAC1b alternative splicing

Polypyrimidine tract-binding protein 1 (PTBP1), also known as hnRNP I, is a member of a subfamily of ubiquitously expressed hnRNPs and contains four RNA recognition motif (RRM) domains that bind to the polypyrimidine track of mRNAs introns (Oberstrass 2005). PTBP1 shuttles between the nucleus and the cytoplasm, intervening in almost all steps of mRNA metabolism, such as alternative splicing, mRNA transport, cytoplasmic localization, translation initiation in internal ribosome entry site (IRES) and regulation of RNA stability (Fu et al. 2018). This RNA-binding protein is also involved in several biological processes, including cell structure and motility, protein targeting and localization, protein metabolism and modification, muscle contraction, cell cycle and immunity (Fu et al. 2018). Numerous studies have reported that PTBP1 is overexpressed in several different types of cancer, including brain, colorectal, ovarian, gastric and breast cancer (McCutcheon et al. 2004; He et al. 2007; Cheung et al. 2009; He et al. 2014; Takahashi et al. 2015; Sugiyama et al. 2016). Furthermore, high expression of PTBP1 has been demonstrated to be associated with aggressive behaviour of several types of cancer, especially in glioma and ovarian tumours (He et al. 2007; Cheung et al. 2009). In colorectal cancer PTBP1 was shown to facilitate cancer migration and invasion activities by promoting the inclusion of cortactin exon 11 (Wang et al. 2017). In another study PTBP1 was associated with metastasis of colorectal cancer cells by downregulating ATG10, an autophagy-related gene (Jo et al. 2017). PTBP1 was also positively associated with cancer progression properties, such as invasion or proliferation, in colorectal cancer through upregulation of PKM2 (plays a central role in metabolism and growth, promoting cell migration) and CD44 (induces a metastatic phenotype in tumour cells) variants (Takahashi et al. 2015). Furthermore, it was shown, in HCT116 (human epithelial colorectal carcinoma cells with KRAS mutation) that the depletion of PTBP1 and PTBP2 promoted the skipping of exon 3b in RAC1 pre-mRNA (Hollander et al. 2016). In this work the effect of PTBP1 on RAC1b splicing event was further analysed in other colorectal cell lines.

Epithelial splicing regulatory protein 1 (ESRP1) is an epithelial cell-specific RNA-binding protein from the hnRNP family that regulates alternative splicing events associated with epithelial phenotypes (Hayakawa et al. 2016; Jeong et al. 2017). ESRP1 binds preferentially to UGG-rich repeats and plays crucial roles during organogenesis (Hayakawa et al. 2016). This protein regulates the alternative splicing of multiple genes, including CD44, CTNND1, ENAH and FGFR2, all transcripts that undergo changes in splicing during the epithelial-to-mesenchymal transition (EMT), a process by which epithelial cells lose their polarity and acquire motile and invasive phenotypes (Kalluri and Weinberg 2009; Jeong et al. 2017). This protein has dual roles in cancer progression, depending on the context of microenvironments surrounding cancer cells. In some situations, ESRP1 expression is favoured as it supports cell survival; in other situations, downregulation of ESRP1 is favoured as this facilitates cell invasion (Hayakawa et al. 2016). Both of these scenarios were observed in colorectal cancer: on one hand downregulation of ESRP1 promoted EMT and consequently tumour progression (Deloria et al. 2016). On the other hand, ESRP1 overexpression enhanced fibroblast growth factor receptor (FGFR1/2) signalling, Akt activation, and Snail upregulation, thus stimulate growth of cancer epithelial cells and promote colorectal cancer progression. Moreover, ESRP1 promoted the ability of colorectal cells to generate macrometastases in mice livers (Fagoonee et al. 2017). In SAS and HSC4 cells, both tongue squamous cell carcinoma, ESRP1 was found to suppress RAC1b expression (Ishii et

al. 2014). In this work, we studied whether the effect of ESRP1 on RAC1b alternative splicing can also be observed in colorectal cells.

Ran-binding protein 2 (RANBP2), also known as Nup358, is a cytosolic component of the filaments that attach to the cytoplasmic ring of the nuclear pore complexes (NPC). NPC are large protein channels that act as mediators of molecular exchange between the nucleus and the cytoplasm of eukaryotic cells (Raices and D'Angelo 2012; Ibarra and Hetzer 2015). RANBP2 plays major roles in nuclear export and import by providing a docking site for Ran and its cofactors. This protein also mediates SUMOylation of Ran cofactor, RanGAP1, as well as of various cargo proteins (Matunis et al. 1998; Pichler et al. 2002; Forler et al. 2004; Bernad et al. 2004; Reverter and Lima 2005; Hutten et al. 2009). In addition, during mitosis, this nucleoporin is found at kinetochores where it is involved in spindle formation and chromosome segregation (Salina et al. 2003; Joseph and Dasso 2008). Furthermore, RANBP2 binds to the kinesin motors KIF5B and KIF5C, linking the NPC to the cytoskeleton (Cai et al. 2001). RANBP2 has been associated with cancer in different, and contradicting, manners. This protein was shown to act as a tumour suppressor due to its role in preventing chromosome segregation errors (Dawlaty et al. 2008; Navarro and Bachant 2008). However, in another study, RANBP2 was identified as a candidate oncogene overexpressed in the subgroup of human colorectal cancers with microsatellite instability (Gylfe et al. 2013). Corroborating with this information, in another study this nucleoprotein was also reported to be overexpressed in human colorectal cancers with microsatellite instability (Dunican et al.). Consistent with an oncogenic function, RANBP2 was found to protect $BRAF^{V600E}$ mutant colon cancers cells from undergoing mitotic cell death (Vecchione et al. 2016). All this information suggested that somehow RANBP2 is involved in the survival of human colorectal cancers with microsatellite instability. Furthermore, the speckled distribution of phosphorylated pre-mRNA processing factors, like SRSF1 and SRSF3 (known regulators of RAC1b splicing event), was found to be regulated by the nucleocytoplasmic transport system in mammalian cells (Saitoh et al. 2012). Although RANBP2 is not a splicing factor itself, the gathered information prompted the investigation of its role in RAC1b overexpression in colorectal cancer cells.

2. Objectives

RAC1b was found to be overexpressed in a subgroup of colon tumours also characterized by the presence of an oncogenic mutation in BRAF (Matos et al. 2008). Together, these two alterations stimulate signalling pathways that promote the proliferation and survival of malignant cells (Matos et al. 2008). Due to its hyperactive properties and selective overexpression in cancerous tissue, RAC1b is a promising therapeutic target, for this subgroup of tumour patients. Therefore, understanding the basic mechanisms underlining its expression regulation is important to identify potential therapeutic agents. PTBP1 and ESRP1 were described in the literature as possible modulators of the alternative splicing of RAC1b while RANBP2 was shown to be a modulator of the distribution of phosphorylated SR proteins, which are known to regulate RAC1b expression (Saitoh et al. 2012; Ishii et al. 2014; Hollander et al. 2016; Vecchione et al. 2016). With the objective of studying whether these factors can regulate RAC1b expression in colorectal cells, this work was divided into 3 different parts:

(1) Provide or generate expression vectors as tools to study the effect of PTBP1, ESRP1, and RANBP2.

(2) Overexpress PTBP1, ESRP1 and RANBP2 in colorectal cells and determine their effect on alternative splicing of a RAC1 minigene.

(3) Deplete the endogenous expression of PTBP1, ESRP1 and RANBP2 in colorectal cells and determine the effect on alternative splicing of RAC1.

3. Experimental Procedures

3.1. Polymerase Chain Reaction (PCR)

Polymerase chain reaction (PCR) is an invaluable tool for molecular biology research since it provides a rapid mean for DNA identification and analysis through the amplification of a specific DNA region/sequence *in vitro* (Wilson and Walker 2009).

Several components are required to perform a PCR. First, a double-stranded DNA (dsDNA) template that includes the target sequence to be amplified is essential. For each strand, is necessary, a small oligonucleotide, also known as primer, that provides a starting point for DNA synthesis. The primers are chemically synthesized according to the known template DNA sequence and flank the region to be amplified. One primer will have the same sequence as the DNA template - forward primer - and the other will be reverse and complementary - reverse primer. When designing primers there are some constraints that should be considered, such as, the primer length (16-28 base pairs (bp)), the primer length difference (± 3 nt), the GC content (40-60%), the melting temperature (Tm, 50-62°C), the melting temperature difference (± 5°C) and the sequence complementarity between the pair of primers or even in the same primer (possibility of secondary structures, like dimers and hairpins) (Wilson and Walker 2009; Pestana et al. 2010; Chuang et al. 2013). Finally, a DNA polymerase is also needed to synthesize a new DNA strand in the presence of deoxyribonucleotides (dNTPs), a buffer that provides a suitable chemical environment for the polymerase performance and magnesium ions as co-factors to increase the yield of the reaction (Pestana et al. 2010).

PCR can be separated in 3 basic steps: DNA denaturation, primer annealing and polymerase extension. In the first step, high temperatures (94-96°C) are used to break the hydrogen bonds that connect the two DNA strands. This results in the separation of the dsDNA originating two single-stranded DNA (ssDNA) templates. The time necessary for denaturation depends on the size of the DNA fragment to be amplified: the longer the fragment, the longer it takes to be denaturated. In the second step the temperature is lowered until the melting temperature of the primers is reached, allowing them to bind to their complementary sequences on the ssDNA template. In the final step, the polymerase synthesizes the missing strands starting from the annealed primers in a process called elongation. The temperature of elongation depends on the used polymerase (for Taq and Pfu the ideal temperature is 72°C) and the time depends both on DNA polymerase efficiency and on DNA fragment length to be amplified (Taq needs approximately 1 minute to elongate 1000 bp, while Pfu only elongates 500 bp in the same time). Hence, in a short time, exact replicates of the target sequence are produced, and at the end of several cycles, the amount of target sequence is significantly increased enabling further analysis (Pestana et al. 2010).

In general, every PCR reaction mix preformed in this experimental work had a final volume of 25 μL. Each primer was used at 0,2 μM and Go Taq G2 Flexi at 0,02 U/μL. The volume of DNA added depended on the type and purpose of the PCR and buffer B (10 mM Tris-HCl pH 9 1,5 M; 50 mM $MgCl_2$; 1,5 mM KCl; 0,1% Bacta Gelatin (DifcoLab)); 0,2 mM of each dNTP) was added to make up the final volume. For all PCRs, a mock reaction was made with water instead of DNA to ensure the amplification resulted from the DNA template and not from genomic DNA or possible contaminants in the mix. The thermocycler C1000 Touch™ Thermal Cycler (Bio-Rad) was used to run the amplification programs.

3.1.1. PTBP1 and ESRP1 amplification

The mRNA sequences of splicing factors PTBP1 (NM_002819.4) and ESRP1 (NM_017697.3) were obtained from the National Center for Biotechnology Information database (NCBI, https://www.ncbi.nlm.nih.gov/) and used as templates to design primers according to the constraints mentioned in *Chapter 3.1*. For both splicing factors, the restriction sequence of BamHI was added to the 5' end of the forward primer and the restriction sequence of NotI to the 5' end of the reverse primer. Thus, after the PCR, splicing factors cDNA sequence was between BamHI and NotI restriction sequences, to later subclone them into the expression vector pcDNA3_myc (procedure described in *Chapter 3.2.1*). The two pairs of primers used are described in **Table 3.1.** ESRP1 sequence was amplified using 2 μL of a cDNA pool of HT29 cells synthesized from 2 μg of total RNA, while PTBP1 was amplified from 100 ng of a Flag-PTBP1 plasmid available at the host lab. The PCR conditions were similar to the ones described in *Chapter 3.1*., however, they were performed with a DNA polymerase mixture of DF-Pfu (BIORON) and Go Taq G2 Flexi (Promega) at the concentration of 0,02 U/μL and 0,01 U/μL, respectively. The optimized PCR programs used were:

ESRP1 PCR program:

94°C – 10 min
94°C – 45 s
64°C – 30 s **x 35**
72°C – 4 min and 30 s
72°C – 10 min

PTBP1 PCR program:

94°C – 10 min
94°C – 1 min
58°C – 1 min **x 40**
72°C – 4 min
72°C – 10 min

Table 3.1 - Primers sequence used for the amplification of ESRP1 and PTBP1 cDNA sequences.

Splicing factor	Primers	
	Name: Sequence 5' → 3'	Tm (°C)
ESRP1	BamHI_ESRP1 F: 5' GGATCCACGGCCTCTCCGGATTACTT 3'	62
	NotI_ESRP1 R: 5' GCGGCCGCTAAATACAAACCCATTCTTTGGG 3'	62
PTBP1	BamHI_PTBP1 F: 5' GGATCCGACGGCATTGTCCCAGATAT 3'	60
	NotI_PTBP1 R: 5' GCGGCCGCTAGATGGTGGACTTGGAGA 3'	60

3.1.2. Agarose gel electrophoresis

After the PCR reaction is complete, PCR products can be identified using agarose gel electrophoresis through the application of an electric current (Wilson and Walker 2009; Lee et al. 2012). This technique is used to separate nucleic acid fragments by size allowing their visualization and purification (Wilson and Walker 2009).

For the agarose gel with pre-casted wells to load the samples, agarose was dissolved in Tris-borate-EDTA buffer (TBE; 89 mM Tris; 89 mM Boric Acid and 2 mM EDTA). The percentage of agarose chosen depended on the size of the DNA fragments to be separated. Higher concentrations of agarose results in smaller pore sizes in the gel, making it, therefore, better to separate smaller DNA fragments. To visualize the DNA molecules in the gel, ethidium bromide (EtBr), a DNA intercalating

dye that emits fluorescence when exposed to UV light, was added to the agarose mix to a concentration of 0.5 μg/ml. Next, the gel mix was poured into a mould and an appropriate comb was introduced to create the wells. After agarose gelification the comb was removed, and the gel was placed in a horizontal electrophoresis chamber filled with TBE. To facilitate the introduction of samples into the gel wells, a loading buffer was added (5X Green GoTaq® Flexi Buffer, Promega) to increase sample density. In addition, this loading buffer also contained two different dyes (a blue dye and a yellow dye) that allow the visual tracking of DNA migration during electrophoresis. After DNA sample application, a constant voltage was applied (70 - 95 V) to the gel for suitable time periods that depended on the DNA migration rate and the electrophoresis purpose. Since DNA has a uniform mass/charge ratio, their molecules are separated by size in a pattern such that the distance travelled in the gel is inversely proportional to the log of its molecular weight (Lee et al. 2012). The size of the PCR product can be determined by simultaneously running a DNA ladder, that contains DNA fragments of known size (Wilson and Walker 2009). At the end of electrophoresis, the gel was placed on a UV light transilluminator to visualize the DNA molecules (Lee et al. 2012). A picture of the gel was then taken resorting to a digital image acquisition system, Fire Reader™.

3.2. Cloning

Cloning is a molecular biology technique used to produce identical copies of recombinant DNA. These molecular clones can be used, for example, to analyse gene sequences, to express the resulting protein in cells, or even to manipulate and mutate the DNA sequence *in vitro* (Wilson and Walker 2009).

Molecular cloning starts with the production of recombinant DNA molecules by introducing a DNA fragment of interest into a circular piece of DNA (plasmid) through the use of restriction enzymes and a DNA ligase (Wilson and Walker 2009). Restriction enzymes recognize one or a few target nucleotide sequences where they cut at or near it. Some restriction enzymes originate products with short single-stranded overhang ends (sticky ends) (Roberts 2005). Thus, if two different DNA molecules, such as the DNA of interest and a plasmid, are digested by the same set of enzymes they will have matching overhangs (Roberts 2005). This means that they can attach together by complementary base pairing, however, to form an unbroken DNA molecule they need to be joined by DNA ligase (Roberts 2005). After producing the recombinant DNA molecules, multiple identical copies of it are obtained by extracting and purifying it from a host organism, such as bacteria, previously transformed with the recombinant DNA.

3.2.1. PTBP1 and ESRP1 cloning

Linear PCR products can not be digested properly by restriction enzymes because they lack nucleotides flanking one side of the restriction sequences. As so, before cloning them into an expression vector, PCR products from PTBP1 and ESRP1 amplifications were first TOPO®-cloned. This cloning technique consists in the direct insertion of Taq polymerase-amplified PCR products that have a terminal adenine overhang into a linear plasmid vector with a complementary T overhang. This product can later be used to reclone the PCR products into an expression plasmid using restriction enzymes.

First, PCR products from PTBP1 and ESRP1 amplifications were inserted in the pCR™2.1-TOPO® vector from the TOPO® TA Cloning® kit (Invitrogen), according to the manufacturer's instructions. Then, the resulting ligations, TOPO_PTBP1 and TOPO_ESRP1, were used to transform competent *Escherichia coli* (*E. coli*, TOP10 strain), included in the TOPO® TA Cloning® kit. The transformation started by adding, on ice, 2,5 μL of each ligation (10% of the bacteria volume) to 25 μL of TOP10 bacteria. This mix was then incubated for one hour on ice, allowing the spreading of

plasmids around the bacterial walls. Next, to promote the entrance of plasmids into the bacteria, a thermal shock was performed by submitting the mix to a temperature of 42°C for 45 seconds, followed by 2 minutes on ice. Subsequent steps were performed under semi-sterile conditions next to a gas flame, creating an upward flow of hot air. 125 µL of S.O.C. medium (5x the bacteria volume) was added to the mix, and then incubated at 37°C under agitation, for one hour. Meanwhile, 40 µL of X-Gal (stock at 40 µg/mL) was spread on a pre-warmed Luria-Bertani (LB) agar plate containing kanamycin (50 µg/mL). In the end, 150 µL of the bacteria suspension was spread on the plate, and later incubated overnight at 37°C, for bacterial growth. The presence of kanamycin in the agar plate ensures that only transformed bacteria can grow, as the pCR™2.1-TOPO® vector contains the kanamycin resistance gene, unlike non-transformed TOP10 bacteria. The X-Gal is used in this context to, later, identify the colonies in which the PCR product was introduced in the pCR™2.1-TOPO® vector. TOP10 bacteria are mutants for the alpha fragment of lac Z (Φ80lacZΔM15), lacking the first residues of β-galactosidase, the α-peptide. However, these bacteria contain the omega fragment of β-galactosidase, which corresponds to the carboxy-terminus of the protein. By themselves both the α and the Ω fragments are non-functional, yet, when the two peptides are expressed together they interact and originate a functional β-galactosidase. The pCR™2.1-TOPO® vector carries the lacZα sequence within the cloning site for PCR products. So, when the PCR product is incorporated into the vector, the lacZα gene is disrupted and a non-functional β-galactosidase is produced (W/out the α-peptide). The presence of an active β-galactosidase can be detected by X-Gal, a colourless analogue of lactose, that can be hydrolysed by β-galactosidase to form a blue precipitate. Therefore, blue colonies indicate that the incorporated vector has no insert in it, while white colonies, where X-Gal was not hydrolysed, indicate the possible presence of the expected insert.

After the overnight incubation, one blue control colony and several white colonies were selected and resuspended in water, to perform a PCR screening using T7 and M13 as primers (**Table 3.3**). Considering that the used primes hybridize upstream and downstream of the vector cloning site, this PCR will originate different sized products according to the presence or absence of insert. From the colonies that tested positive in the screening, bacterial cultures were grown by adding 5 µL of the colony suspension to 3 mL of liquid LB medium supplemented with 50 µg/mL of kanamycin in a 15 mL tube (5x the medium volume of air), and incubated overnight, under agitation, at 37°C. The resulting grown bacterial cultures were used to extract and purify the cloned vectors using the GeneJet Plasmid Miniprep kit (Thermo Fisher Scientific), according to the manufacturer's instructions. To confirm PTBP1 and ESRP1 were correctly inserted in the pCR™2.1-TOPO® vector and that no alterations were present, a sample from each miniprep was prepared for automated DNA sequencing (*Chapter 3.2.4*).

After the confirmation of TOPO_PTBP1 and TOPO_ESRP1 sequences, subcloning into pcDNA3_myc, an expression vector engineered by the host lab, was performed. First, both the pcDNA3_myc vector and the TOPO plasmids, with either PTBP1 or ESRP1 insert, were digested by the restriction enzymes, BamHI (Promega) and NotI (Speedy CciNI (NotI), Nzytech) in NEBuffer 3 (New England Biolabs®) with 0,1 mg/mL of BSA (New England Biolabs®) (**Table 3.2**).

Table 3.2 - Composition of the reaction mixes for the digestion of the TOPO vectors, TOPO_PTBP1 and TOPO_ESRP1, and the pcDNA3_myc vector.

Reaction Mix components	TOPO_PTBP1	TOPO_ESRP1	pcDNA3_myc
NEBuffer 3 10X	5 µL	5 µL	5 µL
Purified BSA 10x	5 µL	5 µL	5 µL
Vector	3000 ng	3000 ng	1000 ng
BamHI (10U/µL)	1 µL	1 µL	0,5 µL
NotI (5U/µL)	1 µL	1 µL	0,5 µL

Add H_2O up to 50 µL

For the PTBP1 subcloning, TOPO_PTBP1 and pcDNA3_myc digestion was performed for 4 h, at 37°C, whereas for ESRP1, TOPO_ESRP1 and pcDNA3_myc digestion was performed overnight, at 37°C. 5 µL of each digestion was confirmed by gel electrophoresis in a 1% agarose gel. The remaining 45 µL of linearized pcDNA3_myc vector was dephosphorylated with 1U of bacterial alkaline phosphatase (BAP, Rapid DNA Dephos & Ligation Kit, Roche) for an hour at 65°C. The dephosphorylated linearized pcDNA3_myc vector was then purified with the NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nagel), according to the manufacturer's instructions. The remaining 45 µL of digested TOPO_PTBP1 and TOPO_ESRP1 vectors were run in a 1% low melting agarose gel and the bands corresponding to either PTBP1 or ESRP1 were cut from the gel and purified with the GeneJET Gel Extraction Kit (Thermo Fisher Scientific), according to the manufacturer's instructions. The ligation reaction between the purified linearized pcDNA3_myc vector and the purified inserts of PTBP1 or ESRP1 was performed in a proportion of 1:7 (v/v), respectively, with the Rapid DNA Dephos & Ligation Kit, according to the manufacturer's instructions.

The resulting ligation reactions were used to transform competent DH5α *E. coli*, prepared in the host lab. First, 5 µL of ligation reaction was added, on ice, to 50 µL of DH5α bacteria. This mix was incubated for an hour on ice, and then submitted to heat shock for 45 seconds at 42°C, followed by 2 minutes on ice. Next, under semi-sterile conditions, 250 µL of LB medium was added to the bacteria mix and incubated at 37°C under agitation for one hour, 150 µL of the bacteria suspension was spread on a pre-warmed LB agar plate containing ampicillin (100 µg/mL), followed by incubation overnight at 37°C. The presence of ampicillin in the agar plate ensures only transformed bacteria can grow, as the pcDNA3_myc vector contains the ampicillin resistance gene, unlike non-transformed DH5α bacteria. On the next day, some colonies were selected and resuspended in water, to perform a PCR screening using PCMV5-1F and PFGH-R1 as primers (**Table 3.3**). This PCR will originate different sized products according to the presence or absence of insert because the pair of primers used hybridize upstream and downstream of the vector cloning site. From the colonies that tested positive in the screening, 5 µL of each colony suspension was added to 3 mL of liquid LB medium supplemented with ampicillin (100 µg/mL) and incubated overnight, under agitation, at 37°C. The resulting grown bacterial culture was then used to extract and purify the vector by using the GeneJet Plasmid Miniprep kit (Thermo Fisher Scientific), according to the manufacturer's instructions. To confirm that both PTBP1 and ESRP1 were correctly inserted into pcDNA3_myc vector, a sample from each miniprep was prepared for automated DNA sequencing (*Chapter 3.2.4*).

All plasmid preparations were quantified in the NanoDrop™ spectrophotometer (Thermo Fisher Scientific) by measuring the absorbance at 260 nm, and the DNA purity was assessed by the ratio of absorbance at 260 nm and 280 nm, being generally accepted a ratio of 1,8 to have “pure” DNA.

Table 3.3 - Primers and PCR programs used for the screening of the integration of PTBP1 and ESRP1 into the pCR™2.1-

Screening for	Primer name: Sequence (5' → 3')	Tm (°C)	PCR Program
TOPO_PTBP1 and TOPO_ESRP1	T7: TAATACGACTCACTATAGGG	56	94°C – 8 min **94°C – 30 s** **54°C – 30 s** **72°C – 90 s** 72°C – 5 min **35 Cycles**
	M13: GGAAACAGCTATGACCATG	52	
pcDNA3_myc_PTBP1 and pcDNA3_myc_ESRP1	PCMV5–1F: GGGACTTTCCAAAATGTCGTA	56	94°C – 8 min **94°C – 30 s** **56°C – 30 s** **72°C – 90 s** 72°C – 5 min **35 Cycles**
	PFGH–R1: TTTATTAGGAAAGGACAGTGGG	58	

TOPO® vector and the pcDNA3_myc expression vector.

3.2.4. Sequencing

DNA sequencing is a laboratory methodology by which the sequence of nucleotides in a portion of DNA is determined (Rye et al. 2017). In this experimental work, sequencing was performed using the Sanger sequencing method, also known as the chain termination method. In Sanger sequencing, the DNA to be sequenced serves as a template for DNA synthesis (Rye et al. 2017). The sequencing reaction mix includes the DNA template, one primer, a DNA polymerase, dNTPs and, in much smaller amounts, the four dye-labelled, chain-terminating dideoxy nucleotides (ddNTPs) (Rye et al. 2017). The mixture is first heated to denature the template DNA and then cooled for primer annealing (Rye et al. 2017). Once the primer has bound, the temperature is raised again, allowing DNA polymerase to synthesize new DNA (Rye et al. 2017). When a ddNTP is incorporated into a chain of nucleotides DNA synthesis terminates. This happens because ddNTPs lack the hydroxyl group on the 3' carbon of the sugar ring, which is required to form the phosphodiester link with the next nucleotide in the chain. Since the ddNTPs are randomly incorporated, synthesis terminates at many different positions, resulting in products with different sizes (Rye et al. 2017). After the sequencing reaction, the resulting products are separated by capillary gel electrophoresis, and, as each fragment reaches the end of the gel matrix tube, a laser excites the dye attached to the ddNTP, allowing its detection (Rye et al. 2017). This way the sequence of the original piece of DNA can be built up one nucleotide at a time, since each nucleotide corresponds to a different color dye, and smaller fragments move quicker through the pores of the gel, while bigger ones move slower (Rye et al. 2017).

For each sequencing reaction, the mix had a final volume of 10 μL, including 350 ng of DNA template, 20 μmol of the chosen primer (**Table 3.4**) and 1 μL of 1x BigDye™ terminator ready mix in sequencing buffer (BigDye™ Terminator v3.1 Cycle Sequencing Kit, Applied Biosystems™). The mix was then placed in a thermocycler and submitted to the program:

96°C – 10 min

96°C – 10 s
52°C – 5 s
60°C – 4 min
x 25

4°C – Hold

The resulting products from the sequencing reaction were sent to the in-house sequencing facility (*Unidade de Tecnologia e Inovação, Departamento de Genética Humana*), where they were purified

and loaded in the automatic DNA sequencer 3130xl Genetic Analyzer (Applied Biosystems™), from which was obtained a chromatogram with the DNA template sequence.

The different sequences obtained for each primer were analysed with the Basic Local Alignment Search Tool (BLAST ®, https://blast.ncbi.nlm.nih.gov/Blast.cgi) by aligning them with the expected sequence.

Table 3.4 - Primers used to sequence the vectors cloned with ESRP1, PTBP1 and RANBP2.

	Primer Name	Primer Sequence (5'→ 3')	Tm (°C)
pCR™2.1-TOPO®	T7	TAATACGACTCACTATAGGG	56
	M13	GGAAACAGCTATGACCATG	52
pcDNA3_myc	PCMV5–1F	GGGACTTTCCAAAATGTCGTA	56
	PFGH–R1	TTTATTAGGAAAGGACAGTGGG	58
ESRP1	BamHI_ESRP1 F	GGATCCACGGCCTCTCCGGATTACTT	62
	ESRP1_Seq251 F	AGCCCTCCGACAGTTTAACC	58
	ESRP1_Seq593 F	TCACAGGTTTTCAGATCCAGAGA	66
	ESRP1_Seq951 F	GCCCAGTTTCTCTCCAAGGA	62
	ESRP1_Seq1318 F	CCCCTACAAATGTTAGAGACTGT	66
	NotI_ESRP1 R	GCGGCCGCTAAATACAAACCCATTCTTTGGG	62
PTBP1	BamHI_PTBP1 F	GGATCCGACGGCATTGTCCCAGATAT	60
	PTBP1_Seq372 F	CCAGCCCATCTACATCCAGT	62
	PTBP1_Seq784 F	ACCAGCCTCAACGTCAAGTA	60
	PTBP1_Seq1388 R	GGGATGTTGGAGAGGTGCA	60
	NotI_PTBP1 R	GCGGCCGCTAGATGGTGGACTTGGAGA	60
RANBP2*	pEGFP-F2	ACATGGTCCTGCTGGAGTTC	62
	pkTol2Chy_RANBP2 R	GCTCAAGGGGCTTCATGATG	62

* RANBP2 expression vector (pkTol2Chy_RANBP2_1-3224) was kindly provided by Dr. Masakazu Hamada from the Mayo Clinic in Rochester, Minnesota, EUA.

3.3. Cell Culture

The culture of human cells consists in growing and maintaining human cells *in vitro* under artificial controlled conditions. This *in vitro* technique is very useful and widely applied in biological research since it allows the study of human metabolism and physiology that is not readily possible *in vivo* (Phelan 1998; Mitry and Hughes 2012).

Cell lines are commercially available and under favourable conditions, they proliferate until they reach confluence (the desired confluence is different from cell line to cell line). At this moment cells must be subcultured (passaged), which means that it is necessary to transfer and dilute them into a second vessel with new growth medium providing more space and nutrients for continued growth (Mitry and Hughes 2012). This cycle of reaching confluence and being passaged, is how cells are maintained in culture. For adherent cells, like the ones used in this experimental work, before subculturing, cells need to be dissociated from each other and from the vessel. This disassociation is typically done using trypsin, a proteolytic enzyme that breaks down proteins that help cells adhere to each other and to the vessels (Mitry and Hughes 2012).

Several conditions are required in order to successfully grow and maintain human cells *in vitro*. Firstly, every cell culture technique must be carried out under strict aseptic conditions, as common contaminants (e.g. bacteria, moulds, and yeast) grow faster than mammalian cells (Mitry and Hughes 2012). Therefore, cell lines must be manipulated inside a laminar flow hood previously sterilized by UV light, and all materials, like pipettes, flasks, dishes and tubes, that come into direct contact with the culture must be sterile. It is also advisable to swab hands, the work surface, and all

non-sterile materials that get into the laminar flow hood with 70% alcohol, frequently (Mitry and Hughes 2012). These aseptic measures will help preventing possible culture contamination. Another way to prevent contamination is by adding to the medium antibiotics like penicillin and streptomycin, however, in this experimental work, this addition could have an interfering effect on the cellular physiology since the study revolves around cell signalling, and so no antibiotics were used in cell culture. Secondly, a favourable artificial environment is also required for cell to grow and maintain *in vitro*. This environment is accomplished by using the right medium, carbon dioxide (CO_2) percentage, humidity and temperature. Each cell line has different physicochemical necessities (pH, osmotic pressure, temperature) and so the culture conditions used will depend on those necessities (Mitry and Hughes 2012). Mediums typically contain essential nutrients, like amino acids, carbohydrates, vitamins and minerals, and they are also usually supplemented with foetal bovine serum (FBS), which provides vital macromolecules and growth factors. The physical conditions are achieved by using a CO_2 - humidified incubator, being the most common used parameters 37°C with 5% of CO_2, however these change depending on the cell line. Another aspect to take into account is the vessel where cells are cultured, most cells are anchorage-dependent and must be cultured in a vessel with a solid or semi-solid substrate (adherent or monolayer culture), while others can be grown floating in the culture medium (suspension culture) (Mitry and Hughes 2012).

Healthy normal cells usually divide for a limited number of times before losing their ability to proliferate, these cell lines are known as finite. However, some cell lines go through transformation, a process that results in cell immortality. This process can occur spontaneously (e.g. Cancer cells) or it can be chemically or virally induced. When a finite cell line undergoes transformation, and acquires the ability to divide indefinitely, it becomes a continuous cell line (Mitry and Hughes 2012). Even though finite cell lines retain most of the characteristics of the original tissue they are difficult to maintain in culture and can have batch-to-batch variation (Mitry and Hughes 2012). Immortalized cells, might not be 100% like the original tissue, since they are in fact transformed, but they are very similar to it. This makes them great models of the original tissue, and therefore extremely important in biological research. It is important to mention that these immortalized cell lines cannot be maintained indefinitely, because with the accumulation of passages cells tend to gather more and more genetic mutations, which might lead to a point where they no longer represent the behaviour of the original tissue. As so, it is essential to maintain the number of passages no higher than 30, even though this number varies from cell line to cell line.

The main used cell line in this experimental work was NCM460, a normal colonic epithelial cell line (Moyer et al. 1996). Other cell lines like HT29, Caco2, HEK 293 and HeLa were used to test, confirm and compare results. Important information about each cell line used in this work, like their origin, desired passage confluence, and appropriate medium, can be found at **Table 3.5**.

Table 3.5 - Information on every cell line used on this experimental work, including their origin, type and culture conditions.

Cell Line	Description	Species of origin	Cell Type	Thawing	Desired passage confluence	Growth Medium
NCM460	Normal colon	*Homo sapiens*	Epithelial	Sensitive	≈ 80%	RPMI[*2] + 10% (v/v) of FBS[*3]
HT29	Colorectal adenocarcinoma	*Homo sapiens*	Epithelial	Non-sensitive	≈ 100%	RPMI + 10% (v/v) of FBS
Caco2	Colorectal adenocarcinoma	*Homo sapiens*	Epithelial	Sensitive	≈ 100%	RPMI + 10% (v/v) of FBS
HEK 293[*1]	Embryonic kidney 293	*Homo sapiens*	Epithelial	Non-sensitive	≈ 100%	DMEM[*4] + 10% (v/v) of FBS
HeLa	Cervix adenocarcinoma	*Homo sapiens*	Epithelial	Non-sensitive	≈ 100%	DMEM + 10% (v/v) of FBS

[*1]HEK 293 disassociate from the flask very easily, as so the trypsinization step is shorter and faster.
[*2]RPMI 1640 Medium, GlutaMAX™ Supplement, Gibco™, Thermo Fisher Scientific
[*3]FBS, Gibco™, Thermo Fisher Scientific
[*4]DMEM + GlutaMAX™, Gibco™, Thermo Fisher Scientific

Mycoplasm contamination is very common in cell culture laboratories, and due to its size (~100 nm) and behavior, it is undetectable by the naked eye or even by optical microscopy (Miller et al. 2003). Therefore, cell lines were regularly tested by a research staff member for mycoplasm contamination. Throughout the experimental work, this tests result was always negative.

In order to keep cell cultures at passages numbers under 30, and for other reasons, like making sure the culture is not lost due to unexpected equipment failure or biological contaminations, cell line stocks were made and stored in liquid nitrogen (-196°C). These stocks were usually made at the lowest passage number as possible (e.g. the first passage after thawing the cells) and assuring that cells are viable and at high concentration to be frozen so that when they are thawed the chance of cell survival is higher. Accordingly, after the cells were passaged, the remaining suspension was centrifuged at 300 x g for 5 minutes and the resulting pellet resuspended in FBS, a growth supplement for cells, with 10% of dimethyl sulfoxide (DMSO), a cryoprotective agent. Without the use of DMSO, it would be lethal for cells to be frozen. This cryoprotective substance reduces the freezing point and allows a slower cooling rate reducing the risk of ice crystal formation, which can damage cells and cause cell death. This suspension was then divided into cryovials (1mL/cryovial) and stored in a cryofreezing container (with isopropanol) at -80°C for at least 24 h. This procedure ensures the cells are slowly frozen by reducing the temperature at approximately 1°C per minute. Although necessary to cryopreservation, DMSO is also very toxic to cells, so when thawing the cells, this compound should be removed promptly. To thaw the cells, the cryovial has to be thawed rapidly at 37°C, and its content resuspended and diluted in the appropriate growth medium. If the cells are sensitive, the diluted suspension was centrifuged at 300 x g for 5 minutes and the supernatant discarded. The remaining pellet was then resuspended in the medium and introduced in a culture flask. If the cell line is less sensitive the centrifugation step is not necessary and the cryovial resuspended content is directly introduced in the flask, however on the day after thawing, the medium has to be changed in order to remove the DMSO (Mitry and Hughes 2012).

3.4. Cell Transfection

Cell transfection consists in the artificial introduction of nucleic acids (DNA or RNA) into eukaryotic cells (Kim and Eberwine 2010). This technique's main purposes are to either produce recombinant proteins or to specifically enhance or inhibit gene expression in transfected cells. As

such, transfection is a powerful analytical tool for the study of the function and regulation of genes or gene products in the context of the cell (Kim and Eberwine 2010).

Cell transfection can be stable if the cell integrates the genetic material into the genome, replicating when the host genome replicates. It can also be transient if the introduced nucleic acid only remains in the cell for a limited period of time and is not integrated into the genome (Kim and Eberwine 2010).

The introduction of foreign nucleic acids into eukaryotic cells can be accomplished by biological, physical or chemical methods. In biological methods, the introduction of nucleic acids into the cell is virus-mediated, also known as transduction (Kim and Eberwine 2010). In physical methods, the nucleic acids are delivered to the cell through the use of physical tools, such as microinjection. Chemical transfection methods are the most commonly used, consisting in the use of a cationic polymer that forms a complex with the negatively charged nucleic acid, and then these positively charged complexes are attracted to the negatively charged cell membrane (Kim and Eberwine 2010). The exact mechanism by which these complexes enter the cell is still unknown, however, it is believed that endocytosis and phagocytosis might be involved in the process. Inside the cell, nucleic acids must translocate into the nucleus, where transcription occurs, however, once again, the mechanism by which this happens is still unknown (Kim and Eberwine 2010). Chemical methods efficiency depends on factors such as nucleic acid/chemical ratio, solution pH and cell membrane conditions and availability, differing from cell line to cell line. In comparison to the other methods, chemical transfection tends to have low efficiency, however, is an easy to use technique, with no package size limit and with relatively low cytotoxicity. (Kim and Eberwine 2010)

Overall, transfection efficiency depends, mainly, on the chosen transfection method, on the cell line, on the health and viability of the cells, on the degree of confluence, on the quality and quantity of the nucleic acid used, and on the presence or absence of serum in the medium. The right confluence to use differs from cell line to cell line, if the number of cells is too low cell cultures grow poorly, whereas too many cells result in contact inhibition, making cells resistant to the uptake of DNA and other macromolecules. Confluence should never be above 80% because cells will not actively divide, making it harder for them to take up DNA and other macromolecules. It is important to use high-quality plasmid DNA that is free of contaminants, such as proteins, RNA, and chemicals. The optimal amount of DNA to use in transfection varies depending on the type of DNA, transfection reagent/method, target cell line, and number of cells. Serum is an important supplement for cell growth and its presence in culture mediums enhances transfection with DNA. However, in cationic lipid-mediated transfection, some proteins from the serum might interfere with DNA/lipid complex formation. To avoid this problem, complex formation is performed in a reduced/free serum medium and later added to the rich serum medium where the cells are cultured.

Enhancing the expression of a specific gene can be achieved by transfecting cells with a DNA plasmid containing the cDNA corresponding to the mRNA sequence encoded in that gene. The opposite process, of inhibiting the expression of a specific gene, can be accomplished through the transfection of chemically synthesized short/small interfering RNAs (siRNAs). The siRNA nucleotide sequence is complementary to a small portion of the target mRNA sequence, and when these small RNAs attach to the RNA-induced silencing complex (RISC), they direct the complex to bind to the target mRNA (Dana et al. 2017). This results in the enzymatic cleavage of the target mRNA and consequent inhibition of its translation, which effectively silences the expression of the gene (Dana et al. 2017).

All transfections and co-transfections (simultaneous transfection with two different nucleic acid molecules) performed in this experimental work were transient and the method used for acid nucleic introduction was chemical, cationic lipid-mediated. Cells were grown in different culture dishes depending on the cell line and on the intended experiment. Optimal cell confluence for

transfections was 60-80% for DNA plasmid and 30-40% for siRNA, in all cell lines with the exception of HT29 (40-60% for DNA plasmid and 20-30% for siRNA). The transfection reagent used was Lipofectamine™ 2000 (Invitrogen, Thermo Fisher Scientific), except for HEK 293 that was Metafectene® (Biontex). The transfection procedure was performed according to the manufacturer's instructions and the conditions for each experiment are resumed in **Tables 3.6 and 3.7**. In the end, transfected culture dishes were incubated at 37°C with 5% of CO_2, for as long as intended (the incubation time depends on the experiment goals).

In every experiment, pcDNA3_EGFP and siGFP were used as controls in plasmid DNA and siRNA transfection, respectively. Controls provide constant variables enabling the correct interpretation of the tested independent variable effect. Without controls, the experiment results would be inconclusive and unreliable: first because there wouldn't be a "standard" against which the results could be compared, and second because there wouldn't be a way to prove that the observed results are a consequence of the independent variable effect and not of experimental conditions. Controls also provide information on whether the reported assay is working properly.

Table 3.6 - Co-transfection conditions of pcDNA3_EGFP, pcDNA3_myc_PTBP1, pcDNA3_myc_ESRP1 or pkTol2Chy_RANBP2_1-3224 with RAC1 minigene (MG) in NCM460, HeLa and HT29 cell lines. Transfection conditions of pcDNA3_EGFP or pkTol2Chy_RANBP2_1-3224 in NCM460, Caco2, HT29 and HEK 293 cell lines. All transfections and co-transfections were performed in 6-well plates.

Cell Line	Cell quantity (in millions)	DNA	DNA quantity	Transfection reagent quantity	Complexes Formation
NCM460	1	GFP RAC1 MG	2 μg 0,5 μg	10 μL of LF2K[*1]	Prepare the mixes: 125 μL of OptiMEM[*2] + LF2K 125 μL of OptiMEM + DNA mix ↓ Incubate for 5 min at room temperature ↓ Combine the two mixes and incubate for 20 min at room temperature (LF2K/DNA complexes formation) ↓ Add ≈ 250 μL of OptiMEM with LF2K/DNA complexes to each well
		PTBP1 RAC1MG	2 μg 0,5 μg	10 μL of LF2K	
		ESRP1 RAC1MG	2 μg 0,5 μg	10 μL of LF2K	
		RANBP2 RAC1MG	2,5 μg 0,5 μg	10 μL of LF2K	
HeLa	1	GFP RAC1MG	1,5 μg 0,5 μg	4 μL of LF2K	
		PTBP1 RAC1MG	1,5 μg 0,5 μg	4 μL of LF2K	
		ESRP1 RAC1MG	1,5 μg 0,5 μg	4 μL of LF2K	
		RANBP2 RAC1MG	2 μg 0,5 μg	4 μL of LF2K	
HT29	0,5	GFP RAC1MG	2 μg 0,5 μg	10 μL of LF2K	
		PTBP1 RAC1MG	2 μg 0,5 μg	10 μL of LF2K	
		ESRP1 RAC1MG	2 μg 0,5 μg	10 μL of LF2K	
		RANBP2 RAC1MG	2,5 μg 0,5 μg	10 μL of LF2K	
Caco2	2	GFP	2,5 μg	10 μL of LF2K	
		RANBP2	2,5 μg	10 μL of LF2K	
NCM460	1	GFP	2,5 μg	10 μL of LF2K	
		RANBP2	2,5 μg	10 μL of LF2K	
HT29	0,5	GFP	2,5 μg	10 μL of LF2K	
		RANBP2	2,5 μg	10 μL of LF2K	
HEK 293	0,5	GFP	2 μg	9 μL of Met[*3]	Instead of 125 μL of OptiMEM it was used 150 μL of DMEM, and so ≈ 300 μL of DMEM with Met/ DNA complexes was added to each well
		RANBP2	2 μg	9 μL of Met	

[*1] LF2K - Lipofectamine™ 2000
[*2] Opti-MEM™ Gibco™, Thermo Fisher Scientific
[*3] Met - Metafectene®

Table 3.7 - Transfection conditions of siGFP (Eurofins Genomics), siESRP1 (Santa Cruz Biotechnology) and siRANBP2 (Santa Cruz Biotechnology) for NCM460, HeLa and HT29 cell lines, all performed in 24-well plates.

Cell Line	Cell quantity (in millions)	DNA	DNA quantity	Transfection reagent quantity	Complexes Formation
NCM460	0,15	siGFP	50 pmol	2 μL of LF2K	Prepare the mixes: 50 μL of OptiMEM + LF2K 50 μL of OptiMEM + DNA ↓ Incubate for 5 min at room temperature ↓ Combine the two mixes and incubate for 20 min at room temperature (LF2K/DNA complexes formation) ↓ Add ≈ 100 μL of OptiMEM with LF2K/DNA complexes to each well
		siESRP1	50 pmol		
		siRANBP2	60 pmol		
HT29	0,075	siGFP	50 pmol	2 μL of LF2K	
		siESRP1	50 pmol		
		siRANBP2	60 pmol		
HeLa	0,15	siGFP	40 pmol	0,4 μL of LF2K	
		siESRP1	40 pmol		
		siRANBP2	50 pmol		

3.5. Cell lysis

Cell lysis refers to the process of disintegration or rupture of the cell membrane resulting in the release of biological material that exists within the cell (lysate). The lysate can then be used to purify or further study cells contents, such as proteins, nucleic acids, and organelles (Wilson and Walker 2009). Cell lysis methods can be divided in two main categories: mechanical and non-mechanical methods. Mechanical techniques resort to machines that use force to generate a lysate, unlike non-mechanical methods that lyse cells by disrupting the lipid membrane and/or cell wall (Wilson and Walker 2009).

Our experimental goals involved the study of gene and protein expression. Therefore, cell lysis was performed to obtain simultaneously both protein and RNA extracts (soft lysis). This was possible with the use of NP-40 lysis buffer (50 mM Tris-HCl, pH 7,5; 2 mM $MgCl_2$; 100 mM NaCl; 10% Glycerol; 1% NP-40), that contains NP-40, a mild nonionic detergent. This surfactant solubilizes both the plasma membrane and the internal membranous organelles but not nuclear membranes, making it possible to obtain both nuclear and cytoplasmic fractions. Proteases inhibitors were added to the NP-40 lysis buffer to avoid degradation of the proteins of interest. As RANBP2 is a nuclear pore complex protein, the nuclear and the cytoplasmic fractions were not separated by centrifugation, and consequently, protein lysates contained nucleic acids. Ergo, the nuclease benzonase (Thermo Fisher Scientific) had to be added to the sample buffer 5x (SB 5x: 250 mM Tris-HCl, pH6,8; 25% Glycerol; 10% SDS, 325 mM Dithiothreitol (DTT); 1,25 mg Bromophenol blue) that was added to the portion of lysate used to analyse the cellular proteins.

Initially, when performing cell lysis, the culture medium was discarded, and cells were washed twice with free-serum medium to remove every serum protein present in the dish. After thoroughly washing the cells, the culture dish was placed in ice and NP-40 lysis buffer with proteases inhibitors (1 μL of each cocktail per 100 μL of NP-40 lysis buffer → Cocktail 1: 1,5 μM aprotinin, 23 μM leupeptin, 10 μM E64, 1 mM EGTA; Cocktail 2: 15 μM pepstatin A, 1 mM PMSF, 1 mM 1,10-

phenanthroline) was added to the cells. After five minutes cells were scraped, and the lysate was divided into two portions, one to extract RNA with the RNA extraction kit NucleoSpin® RNA (Macherey Nagel) and other, to which was added SB 5x with benzonase (0,5 U/μL) to obtain protein extract. Next, the protein extract was heated at 95°C, for 10 minutes, in a thermocycler (Biometra®, Tpersonal), in order to denature the proteins. The sodium dodecyl sulphate (SDS), an anionic detergent, and the DTT, a reducing agent, present in the sample buffer also contributed to the protein denaturation process by chemically disrupting the proteins tertiary structure. Cell lysis conditions are described in **Table 3.8**.

To test the transfection efficiency in different cell lines, only protein extracts were necessary to access the results, thus total lysis was performed with 100 μL of SB2x per p35 dish. After discarding the culture medium and thoroughly washing the cells, 100 μL of SB2x with benzonase was added to each dish. After 5 minutes cells were scraped, and the lysate heated at 95°C, for 10 minutes, in a thermocycler (Biometra®, Tpersonal).

Table 3.8 - Cell lysis conditions for every transfection preformed in this experimental work.

Transfected Material	Culture dishes type	Time of lysis after transfection	Lysis buffer* (μL)	Lysate division	SB5x + Benzonase (μL)
DNA plasmid	6-well plates	24h	80	40 μL for RNA 40 μL for Protein	10
siRNA's	24-well plates	24h	30	15 μL for RNA 15 μL for Protein	4
		48h	35	17,5 μL for RNA 17,5 μL for Protein	4,5
		72h	40	20 μL for RNA 20 μL for Protein	5
		96h	45	22,5 μL for RNA 22,5 μL for Protein	5,5

* NP-40 lysis buffer with proteases inhibitors

3.6. SDS-PAGE and Western blot

The combination of sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) and Western blot (WB) techniques is widely used to study the proteins of interest in a sample.

SDS-PAGE is an electrophoretic technique used to separate proteins in a gel according to their molecular weight, in the presence of an SDS buffer. SDS is an anionic detergent that in addition to disturbing the non-covalent forces in proteins, also masks proteins intrinsic charges by binding uniformly to their linear/denatured form (Saraswathy and Ramalingam 2011). Consequently, every linear protein, independently of its amino acid charge, will have a uniform negative charge, meaning all proteins will migrate towards the positive anode. However, depending on their molecular weight proteins will have different migration rates, bigger proteins will move slower through the gel matrix comparing to smaller proteins. Thus, to guarantee that the separation of denatured proteins in an electrical field is based only on their molecular weight, SDS is used to make the sample buffer, the electrophoresis buffer and the polyacrylamide gel.

The polyacrylamide gel was divided in two layers: the stacking and the resolving layer. When an electric field is applied, the stacking layer concentrates the linear protein molecules, ensuring that every protein is found at the same position when the separation is initiated. The resolving layer

separate the proteins based on their molecular weight. In this experimental work, staking layers had 4% of polyacrylamide and the resolving gel had 8% or 12% of polyacrylamide, depending on the sizes of the proteins that were intended to be separated. Bigger percentages of polyacrylamide originate smaller gel matrix pores, which separates better smaller proteins. The mix for each gel layer was prepared with bi-distilled water, Tris-HCl buffer (Tris-HCl 0,5 M pH 6,8 for the stacking gel and Tris-HCl 1,5 M pH 8,8 for the resolving gel), 40% acrylamide/bis solution 37,5:1 (BIO-RAD) and SDS 10% (Sigma-Aldrich), as described in **Table 3.9**. The polymerization process was assured by adding ammonium persulfate (APS, Sigma-Aldrich) and tetramethylethylenediamine (TEMED, VWR). TEMED accelerates the rate of formation of free radicals from APS and these in turn catalyse polymerization (Yang and Mahmood 2012). The polymerized polyacrylamide layered gel between the two glasses was placed in the electrode assembly and inserted into the electrophoresis chamber (Mini-PROTEAN®, BIO-RAD) according to the manufacturer's instructions. Denatured protein samples, obtained as described in *Chapter 3.5*, were run for approximately 1 hour at a fixed amperage of 20 mA/gel in SDS buffer (8,7 mM SDS, 63 mM Tris, 480 mM Glycine). The migration rate of the proteins can be monitored by simultaneously running a prestained protein marker (PageRuler™ Plus Prestained Protein Ladder, Thermo Fisher Scientific Fisher ^{TM}) that contains stained proteins of known molecular weight. This marker also allows to assess transfer efficiency and to estimate the approximate size of the separated proteins.

Table 3.9 - Recipes of the polyacrylamide gels

	Resolving gel		**Stacking gel**
Polyacrylamide percentage	8%	12%	4%
ddH_2O (mL)	2,70	2,2	1,5
Tris-HCl buffer (mL)	1,25	1,25	0,25
40% acrylamide/bis 37,5:1 (mL)	1,00	1,5	0,2
SDS 10% (mL)	0,05	0,05	0,02
APS (mL)	0,05	0,05	0,04
TMED (mL)	0,0025	0,0025	0,002

Western blotting corresponds to the electrophoretic transfer of proteins from polyacrylamide gels to a polyvinylidene fluoride membrane (PVDF) or nitrocellullose membrane, followed by immunodetection of proteins using antibodies with chemiluminescent detection (Yang and Mahmood 2012). Initially, a primary antibody is used to specifically recognize the target antigen (protein of interest), and then a secondary enzyme-conjugated antibody will recognize the heavy chains of the specie of the primary antibody. Since several secondary antibodies can bind to the primary antibody there will be a signal amplification that will allow the detection of the proteins of interest. The conjugated enzyme will catalyse a reaction that emits light, which can then be captured with X-ray film.

At the end of SDS-PAGE the polyacrylamide gel was removed from the gel cassette and the stacking layer cut off. The proteins in the resolving gel were transferred to a PVDF membrane (BIO-RAD), previously hydrated, using a wet/tank blotting system Mini Trans-Blot® Cell (BIO-RAD) according to the manufacturer's instructions (**Figure 3.1**).

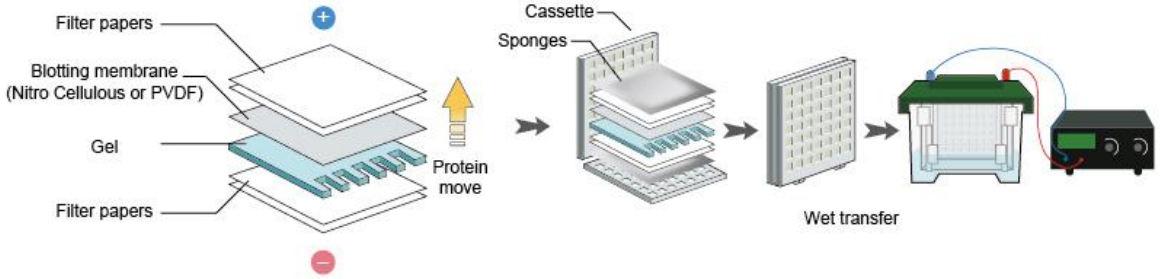

Figure 3.1 – Western blot transference. Schematic representation of a western blot transference. On the left there is the components of the transference sandwich that are placed between the sponges of the cassette, that is then closed and introduced in the tank for protein wet transfer to the PVDF membrane. Image adapted from https://www.creative-diagnostics.com/Electrophoresis-Protein-Transfer.htm.

The transfer occurred for 1 h at the fix voltage of 100 V in transfer buffer (48 mM Tris, 38,6 mM Glycine, 1,41 mM SDS, 20% (v/v) Methanol). In the end, the proteins on the membrane were stained and fixed with a solution of Coomassie (0,25% (m/v) Coomassie Brilliant Blue R250 dye, 45% (v/v) Methanol, 10% (v/v) acetic acid) and the excess was washed, under agitation, with a destaining solution (45% (v/v) Methanol, 10% (v/v) acetic acid). Next, the membrane was washed 3 times with a TBST solution (20 M Tris, 6,7 M NaCl, 3% (v/v) HCl 37%, 0,05% (v/v) Triton X-100 (Sigma-Aldrich) pH 7,6) followed by 1-hour incubation, under agitation, with a blocking solution (5% (m/v) non-fat milk powder (Molico, Nestle) in TBST). The milk proteins will bind to the free spaces of the membrane preventing antibodies from binding to the membrane nonspecifically. After blocking, the membrane was cut according to the proteins that were intended to be detected. Then, each part of the membrane was incubated over-night, under agitation, with the appropriate dilution of primary antibody in blocking solution (**Table 3.10**). Next, the membrane parts were washed 3 times with TBST, under agitation, for 10, 5 and 5 minutes, respectively, to remove the excess of primary antibody that did not bind to the membrane. Afterwards, they were incubated with the secondary antibody, conjugated with horseradish peroxidase (HRP) (**Table 3.10**), under agitation for an hour, and washed again 3 times with TBST, under agitation, for 10 minutes each time to remove the excess of secondary antibody. Finally, visualization of the proteins of interest was done in a dark chamber based on a chemiluminescent method where membranes are exposed to a mix of two solutions (Solution 1- 0,1 M Tris-HCI pH 8,8, 37,5 pM Luminol (Roche), 0,4 mM p-coumaric acid (Roche) and Solution 2 - 0,1 M Tris-HCI pH 8,8, 49 mM Hydrogen peroxide (H_20_2, Sigma-Aldrich)) for 1 minute. During this minute, HRP, in the presence of peroxide, will oxidize luminol to an excited product called 3-aminophthalate that emits light (Thorpe and Kricka 1986). This light was then captured by making multiple exposures to X-ray films in an appropriate cassette. The exposed films were later processed in an automatic processor (Medical X-ray Processor, KODAK).

Table 3.10 - Antibodies used in Western blot and respective dilutions.

Primary antibody	Brand	Dilution	Secondary antibody	Brand	Dilution
Rabbit Anti-GFP antibody	Abcam	1:2000	Goat Anti-Rabbit IgG Conjugated to HRP	BIO-RAD	1:3000
Rabbit Anti-RAC1b	MilliPore ™	1:4000			
Mouse Anti-c-Myc, clone 9E10	Sigma-Aldrich	1:2000	Goat Anti-Mouse IgG Conjugated to HRP		
Mouse Anti-α-Tubulin	Sigma-Aldrich	1:12000			
Mouse Anti-RANBP2	Santa Cruz Biotechnology	1:500			

3.7. Semi-quantitative RT-PCR

Reverse transcription polymerase chain reaction (RT-PCR) is a rapid and sensitive technique used in the detection and/or quantification of mRNA, being, therefore, a powerful tool in the study of gene expression (Farrell 2017). RT-PCR is a two-step process, in which purified RNA is, first, reverse transcribed into single-stranded cDNA molecules. And then, the cDNA, that is much less prone to degradation than RNA, is amplified by a standard PCR procedure (Farrell 2017). Semi-quantitative PCRs allow the comparison of an RNA transcript expression level under different experimental conditions because, in parallel to the PCR amplification of the cDNA of interest, a transcript of a housekeeping gene is also amplified. Since the housekeeping gene has a ubiquitous expression, its quantification can be used to normalize the mRNA of interest expression levels, making it, possible to compare the different experimental conditions. The relative quantification of the different transcripts is only possible by making serial dilutions of control cDNAs to establish the equation of the linear amplification phase for each PCR reaction (Ferre 1992).

In this experimental work, RNA extraction and purification was done using the RNA extraction kit NucleoSpin® RNA (Macherey Nagel), according to the manufacturer's instructions. RNA extracts were quantified in the NanoDrop™ spectrophotometer by measuring the absorbance at 260 nm, and the RNA purity was assessed by the ratio of absorbance at 260 nm and 280 nm, being generally accepted a ratio of 2.0 to have "pure" RNA. cDNA was synthesized from 1 μg of RNA using the qScript™ XLT cDNA SuperMix (Quanta Bio) according to the manufacturer's instructions. PCR conditions are similar to the ones described in *Chapter 3.1*. The primers and PCR programs used are described in **Table 3.11**. PCR products were separated on 2 - 2,5% agarose gels, and images were captured using the digital image acquisition system, Fire Reader.

Table 3.11 - Primers and PCR programs used for the semi-quantitative RT-PCRs of RAC1b from the minigene (MG), endogenous RAC1b, ESRP1 and RANBP2.

Target	Primer Name: Sequence 5' → 3'	Tm (°C)	PCR Program	
MG-RAC1b	KtagBamS: CATGATCGACTACGACGTTCCTGATTATGCGG	80	94°C – 5 min **94°C – 30 s** **60°C – 30 s** **72°C – 30 s** 72°C – 5 min	**28 Cycles**
	RACRealEx3b-R: ATATCCTTACCGTACGTTTCTCCAA	70		
MG-RAC1	KtagBamS: CATGATCGACTACGACGTTCCTGATTATGCGG	80	94°C – 5 min **94°C – 30 s** **54°C – 30 s** **72°C – 30 s** 72°C – 5 min	**29 Cycles**
	RACRealJunc3-4R: ACAAGCAAATTGAGAACACATCTGTT	70		
Endogenous RAC1b	RACBamEx3-F: GGATCCTTTGACAATTATTCTGCCAATG	62	94°C – 5 min **94°C – 30 s** **60°C – 30 s** **72°C – 30 s** 72°C – 5 min	**28 Cycles**
	RACRealEx3b-R: ATATCCTTACCGTACGTTTCTCCAA	70		
Endogenous RAC1	RACBamEx3-F: GGATCCTTTGACAATTATTCTGCCAATG	62	94°C – 5 min **94°C – 30 s** **54°C – 30 s** **72°C – 30 s** 72°C – 5 min	**29 Cycles**
	RACRealJunc3-4R: ACAAGCAAATTGAGAACACATCTGTT	70		
RANBP2	RANBP2 F: CCATGAGGCAGAGAGGAACA	62	94°C – 5 min **94°C – 30 s** **60°C – 30 s** **72°C – 45 s** 72°C – 5 min	**26 Cycles**
	RANBP2 R: GGTCACAGGCCATCATTTCC	62		
ESRP1	ESRP1_Seq1318F: CCCCTACAAATGTTAGAGACTGT	66	94°C – 5 min **94°C – 30 s** **64°C – 30 s** **72°C – 60 s** 72°C – 5 min	**35 Cycles**
	NotI_ESRP1 R: GCGGCCGCTAAATACAAACCCATTCTTTGGG	62		
Pol2	QPol2-F: CGCAATGAGCAGAACGGCGC	66	94°C – 5 min **94°C – 30 s** **64°C – 30 s** **72°C – 30 s** 72°C – 5 min	**28 Cycles**
	QPol2-R: TCTGCATGGCACGGGGCAAG	66		

3.8. Immunofluorescence assay

Immunofluorescence assay (IF) is a laboratory technique, mainly used on biological samples, that relies on the use of antibodies to label a specific target antigen with a fluorescent dye (fluorophore) (Odell and Cook 2013). The antigen can then be visualized under a fluorescent microscope, making it possible to determine the target antigen localization and expression levels in the sample. The fluorescent readout is done by using a light source to excite the fluorophore, that absorbs the radiation, becoming transiently excited. When the fluorophore returns to its ground state, it produces a light emission that is detected with a specialized reader. Since some of the absorbed energy by the fluorophore is lost due to molecular vibrations, the emitted light has a higher wavelength than the one used for excitation. The antigen can be either directly or indirectly detected, depending on whether the fluorophore is conjugated to the primary or the secondary antibody (**Figure 3.2**). In direct methods, a single primary antibody conjugated with a fluorophore is used to detect the target antigen.

Whereas, in indirect methods, a primary antibody is initially used to specifically recognize the target antigen, and then a secondary antibody conjugated with a fluorophore is added to bind to the primary antibody. This last method results in signal amplification since several secondary antibodies can bind to the primary antibody (**Figure 3.2**) (Odell and Cook 2013).

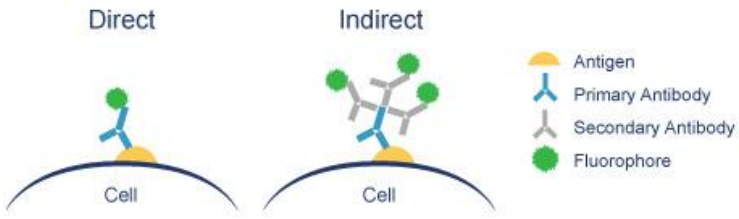

Figure 3.2 – Antigen detection methods. In the left side there is the schematic representation of a direct antigen detection. In the right side there is the schematic representation of an indirect antigen detection. Image from https://www.abcam.com/secondary-antibodies/direct-vs-indirect-immunofluorescence.

Subcellular localization of PTBP1 and ESRP1 were determined by indirect IF (using the Myc tag as target for the primary antibody), while RANBP2 localization was detected without the use of antibodies since it was tagged with GFP, a protein that emits fluorescence naturally. To perform the IF, NCM460 and HeLa cell lines were cultured on glass coverslips (sterilized by flaming) inside 6-well plates, where the tagged plasmids of interest were transfected. These coverslips were washed 3 times with PBS (14 μM NaCl; 0,2 μM KCl; 0,2 μM $Na_2HPO_4.2H_2O$; 0,2 μM KH_2PO_4 and 0,2% (m/v) Sodium Azide) and fixed for 30 minutes with 4% formaldehyde (PFA 4% - 10% (v/v) PBS 10x; 11% (v/v) formaldehyde 35%). This fixation step will preserve the cells in their current state, avoiding their degradation and thus allowing a future analysis. Then, cells were washed 3 times with PBS and permeabilized for 30 minutes with PBS + 0,5% Triton X-100. Next, cells were washed 3 times with PBST (PBS + 0,001% Triton X-100) and incubated 1 hour with mouse anti-Myc antibody (1:100, in PBST, Sigma-Aldrich), in a humidified chamber (except for the GFP_RANBP2 that remained in PBST). After 3 washes of 5 minutes each, with PBST under agitation, the coverslips were incubated for 30 minutes with Alexa Fluor 488 goat anti-mouse (1:250, in PBST, Invitrogen) and phalloidin-TRITC antibody (1:250, in PBST, Sigma), in a humidified chamber. Coverslips with cells transfected with GFP or GFP_RANBP2 were also incubated for the same time but only with phalloidin-TRITC antibody (1:250, in PBST). Phalloidin is a high-affinity filamentous actin probe with an emission light in the red zone of the spectrum. All the coverslips were then washed 3 times for 5 minutes each, under agitation with PBST, followed by DAPI (4',6-diamidino-2-phenylindole, 0,125μg/ml, in PBST) staining during few seconds. DAPI binds to A-T rich regions of the DNA double helix and emits light in the blue zone of the spectrum. After 3 more washes with PBST, cells were post-fixed with 4% PFA for 15 minutes. Finally, coverslips were washed 3 more times with PBST and assembled over EverBrite Mounting Medium (Biotium) in a microscope slide, being sealed with transparent nail polish. Cells were later visualized, and images recorded by confocal microscopy (TCS-SPE, Leica) where they were excited with a 488 nm laser to see the transfected tagged plasmids, whereas phalloidin and DAPI were exited with a 532 nm and 405 nm laser, respectively.

3.9. Data Treatment

Densitometric analysis, of the digitalized X-ray films and the digitally recorded images from the RT-PCRs bands, was performed resorting to the ImageJ software (Download: https://imagej.nih.gov/ij/download.html).

Protein abundance and percentage of mRNA depletion were determined by calculating the ratio between the protein or cDNA of interest and the housekeeping protein or transcript. Then, in order to compare and analyse the ratios, normalization to the experiment control ratio had to be done. For RT-PCR semi-quantification, before calculating the ratios, DNA concentration in each band was determined by using the equation of linear amplification phase from the PCR. For protein and mRNA depletion analysis, tubulin and polymerase II were used as the housekeeping protein and transcript, respectively. Since for RAC1b mRNA expression the goal was to determine the differences in RAC1 alternative splicing in the different experimental conditions, RAC1b/RAC1 ratio was calculated instead of the RAC1b/housekeeping gene ratio.

Results significance was assayed by using Student's t-test, a statistical hypothesis test that compares the means of two groups (Kim 2015). In this experimental work, a two-sample assuming equal variances t-test was used, in which the null hypothesis (H_0) was that the means were equal ($\mu 1 - \mu 2 = 0$) and the alternative hypothesis (H_a) was that the means were different ($\mu 1 - \mu 2 \neq 0$). The probability that the results occurred by chance is given by the p-value. In this experimental work, the p-value from which the null hypothesis was rejected was 0,05 (5% of probability that the results occurred by chance).

4. Results

4.1. Production of expression vectors for PTBP1 and ESRP1

As described in *Chapter 3.1.1*, primers were designed to amplify the complete cDNA sequence of ESRP1 and PTBP1 (**Table 3.1**). With the goal of cloning both into an expressing vector to later express the encoded proteins in cells, a high-fidelity DNA polymerase was used to avoid/reduce the introduction of errors in the DNA sequence of the PCR products. Therefore, Pfu, a thermostable DNA polymerase with 3' → 5' exonuclease-dependent proofreading activity, was first chosen to perform the PCR amplification (Cline 1996). However, no amplification was obtained using only Pfu, probably due to this polymerase's lower efficiency. Since Taq, a non-proofreading DNA polymerase, has higher efficiency than Pfu, a mixture of Pfu and Taq was used in the following PCRs (Cline 1996). This way a compromise between the fidelity and efficiency of the PCR reaction should be achieved. As shown in **Figure 4.1**, using the polymerase mix, bands with the expected sizes for ESRP1 and PTBP1 (approximately 2000 kb and 1600 kb, respectively) were obtained. While PTBP1 cDNA sequence was amplified from a Flag-PTBP1 plasmid available at the host lab, ESRP1 cDNA sequence had to be amplified from a cDNA pool prepared from cell lysates. ESRP1 transcript expression was first assayed in different cell lines (HT29, NCM460, and SW480 (data not shown)), and the resulting amplification was more efficient for the cDNA pool of HT29 cells. Thus, this cDNA pool of HT29 cells was chosen to proceed with ESRP1 amplification.

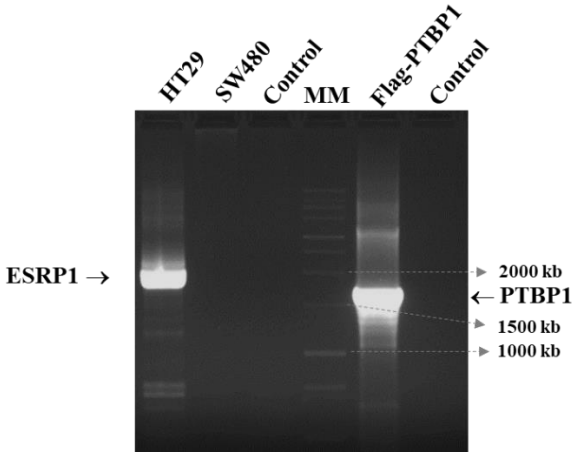

Figure 4.1 – PTBP1 and ESRP1 cDNA amplification. Digitally recorded image of the 2% agarose gel with the PCR products from ESRP1 and PTBP1 cDNA amplifications using a mix of Pfu and Taq. HT29 and SW480 cDNA pools were used as templates for ESRP1 cDNA amplification while a Flag-PTBP1 plasmid at 50 ng/μL was used for PTBP1 amplification. The control lanes correspond to the PCR mix without DNA. ESRP1 and PTBP1 bands have approximately 2000 kb and 1600 kb, respectively. MM: Molecular Marker (Promega)

After amplification, the ESRP1 and PTBP1 cDNA sequences were first cloned into the pCR™2.1-TOPO® vector using the TA-cloning strategy and not directly into pcDNA3_myc expression vectors because the digestion of linear PCR products by restriction enzymes is inefficient, due to the lack of nucleotides flanking one side of the restriction sequences. Following each TOPO cloning reaction, bacteria were transformed, previously described in *Chapter 3.2.1*, and 17 white colonies and 1 blue colony were randomly chosen to perform a PCR screening using the appropriate pair of primers (**Figure 4.2 A**). From the colonies that tested positive in each screening, three were grown in mini-cultures for plasmid purification (white arrows in **Figure 4.2 A**), followed by the reaction for automated DNA sequencing to determine if the sequences were correctly inserted into the TOPO vector. After confirming that PTBP1 and ESRP1 sequences were correctly inserted in the TOPO vector with no alterations compared to the database sequence, the most concentrated mini-prep

of each was chosen for subcloning into pcDNA3_myc expression vector, as previously described in *Chapter 3.2.1*. To confirm the subcloning, a PCR screening of 18 randomly chosen colonies was performed with the appropriate pair of primers to determine which colonies had the cDNA sequence of interest inserted into pcDNA3_myc vector (**Figure 4.2 B**). Three positively tested colonies of each subcloning were chosen to purify the plasmid (white arrows in **Figure 4.2 B**), and then sequence it to determine if the sequences were correctly inserted into pcDNA3_myc vector. All six had PTBP1 and ESRP1 sequences correctly inserted in the pcDNA3_myc vector with no alterations, therefore the most concentrated mini-prep of each cloned vector was chosen to proceed with the experiments.

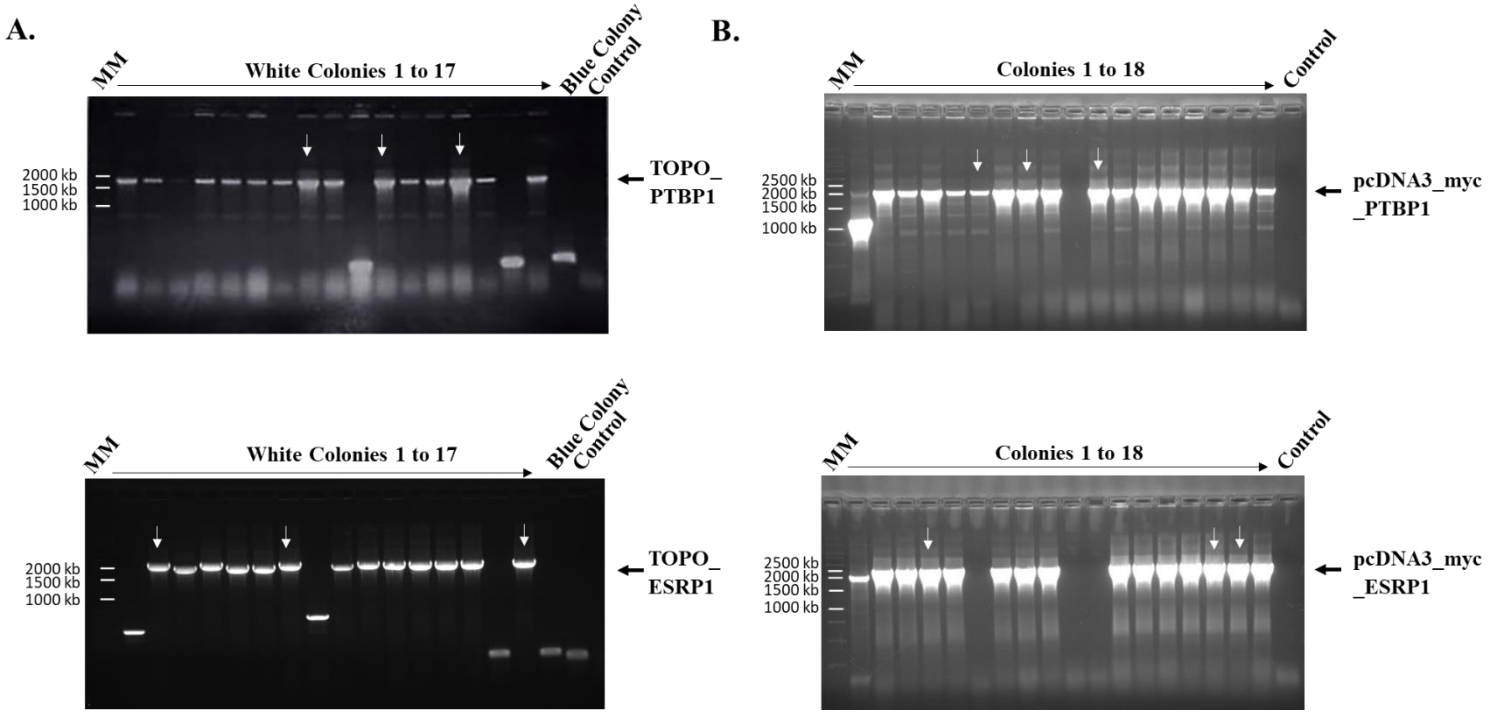

Figure 4.2 – PCR screenings. A. Digitally recorded image of the agarose gels with the PCR products from the screening performed for TOPO_PTBP1 (above) and TOPO_ESRP1 (below). 17 white colonies were chosen for the screening and a blue colony was chosen as a negative control. The control lanes correspond to the PCR mix without DNA. Colonies with TOPO vectors with the inserted ESRP1 and PTBP1 have a PCR product size around 2200 kb and 1800 kb, respectively. **B.** Digitally recorded image of the agarose gels with the PCR products from the screening performed for pcDNA3_myc_PTBP1 (above) and pcDNA3_myc_ESRP1 (below). 18 colonies were chosen for the screening. The control lanes correspond to the PCR mix without DNA. Colonies with pcDNA3_myc vectors with the inserted ESRP1 and PTBP1 have a PCR product size around 2200 kb and 1800 kb, respectively. MM: Molecular Marker (Promega). White arrows mark the chosen positive clones.

4.2. The effect of PTBP1, ESRP1 and RANBP2 overexpression in the alternative splicing of RAC1

For the overexpression experiments, a RAC1 minigene (**Figure 4.3**), previously shown to recapitulate the endogenous splicing decisions in colorectal cells (Gonçalves et al. 2009), was used to assay the effect of the three proteins, PTBP1, ESRP1 and RANBP2 on the inclusion of the alternative exon 3b into the RAC1 minigene-derived transcript. Since transient transfection efficiency of expression vectors in cell lines never reaches 100%, the effects of the transfected proteins on endogenous RAC1 transcripts could be masked by the non-transfected cells. Thus, by co-transfecting RAC1 minigene along with the expression vectors of interest, the effect of the corresponding proteins on exon 3b inclusion can be accessed more accurately.

The regulation of RAC1 alternative splicing was intended to be studied in colorectal tumour cells. So, HT29, a colorectal adenocarcinoma cell line with high expression of RAC1b, was chosen to

perform the co-transfection experiments. However, the transfection of pkTol2Chy_RANBP2_1-3224 vector into HT29 cells did not result in detectable RANBP2 expression. To understand if RANBP2 protein was being expressed correctly, HEK 293 cells, which are known to have high transfection efficiency, were transfected with pkTol2Chy_RANBP2_1-3224: This vector has an in frame GFP tag upstream of the RANBP2 sequence so that RANBP2 expression can be visualized in the cells by fluorescence microscopy and by preforming a western blot with α-GFP antibody. Both techniques confirmed (data not shown) that GFP_RANBP2 protein was being correctly expressed from the pkTol2Chy_RANBP21-3224 transfected vector. Apparently HT29 cells transfection efficiency was too low to detect GFP_RANBP2 transfected protein. Therefore, alternative colorectal cell lines, NCM460 and Caco2 cells (both expressing RAC1b), were transfected with pkTol2Chy_RANBP2_1-3224. After western blot analysis it was found that only NCM460 cells had a significant expression of GFP_RANBP2 transfected protein, being therefore, the chosen cell line to proceed with the study.

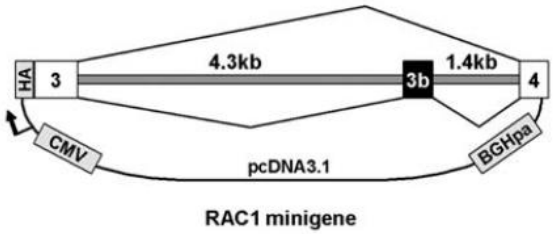

Figure 4.3 – RAC1 minigene. Schematic representation of the RAC1 minigene, containing the genomic region between exon 3 and exon 4 of the human RAC1 gene. Image from (Gonçalves et al. 2009).

In order to understand if PTBP1, ESRP1 and RANBP2 regulate RAC1b alternative splicing, pcDNA3_myc_PTBP1, pcDNA3_myc_ESRP1, pkTol2Chy_RANBP2_1-3224 were each co-transfected with the RAC1 minigene into NCM460 cells. The pcDNA3_GFP was used in this experiment as a control since GFP is a fluorescent protein that does not affect RAC1b expression. The expression of the transfected tagged proteins was assessed by western blot analysis using the specific tag antibody, as shown in **Figure 4.4**. Every transfected vector was successfully expressed in NCM460 cells.

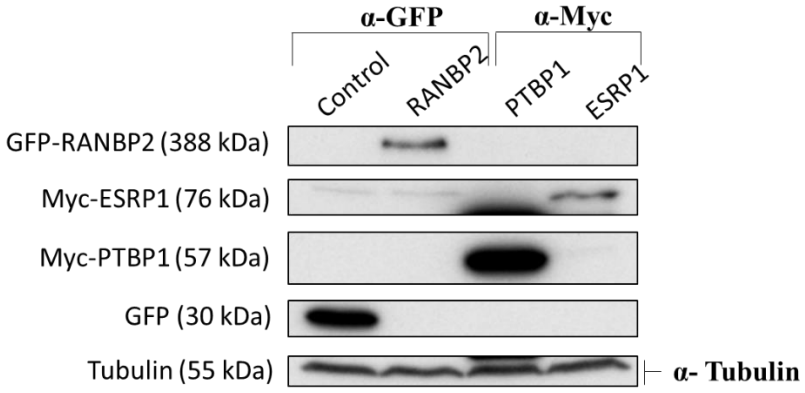

Figure 4.4 – Western blot analysis of the transfected tagged-proteins expression. pcDNA3_GFP, pcDNA3_myc_PTBP1, pcDNA3_myc_ESRP1 and pkTol2Chy_RANBP21-3224 were each co-transfected with the RAC1 minigene into NCM460 cells. Transfected GFP and RANBP2 were detected by using an anti-GFP antibody, while transfected PTBP1 and ESRP1 were detected by using an anti-Myc antibody. Tubulin was detected as loading control.

To analyse the roles of the transfected proteins in the inclusion of exon 3b, a semi-quantitative RT-PCR, with specific primers for RAC1 and RAC1b transcripts derived from the RAC1 minigene, was performed (**Figure 4.5 A**, upper panel). PTBP1 overexpression does not affect significantly exon

3b inclusion while ESRP1 and RANBP2 decrease that inclusion when overexpressed (**Figure 4.5 A**, lower panel**)**. The WB results for endogenous RAC1b (**Figure 4.5 B**) correlate with the results obtained in the RT-PCR for the RAC1 minigene approach. Thus, in conclusion, ESRP1 and RANBP2 are promoting the skipping of the alternative exon 3b while PTBP1 does not seem to have a significant effect on RAC1b splicing in these cells.

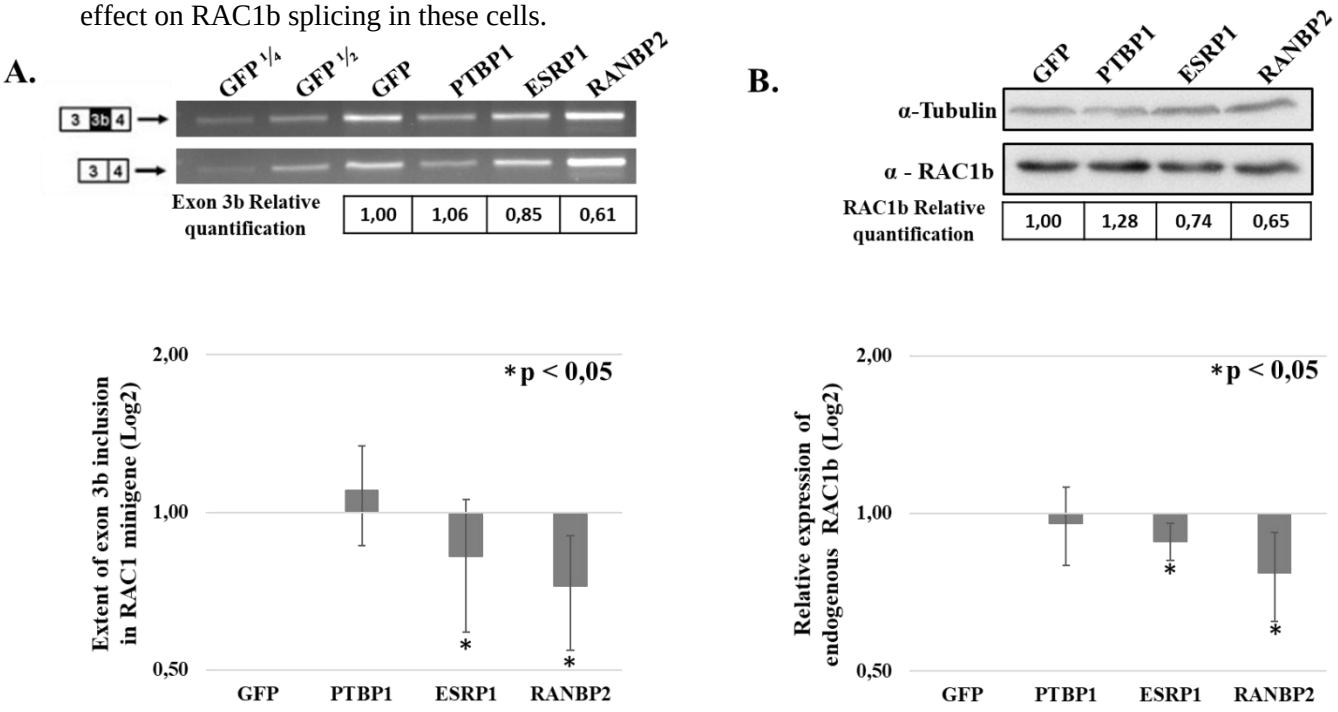

Figure 4.5 – Analysis of RAC1 exon 3b inclusion after PTBP1, ESRP1 and RANBP2 overexpression in NCM460 cells. A. Effect on RAC1 minigene alternative splicing in NCM460 cells. Representative semi-quantitative RT-PCR showing the densitometric analysis of the minigene-derived transcripts (upper panel). Exon 3b relative inclusion analysis was done as described in *Chapter* 3.9. In the bottom panel the results from semi-quantitative RT-PCR experiments were quantified and are graphically displayed on a logarithmic scale for proportional visualization of positive and negative changes. **B.** Effect on endogenous RAC1 alternative splicing in NCM460 cells. RAC1b and tubulin were detected by western blot with specific antibodies and densitometrically quantified by ImageJ. RAC1b relative expression was done as described in *Chapter* 3.9. In the bottom panel the results from western blot experiments were quantified and are graphically displayed on a logarithmic scale for proportional visualization of positive and negative changes. Asterisks (*) indicate the results considered significant by a two-way t-test assuming equal variances, in which n=4 and p<0,5.

To understand if these results were dependent on the cell line (NCM460), pcDNA3_myc_PTBP1, pcDNA3_myc_ESRP1, pkTol2Chy_RANBP21-3224 and pcDNA3_GFP as control, were each co-transfected with RAC1 minigene into HeLa cells, which also express RAC1b but do not have a colorectal origin. By western blot analysis, the expression of the transfected tagged proteins in HeLa cells was found to be similar to the observed for NCM460 cells (data not shown). A semi-quantitative RT-PCR specific for the RAC1 minigene-derived transcripts revealed that ESRP1 and PTBP1 overexpression decrease exon 3b inclusion while RANBP2 does not have a significant effect on inclusion (**Figure 4.6 A**). For endogenous RAC1b, the obtained results did not reveal significant alterations (**Figure 4.6 B**), probably due to non-transfected cells interference.

The increase of exon 3b skipping observed upon RANBP2 overexpression was specific of NCM460 cells, while ESRP1 overexpression enhanced exon 3b skipping in both cell lines. In fact, this skipping is apparently more pronounced in HeLa cells. PTBP1 overexpression also promoted exon 3b skipping but only in HeLa cells. This way it seems that ESRP1 is the only cell-type independent regulator of RAC1b splicing.

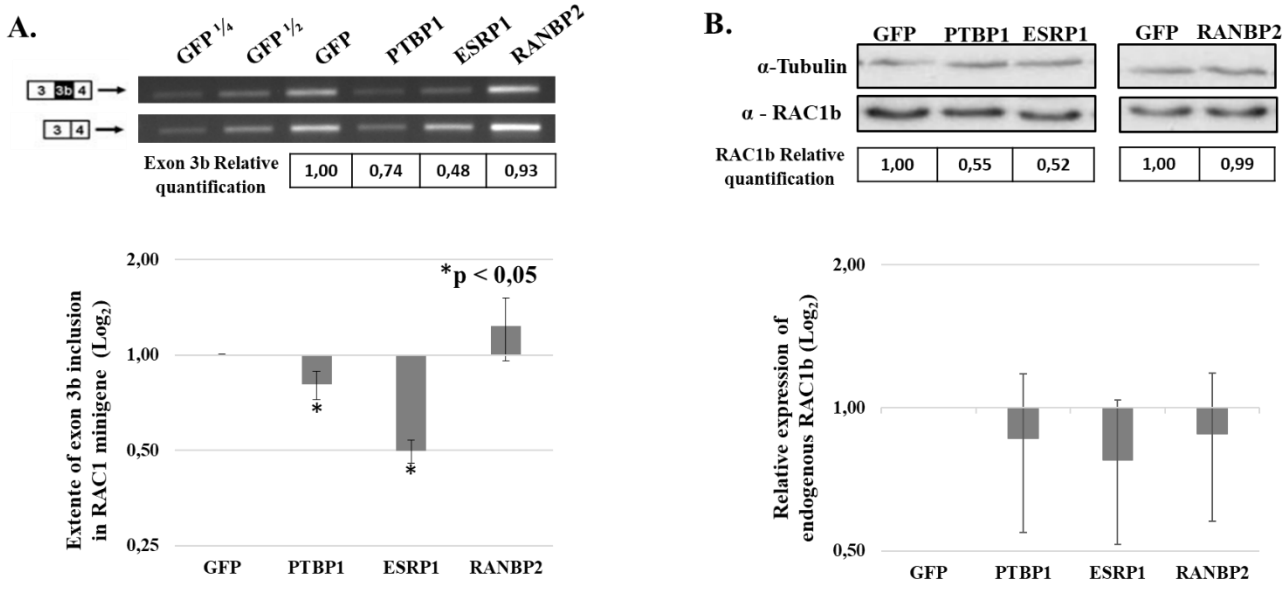

Figure 4.6 – Analysis of RAC1 exon 3b inclusion after PTBP1, ESRP1 and RANBP2 overexpression in HeLa cells. A. Effect on RAC1 minigene alternative splicing in HeLa cells. RAC1 minigene semi-quantitative RT-PCR showing the densitometric analysis of the minigene-derived transcripts (upper panel). Exon 3b relative inclusion analysis was done as described in *Chapter* 3.9. In the bottom panel the results from semi-quantitative RT-PCR experiments were quantified and are graphically displayed on a logarithmic scale for proportional visualization of positive and negative changes. **B.** Effect on endogenous RAC1 alternative splicing in HeLa cells. RAC1b and tubulin were detected by western blot with specific antibodies and densitometrically quantified by ImageJ. RAC1b relative expression was done as described in *Chapter* 3.9. In the bottom panel the results from western blot experiments were quantified and are graphically displayed on a logarithmic scale for proportional visualization of positive and negative changes. Asterisks (*) indicate the results considered significant by a two-way t-test assuming equal variances, in which n=2 and p<0,5.

4.2.1. Subcellular localization of PTBP1, ESRP1 and RANBP2

To determine the subcellular localization of transfected PTBP1, ESRP1 and RANBP2, an immunoflouresce assay was performed in NCM460 and HeLa cells. **Figure 4.7** shows a very similar subcellular localization of the transfected proteins between the two cell lines. PTBP1 and ESRP1 can be found both in the nucleus and cytoplasm, being PTBP1 more localized in the nucleus in HeLa cells when compared to NCM460 cells. RANBP2 is found at the nuclear membrane and in the cytoplasm. PTBP1 and ESRP1 are both splicing factors and RANBP2 is part of the nucleopore complex, which correlates perfectly with their determined localization in the nucleus or at the nuclear membrane, respectively. The cytoplasmic localization is explained by the overexpression experiment itself and, for PTBP1 and ESRP1 also due to the known shuttling of some splicing factors between both compartments. Besides giving information on the proteins' localization, this experiment also allows to evaluate transfection efficency, which appears to be similar for both cell lines, as shown in **Figure 4.7**.

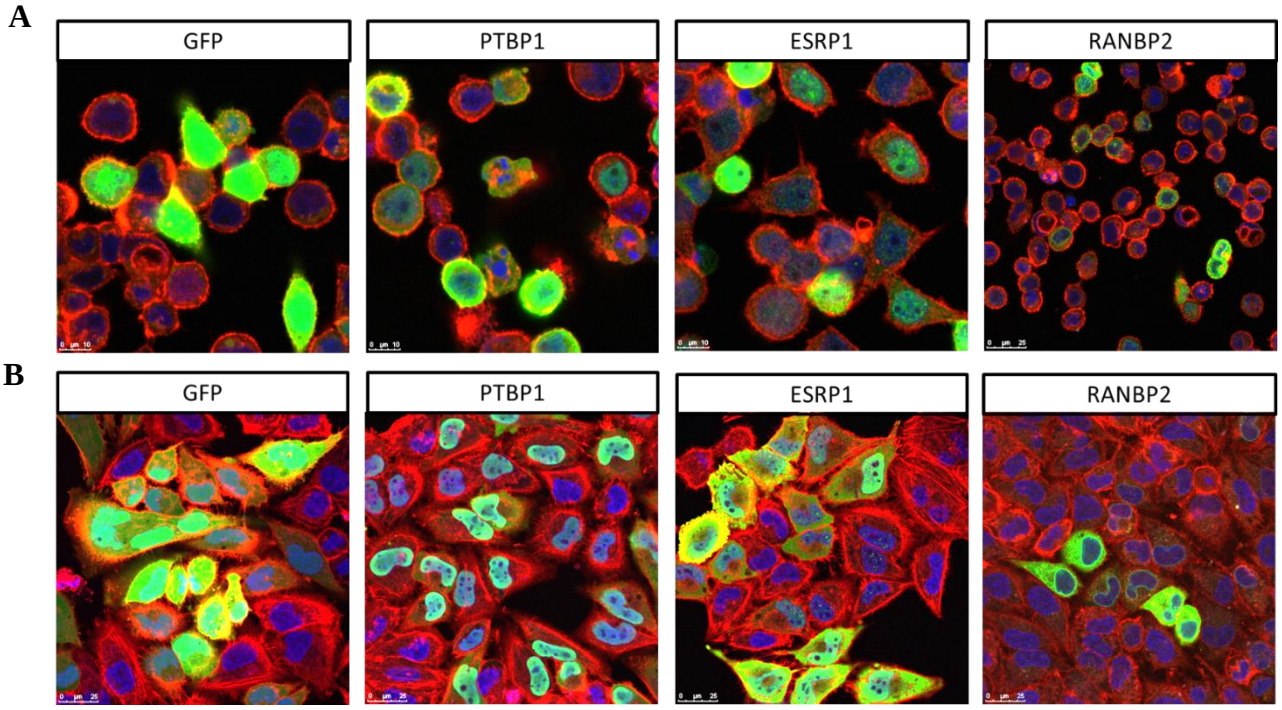

Figure 4.7 – Subcellular localization of transfected PTBP1, ESRP1 and RANBP2. NCM460 cells **(A)** and HeLa cells **(B)** were transfected with pcDNA3_GFP (GFP), pcDNA3_myc_PTBP1 (PTBP1), pcDNA3_myc_ESRP1 (ESRP1) and pkTol2Chy_RANBP21-3224 (RANBP2). 24 h later cells were fixed and treated as described in *Chapter* 3.8 to be analysed by confocal immunofluorescence microscopy. Shown are overlay images from cell staining for the indicated transfected proteins (green), actin (red), and DAPI (blue).

4.3. The effect of ESRP1 and RANBP2 depletion in the alternative splicing of endogenous RAC1 transcripts

To confirm the observed results of the minigene approach for ESRP1 and RANBP2 overexpression, the endogenous expression of each of these proteins was depleted in NCM460 cells using commercially available siRNAs. First, we performed a pilot experiment to determine the time of incubation after transfection at which depletion was the most efficient. For that, a set of transfection experiments with different periods of incubation was performed. After 24, 48, 72 and 96 hours of incubation transfected NCM460 cells were lysed, total RNA was extracted, and transcript expression was assayed by RT-PCR. As shown in **Figure 4.8** the suppression of ESRP1 and RANBP2 transcripts was more efficient after 72 h of incubation reaching 95% and 61% of depletion, respectively. Ideally, depletion efficiency should also be assessed at the protein level to ensure that the siRNA is working properly and that the effect observed results from the target protein depletion and no other possible non-specific effect of the siRNA. ESRP1 depletion efficiency could not be assessed at the protein level because there was no ESRP1 antibody available. RANBP2 depletion efficiency was successfully confirmed by western blot analysis, however in this particular cell lysate technical problems with the gel run did not allow proper documentation.

Subsequently, the expression of endogenous RAC1b protein and transcript was assayed under the same experimental conditions. The respective results of Western blot and semi-quantitative RT-PCR showed that ESRP1 depletion decreased endogenous RAC1b protein and transcript levels, whereas RANBP2 depletion increased endogenous RAC1b expression (**Figure 4.9**). Note that the effect of ESRP1 and RANBP2 depletion on RAC1b protein and transcript levels increased with incubation time, except for RANBP2 protein depletion at 96 h (**Figure 4.9 B**).

Based on the previous data for both siRNAs efficiency, the chosen incubation time to repeat the experiment in order to verify the results was 72 h.

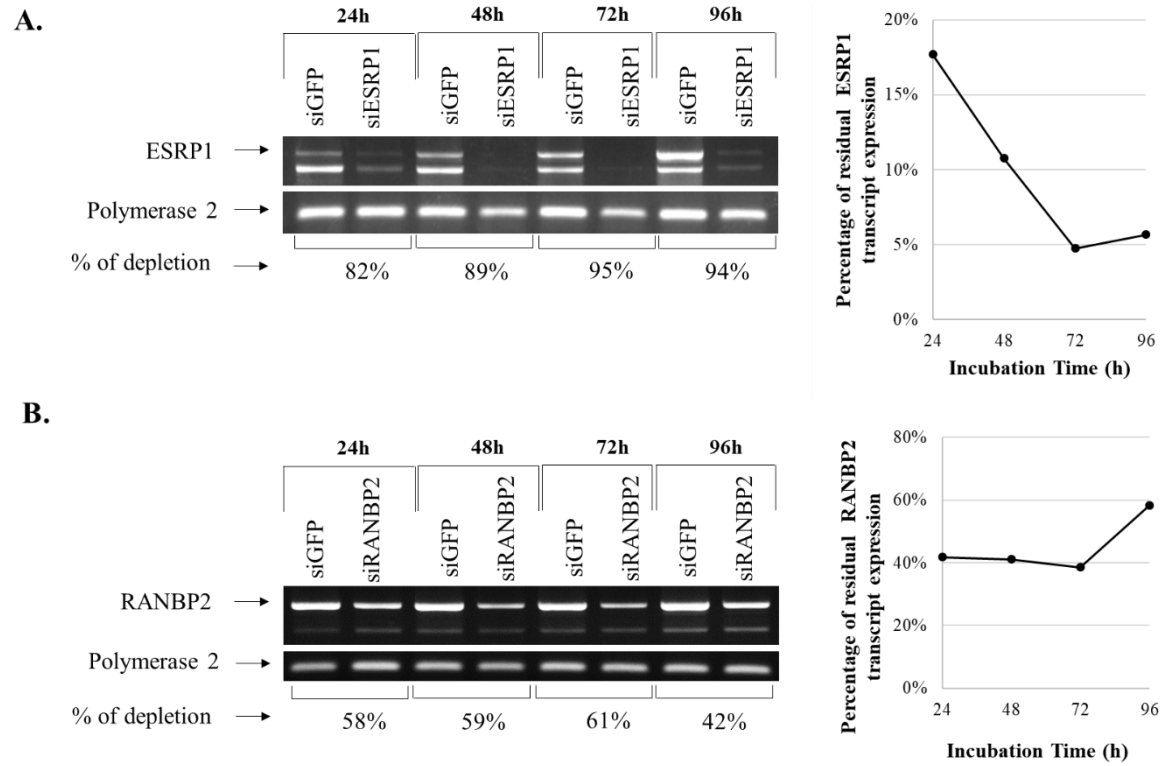

Figure 4.8 – Time course experiment of siRNA-mediated depletion of ESRP1 and RANBP2. NCM460 cells were transfected with siESRP1 (**A)** and siRANBP2 (**B**), or siGFP as control. Cell lysis was performed after 24, 48, 72 and 96 h. Left panels: Digitally recorded images of the 2% agarose gel with the RT-PCR products specific for ESRP1 (**A**) RANBP2 (**B**) and Polymerase 2 (Pol2) as control. Right panels: Graphic display of the percentage of residual of ESRP1 (**A**) and RANBP2 (**B**) transcript expression during the indicated incubation period. Percentage of transcript depletion was calculated as described in Chapter 2.9.

The effect of ESRP1 and RANBP2 depletion on RAC1 alternative splicing was also studied in HeLa and HT29 cells, in order to understand if the results obtained were independent of the cell line. Therefore, NCM460, HeLa and HT29 cells were transfected with specific siRNAs against GFP, ESRP1 and RANBP2, lysed after 72 h of incubation, and endogenous RAC1b protein and transcript levels determined. As shown in **Figure 4.10**, the specific siRNAs depleted both ESRP1 and RANBP2 in all cell lines under the previously chosen experimental conditions. For RANBP2, the protein depletion was also confirmed by western blot (**Figure 4.10 B,** lower panel) because there was a specific endogenous antibody available at the lab. The difference observed in the depletion efficiency at the transcript and protein level will be discussed in *Chapter 5*. In NCM460 cells, ESRP1 depletion decreased endogenous RAC1b expression and RANBP2 depletion increased endogenous RAC1b, confirming the previous results shown in **Figure 4.9**. The results obtained for HeLa and HT29 cells, both at the transcript and protein levels, followed the same tendency observed in NCM460 cells. Overall, depleting ESRP1 decreased endogenous RAC1b expression while RANBP2 depletion increased it (**Figure 4.11**), independently of the cell line used. The non-significance observed in the t-test values for some of the depletion experiments performed on HeLa and HT29 cells can be explained by the larger variations observed between experiments (discussed in *Chapter 5*).

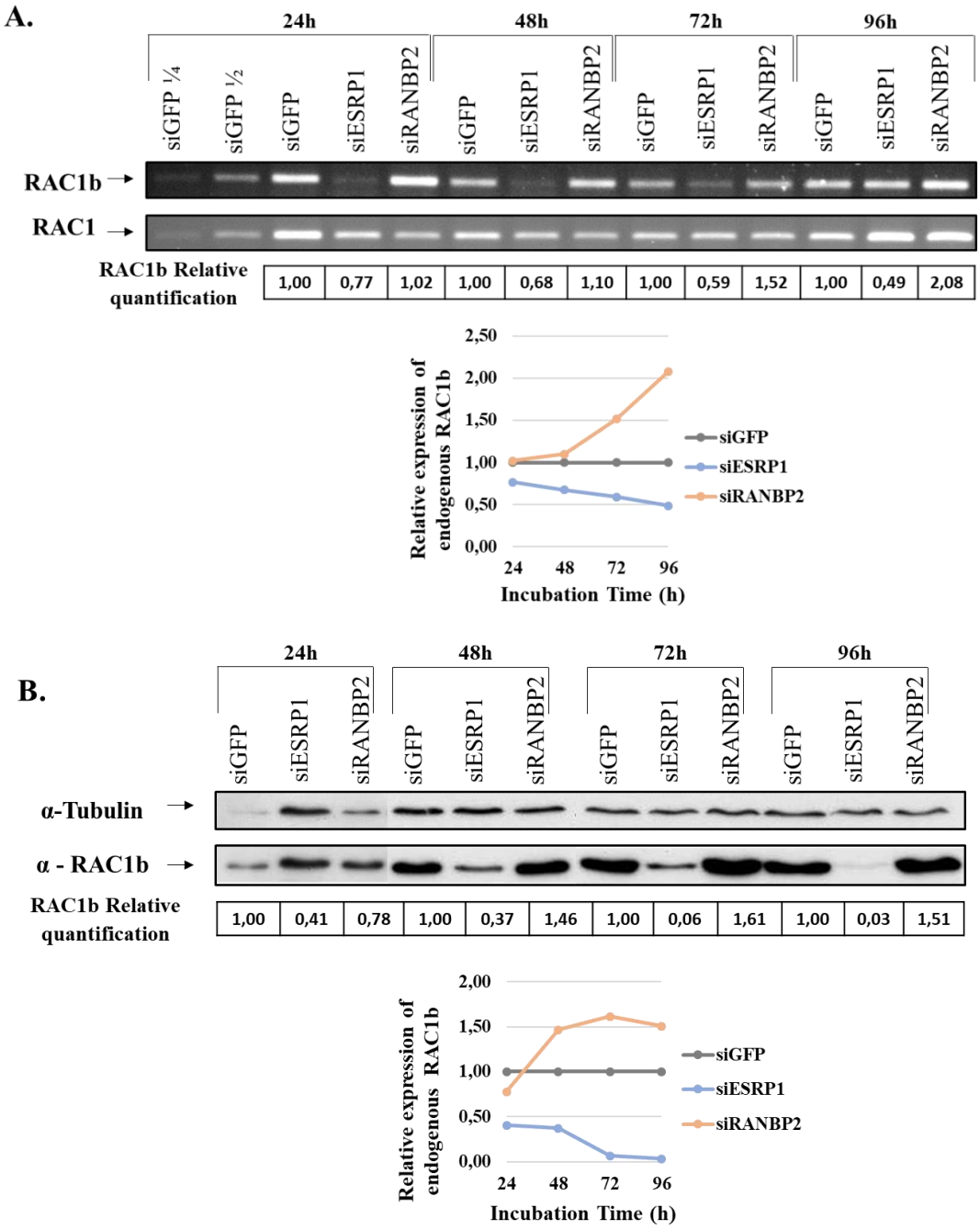

Figure 4.9 – Analysis of endogenous RAC1b expression after depletion of ESRP1 and RANBP2 in NCM460 cells. A. Effect of ESRP1 and RANBP2 depletion on RAC1 alternative splicing at the transcript level. A specific semi-quantitative RT-PCR for the endogenous RAC1 splice variants was preformed (upper panel), and densitometrically quantified by Image J. RAC1b relative quantification was done as described in *Chapter* 3.9. On the bottom panel is the graphic display of the relative quantification of RAC1b endogenous transcript during the indicated incubation period. **B.** Effect of ESRP1 and RANBP2 depletion on RAC1 alternative splicing at the protein level. RAC1b and tubulin were detected by western blot (upper panel) and densitometrically quantified by ImageJ. RAC1b relative quantification was done as described in *Chapter* 3.9. The bottom panel shows the graphic display of the relative quantification of RAC1b endogenous protein during the indicated incubation period.

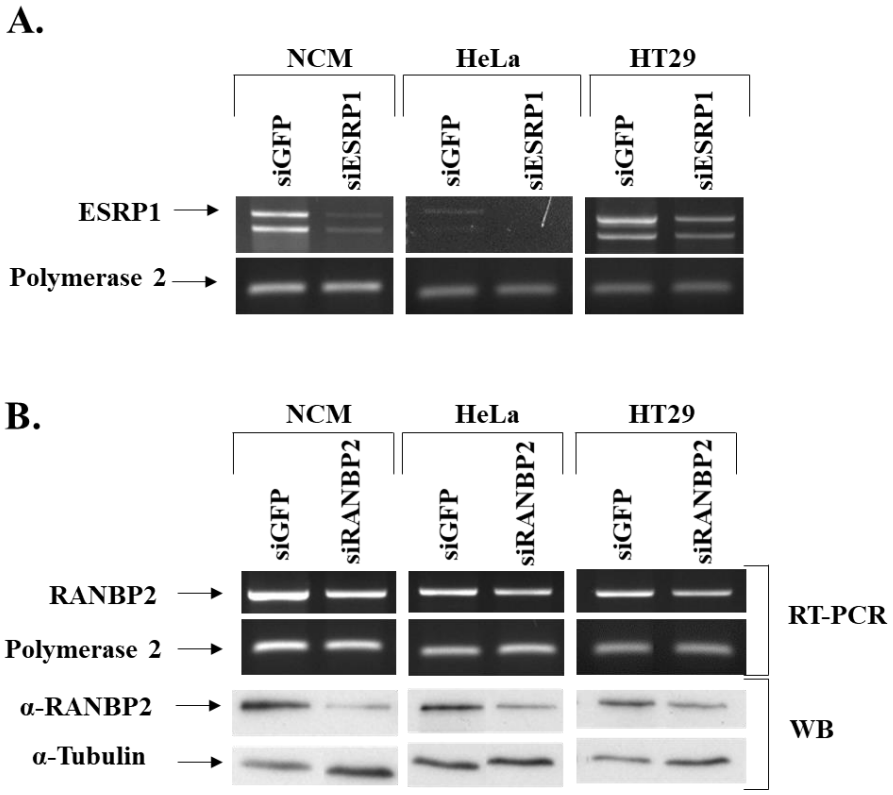

Figure 4.10 – ESRP1 and RANBP2 depletion in NCM460, HeLa and HT29 after 72 h incubation. A. Confirmation of ESRP1 transcript depletion in NCM460, HeLa and HT29 cells. Shown are digitally recorded images of the 2% agarose gel with the RT-PCR products specific for ESRP1 and Polymerase 2 as control. **B.** Confirmation of RANBP2 transcript and protein depletion in NCM460, HeLa and HT29 cells. Upper panel: Digitally recorded images of the 2% agarose gel with the RT-PCR products specific for RANBP2 and Polymerase 2 as control. Lower panel: Western blot detection of RANBP2 and tubulin (loading control) with specific antibodies.

Altogether, the results obtained in the RAC1 minigene approach for RANBP2 in NCM460 cells were sustained by the results found for endogenous RAC1b, in both overexpression and depletion experiments. Thus, RANBP2 emerged as a candidate regulator of RAC1 alternative splicing that promotes the skipping of exon 3b in NCM460 cells. By contrast, the results for ESRP1 overexpression were not corroborated by the depletion experiments, this can be explained by several factors, which will be discussed in *Chapter 5*. Even though, in overexpression experiments RANBP2 effect was NCM460 cell-dependent and ESRP1 effect was cell line-independent, in depletion experiments both factors acted cell line- independent, reinforcing their effects in the regulation of the alternative splicing of RAC1.

Overall, RANBP2 was identified as a candidate regulator of RAC1 alternative splicing in colorectal cells, while for ESRP1 strong evidences were provided supporting its possible role as a regulator of RAC1b expression.

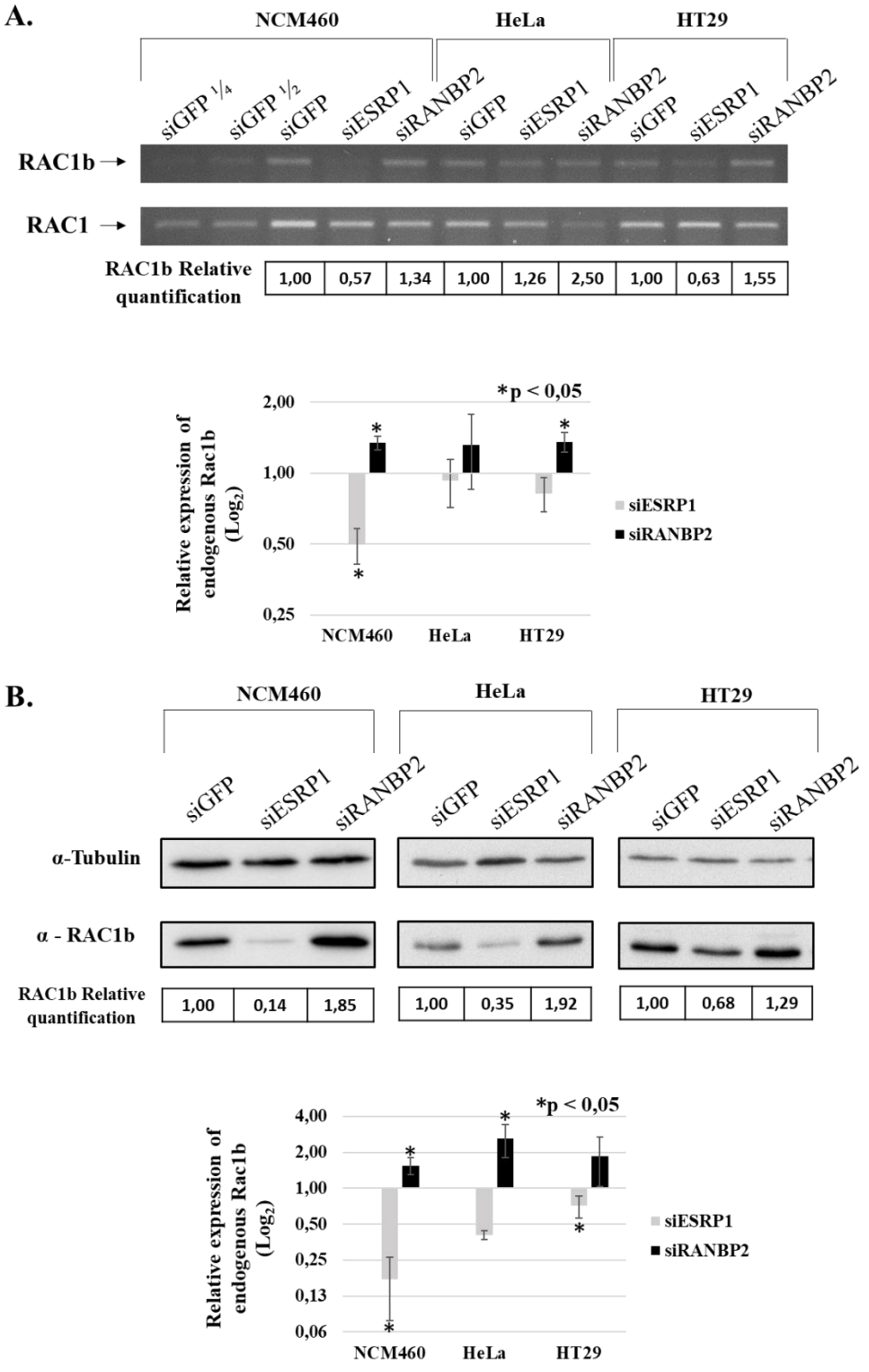

Figure 4.11 – Analysis of endogenous RAC1b expression in NCM460, HeLa and HT29 cells. A. Effect of ESRP1 and RANBP2 depletion on RAC1 alternative splicing at the transcript level. A specific semi-quantitative RT-PCR for endogenous RAC1 splice variants was preformed (upper panel), and the endogenous transcripts were densitometrically quantified by Image J. **B.** Effect of ESRP1 and RANBP2 depletion on RAC1 alternative splicing at the protein level. RAC1b and tubulin were detected by western blot (upper panel) and densitometrically quantified by ImageJ. **A, B.** RAC1b relative quantification was done as described in *Chapter* 3.9. On the bottom panel the results from semi-quantitative RT-PCR and western blot experiments are displayed graphically on a logarithmic scale for proportional visualization of positive and negative changes. Asterisks (*) indicate the results considered significant by a two-way t-test assuming equal variances, in which for NCM460 n=3 and for HeLa and HT29 n=2, p<0,5.

5. Discussion, Conclusions and Future Perspectives

The mammalian RAC1 gene can originate two alternative transcripts, RAC1 and RAC1b. Usually, exon 3b is skipped and RAC1 is the predominant transcript produced (Gonçalves et al. 2009). However, the 3b-containing variant, RAC1b, was found to be overexpressed in several malignant tumours including colorectal, breast, lung, thyroid and pancreas (Schnelzer et al. 2000; Matos and Jordan 2008; Stallings-Mann et al. 2012; Silva et al. 2013; Mehner et al. 2014). Due to its hyperactive properties and selective overexpression in cancerous tissue (Matos and Jordan 2008), RAC1b is considered a promising therapeutic target. Identification and characterization of the mechanisms involved in the regulation of RAC1 alternative splicing will provide essential knowledge regarding the cellular events that lead to aberrant RAC1 signalling. Eventually, this information will be useful for a better understanding of tumour progression and the development of effective pharmacological modulators able to restore normal RAC1 signalling. So far, the host lab has already identified SRSF1 and SRSF3 as antagonistic regulators of RAC1 alternative splicing in colorectal cells (Gonçalves et al. 2009). Also, SRSF1 protein levels were found to increase with the inhibition of the PI3K pathway, while SRSF3 expression was shown to increase with the activation of β-catenin/TCF4 (Gonçalves et al. 2009). Furthermore, kinases SRPK1 and GSK3β were found required to sustain RAC1b levels and both were shown to act upon the phosphorylation of splicing factor SRSF1 (Gonçalves et al. 2014). It can be expected that other splicing factors also contribute to RAC1 alternative splicing regulation.

With this work the two candidate splicing regulators PTBP1 and ESRP1 that have been related to RAC1b expression in recent publications, as well as the nucleo-cytoplasmic transport regulator RANBP2, were studied in colorectal cells (Saitoh et al. 2012; Ishii et al. 2014; Hollander et al. 2016; Vecchione et al. 2016). To do so, the possible effects of PTBP1, ESRP1 and RANBP2 overexpression on RAC1 alternative splicing were first determined resorting to the RAC1 minigene (Gonçalves et al. 2009). According to the results, significant effects were then confirmed at the endogenous level through the use of a commercially available siRNAs to deplete the regulators. The experiments were mainly performed in NCM460 colon cells (Moyer et al. 1996) but confirmed using HeLa and HT29 cells to determine if the observed results were cell line dependent.

With regard to RANBP2, the results from the overexpression experiments in NCM460 cells were corroborated by the results from the depletion experiments, which led to the conclusion that this nucleoporin is promoting the skipping of exon 3b in NCM460 cells (**Figures 4.5 and 4.11**). These results are in agreement with the finding that RANBP2 enhances transcriptional activity by increasing the nuclear import of TCF-4 and β-catenin in HCT116 and DLD1 cells (Shitashige et al. 2008). Since β-catenin/TCF4 activation is known to promote SRSF3 expression, it is plausible that RANBP2 might be promoting exon 3b skipping by promoting the splice silencer SRSF3 through the activation of β-catenin/TCF4. This could be tested in future experiments by determining the SRSF3 expression levels under these conditions, both at the transcript and protein level. Another interesting report that concurs with this work's findings, is the discovery that the levels of RANBP2 were elevated in transgenic mouse models of prostate cancer constitutively expressing a PI3K catalytic subunit (PIK3CA) and treating the animals with a PI3K inhibitor decreased RANBP2 protein abundance (Renner et al. 2007). This means that RANBP2 expression is influenced by PI3K signalling. Interestingly, RAC1b and SRSF1 expression increased upon PI3K inhibition (Gonçalves et al. 2009), a result antagonistic to the effect on RANBP2. It is possible that the effect of PI3K signalling on SRSF1 might be mediated by RANBP2 expression because RANBP2 is involved in nuclear protein import and might be inhibiting SRSF1 expression and consequently promoting exon 3b skipping. This could be tested in future experiments by treating colorectal cells with PI3K inhibitors and then correlating the RANBP2 and SRSF1 expression levels, both at the transcript and protein level. Another way that RANBP2 could be influencing RAC1b expression is through its role in the regulation of the speckled distribution of

phosphorylated pre-mRNA processing factors (Saitoh et al. 2012). In RANBP2 knockdown cells, SRSF1 was shown to be localized at both the cytoplasmic granules and nucleus. It is possible that RANBP2 knockdown resulted in the maintenance of SRSF3 in cytoplasmic granules, reducing is amount at the nucleus. The lack of SRSF3 and the presence of SRSF1 in the nucleus mediated by the absence of RANBP2 results in the increase of RAC1b expression. Altogether, RANBP2 might be promoting exon 3b skipping by either activating SRSF3 and/or inhibiting SRSF1, the already known regulators of RAC1 alternative splicing. Further experiments must be performed in order to understand the relation between RANBP2, SRSF1 and SRSF3.

ESRP1 is an epithelial cell-specific protein that along with ESRP2, enforces genome-wide epithelial splicing programs in diverse epithelial cell types (Yang and Carstens 2017). This protein is normally highly express in both colon and rectal tissues (https://www.proteinatlas.org/ENSG00000104413-ESRP1/tissue), which indicates that ESRP1 may have an important role in the maintenance of these tissues. When ESRP1 was overexpressed in NCM460 cells, exon 3b skipping was promoted (**Figure 4.5**). As so, logically, in the depletion experiments of ESRP1 it would be expected to see an increase in the exon 3b inclusion, however, that was not the case. Interestingly, with ESRP1 depletion exon 3b skipping was also promoted (**Figure 4.11**). Essentially, both overexpression and depletion experiments of ESRP1 resulted in RAC1b expression decrease. Knowing that ESRP1 expression is responsible for maintaining the epithelial phenotype, it is not farfetched to say that this protein might be tightly regulated. It is likely that cells keep ESRP1 concentration at a certain level, keeping it from rising above a defined concentration. This type of regulation can be accomplished through a negative feedback mechanism, in which ESRP1 regulates its own expression. Thus, by overexpressing ESRP1 we might be inhibiting its expression, and consequently, the effect observed in overexpression experiments would correspond to the depletion experiments effect. This type of regulation has been reported for the splicing factor hnRNP L, that when in excess activates NMD of its own mRNA by including the 'poison exon' in its final transcript, creating therefore a negative autoregulatory feedback loop responsible for keeping the homeostasis of hnRNP L levels (Rossbach et al. 2009). In future studies, this hypothesis can be tested by comparing ESRP1 protein levels in a control situation with a ESRP1 overexpressing situation. Interestingly, data shared with the host lab by Doctor Russ P. Carstens from the department of genetics of the Perelman school of medicine, showed that in RNAseq data from combined small and large intestine epithelial cells from ESRP1/2 knock-out mice (Yang and Carstens 2017), RAC1b expression appears to be abolished when compared to the controls, although the overall RAC1b expression level in the control cells was low. This information supports the result found in this work in colorectal cells, encouraging more studies regarding ESRP1 and RAC1b. Contrary to the result observed in colorectal cells, ESRP1 was found to promote the skipping of exon 3b in the RAC1 transcript in SAS and HSC4 cells, both tongue squamous cell carcinoma (Ishii et al. 2014). The difference in the results is probably related with the fact that the cells are from different tissues and consequently differ in the regulation mechanisms. Taking all this information into account, ESRP1 can be considered in NCM460 cells as a possible candidate regulator of RAC1 alternative splicing that promotes the inclusion of exon 3b. It should also be noted that an ESRP-binding motif (UGGUGG) is present in the intron upstream of exon 3b suggesting that ESRP1 directly regulates the alternative splicing of RAC1 mRNA (Ishii et al. 2014).

PTBP1 and PTBP2 depletion was previously reported to promote the skipping of exon 3b in RAC1 pre-mRNA in HCT116 (human epithelial colorectal carcinoma cells with KRAS mutation) (Hollander et al. 2016). As shown in **Figure 4.5**, the overexpression of PTBP1 in NCM460 cells had no significant effect on RAC1 alternative splicing. Both, NCM460 and HCT116 cells are human epithelial colorectal cells, however, HCT116 cells are carcinogenic unlike NCM460 cells. This means that these cells have different genetic properties, which can explain a different response to PTBP1

overexpression. In particular, HCT116 cells have an oncogenic mutation in the KRAS gene, in which cell proliferation is promoted via the ERK pathway and cell survival is stimulated via PI3K and RACl signalling (Seruca et al. 2009). In tumours with an oncogenic BRAF mutation ERK pathway is still stimulated, however, BRAF lies downstream of KRAS and by itself cannot activate RACl. In these tumours, the survival stimulus is achieved by overexpression of hyperactive RAC1 variant, RAClb (Matos et al. 2008). So, RAC1b expression seems to be differentially regulated depending on the cell needs, and it is likely that PTBP1 can regulate exon 3b inclusion in HCT116 cells but not in NCM460 cells. Nevertheless, it is possible that the lack of effect upon overexpression of PTBP1 in NCM460 cells might be related to the fact that the protein is already highly expressed in colon and rectal tissues just like ESRP1 (https://www.proteinatlas.org/ENSG00000011304-PTBP1/tissue). Taking this into account and the fact that the published PTBP1 results in HCT116 cells were only observed with depletion experiments, it would be cautious to also perform an siRNA experiment for PTBP1 in future experiments.

To understand if the effects of PTBP1, ESRP1 and RANBP2 on RAC1b expression were independent of the NCM460 cell line, both the overexpression and depletion experiments were performed in HeLa cells, which also express RAC1b but are of cervix adenocarcinoma origin. The overexpression experiments described in the **Figures 4.5 and 4.6** suggest that ESRP1 is a general RAC1b expression regulator, while the effect of RANBP2 was specific for NCM460 cells and that of PTBP1 for HeLa cells. However, subsequent depletion experiments, also performed in HT29 cells as a further specificity control, gave the indication that both ESRP1 and RANBP2 are more ubiquitous RAC1b expression regulators.

In the depletion experiments using HeLa cells, the effects of ESRP1 on RAC1b expression at both transcript and protein level were considered not statistically significant by the t-test. These results can be explained by the fact that HeLa cells express very little ESRP1, in agreement with published data (Ishii et al. 2014) and the results shown in Figure 4.10A. Therefore, lowering the concentration of a poorly expressed protein might not result in a significant effect. RANBP2 depletion effect on RAC1 alternative splicing at the transcript level was also considered not significant by the t-test, however, reach statistical significance at the protein level. In part, this lack of significance can also be explained by the fact that the incubation period with the siRNA used on HeLa cells was optimized in NCM460 cells. As mentioned before, these two cell lines have different metabolic and cell division rates, which will influence the siRNA half-life and thus the ideal incubation period with the siRNA to observe an effect on RAC1 alternative splicing. So, what may be happening is that RANBP2 depletion was losing efficiency after 72 h by the dilution effect during cell division and/or by degradation, leading to the recovery of the protein expression. Consequently, the effect of RANBP2 depletion at the RAC1b transcript level may be already declining while at the protein level it is still detectable. This problem can also be happening for ESRP1 depletion, but in this case the declining of the effect already reached the protein level. This phenomenon of declining depletion was observed in **Figure 3.8B** with RANBP2 protein level in NCM460 cells after 96 h of incubation (This effect might happen earlier in HeLa and HT29 cells). Another information that is important to consider is that transfection efficiencies were not the same for both independently performed experiments. This difference of efficiency resulted in a wide variation of the results, which also led to the observed lack of significance by the t-test.

The depletion experiments were also performed in HT29, a colorectal adenocarcinoma cell line with BRAF mutation and high expression of RAC1b. While NCM460 are normal epithelial colorectal cells, HT29 are carcinogenic epithelial colorectal cells, with different properties. ESRP1 depletion effect on RAC1 alternative splicing was only considered significant by the t-test at the protein level, while RANBP2 depletion effect was only significant at the transcript level. This might be explained, again, by the fact that the incubation periods with the siRNA used on HT29 cells were

optimized in NCM460 cells. Thus, it is possible that ESRP1 depletion effect was already reverting at the RAC1b transcript level but not at the protein. Contrarily, RANBP2 depletion effect was detectable at the RAC1b transcript but not the protein level. As mention above, HT29 cell are characterized by a high RAC1b expression, essential for cell cycle and survival. This which might mean that RAC1b expression is tightly regulated in HT29 cells. It is possible that RANBP2 depletion is increasing exon 3b inclusion (observed at the transcript level, **Figure 4.11**), however, post-translational regulation mechanisms may keep RAC1b protein levels at a steady concentration, inhibiting the effect at the protein level. Efficiency difference may also be contributing to the results variation, which influences the significance attributed by the t-test.

To confirm the tendencies observed in HeLa and HT29 more experiments need to be performed to determine the right incubation time with the siRNA for each cell line. Afterwards, several independent experiments must be done to raise the statistical significance. ESRP1 depletion efficiency was not assessed at the protein level because there was no ESRP1 antibody available. When depleting a protein with siRNAs it is important to understand that it takes time from the moment that the siRNA decreases the mRNA amount to the moment protein level reduction is observed. This was very evident when the depletion efficiency at the transcript and protein level were observed for RANBP2. Future experiments should therefore include detection of ESRP1 protein level upon depletion in order to ensure that the siRNA is working properly and that the effect observed on RAC1b expression results from ESRP1 protein depletion and no other possible non-specific effect of the siRNA.

RANBP2 was identified as a candidate oncogene overexpressed in the subgroup of human colorectal cancers with microsatellite instability (Gylfe et al. 2013). Corroborating with this information, in another study this nucleoprotein was also reported to be overexpressed in human colorectal cancers with microsatellite instability (Dunican et al.). In addition, RANBP2 was found to protect $BRAF^{V600E}$ mutant colon cancers cells from undergoing mitotic cell death (Vecchione et al. 2016). All this information leads to the conclusion, that somehow RANBP2 is involved in the survival of human colorectal cancers with microsatellite instability. $BRAF^{V600E}$-positive colorectal tumours are often characterized, among other aspects, by the presence of microsatellite instability phenotype and frequent overexpression of RAC1b (Matos et al. 2016). RANBP2 role in $BRAF^{V600E}$-positive colorectal tumours could be RAC1b dependent, since both seem to be promoting colorectal tumours survival. However, in this work, the depletion of RANBP2 in $BRAF^{V600E}$-positive HT29 cells led to an increase of exon 3b inclusion, the opposite of what was expected. This indicates that the described effect of RANBP2 on the protection of $BRAF^{V600E}$ mutant colon cancers cells from undergoing mitotic cell death are not RAC1b related. Probably in the absence of RANBP2, $BRAF^{V600E}$ positive cells might promote RAC1b expression in order to increase their survival chances. However more studies need to be done in order so understand the possible mechanisms by which RANBP2 is influencing RAC1b expression in HT29 cells.

Even though more experiments need to be done to confirm, the results from this thesis provided strong evidence that ESRP1 and RANBP2 are involved in RAC1b expression regulation in colorectal cells and identified for the first time other factors besides SRSF1 and SRSF3 that are involved in the regulation of RAC1b. Further experiments are needed to clarify how these proteins are regulating RAC1 alternative splicing, for example, whether ESRP1 binds directly to the RAC1 pre-mRNA or whether RANBP2 acts by modulating the nuclear levels of SRSF1 or SRSF3. This knowledge will be useful to characterize RAC1 alternative splicing regulation mechanisms and eventually to develop effective pharmacological modulators that can restore normal RAC1 signalling in tumour cells.

6. References

Alonso-Espinaco V, Cuatrecasas M, Alonso V, Escudero P, Marmol M, Horndler C, Ortego J, Gallego R, Codony-Servat J, Garcia-Albeniz X, Jares P, Castells A, Lozano JJ, Rosell R, Maurel J (2014) RAC1b overexpression correlates with poor prognosis in KRAS/BRAF WT metastatic colorectal cancer patients treated with first-line FOLFOX/XELOX chemotherapy. Eur J Cancer 50:1973–1981. doi: 10.1016/j.ejca.2014.04.019

Bernad R, van der Velde H, Fornerod M, Pickersgill H (2004) Nup358/RanBP2 attaches to the nuclear pore complex via association with Nup88 and Nup214/CAN and plays a supporting role in CRM1-mediated nuclear protein export. Mol Cell Biol 24:2373–2384.

Blaustein M, Pelisch F, Srebrow A (2007) Signals, pathways and splicing regulation. Int J Biochem Cell Biol 39:2031–2048. doi: 10.1016/j.biocel.2007.04.004

Braunschweig U, Gueroussov S, Plocik AM, Graveley BR, Blencowe BJ (2013) Dynamic Integration of Splicing within Gene Regulatory Pathways. Cell 152:1252–1269. doi: 10.1016/j.cell.2013.02.034

Brenner H, Kloor M, Pox CP (2014) Colorectal cancer. The Lancet 383:1490–1502. doi: 10.1016/S0140-6736(13)61649-9

Cai Y, Singh BB, Aslanukov A, Zhao H, Ferreira PA (2001) The Docking of Kinesins, KIF5B and KIF5C, to Ran-binding Protein 2 (RanBP2) Is Mediated via a Novel RanBP2 Domain. J Biol Chem 276:41594–41602. doi: 10.1074/jbc.M104514200

Chen M, Manley JL (2009) Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. Nat Rev Mol Cell Biol 10:741–754. doi: 10.1038/nrm2777

Cheung HC, Hai T, Zhu W, Baggerly KA, Tsavachidis S, Krahe R, Cote GJ (2009) Splicing factors PTBP1 and PTBP2 promote proliferation and migration of glioma cell lines. Brain J Neurol 132:2277–2288. doi: 10.1093/brain/awp153

Chuang L-Y, Cheng Y-H, Yang C-H (2013) Specific primer design for the polymerase chain reaction. Biotechnol Lett 35:1541–1549. doi: 10.1007/s10529-013-1249-8

Cieply B, Carstens RP (2015) Functional roles of alternative splicing factors in human disease: Functional roles of alternative splicing factors in human disease. Wiley Interdiscip Rev RNA 6:311–326. doi: 10.1002/wrna.1276

Cline J (1996) PCR fidelity of pfu DNA polymerase and other thermostable DNA polymerases. Nucleic Acids Res 24:3546–3551. doi: 10.1093/nar/24.18.3546

Daguenet E, Dujardin G, Valcarcel J (2015) The pathogenicity of splicing defects: mechanistic insights into pre-mRNA processing inform novel therapeutic approaches. EMBO Rep 16:1640–1655. doi: 10.15252/embr.201541116

Dana H, Chalbatani GM, Mahmoodzadeh H, Karimloo R, Rezaiean O, Moradzadeh A, Mehmandoost N, Moazzen F, Mazraeh A, Marmari V, Ebrahimi M, Rashno MM, Abadi SJ, Gharagouzlo E (2017) Molecular Mechanisms and Biological Functions of siRNA. Int J Biomed Sci IJBS 13:48–57.

Dawlaty MM, Malureanu L, Jeganathan KB, Kao E, Sustmann C, Tahk S, Shuai K, Grosschedl R, van Deursen JM (2008) Resolution of Sister Centromeres Requires RanBP2-Mediated SUMOylation of Topoisomerase IIα. Cell 133:103–115. doi: 10.1016/j.cell.2008.01.045

De Conti L, Baralle M, Buratti E (2013) Exon and intron definition in pre-mRNA splicing: Exon and intron definition in pre-mRNA splicing. Wiley Interdiscip Rev RNA 4:49–60. doi: 10.1002/wrna.1140

Deloria AJ, Höflmayer D, Kienzl P, Łopatecka J, Sampl S, Klimpfinger M, Braunschmid T, Bastian F, Lu L, Marian B, Stättner S, Holzmann K (2016) Epithelial splicing regulatory protein 1 and 2 paralogues correlate with splice signatures and favorable outcome in human colorectal cancer. Oncotarget 7:45. doi: 10.18632/oncotarget.12070

Dunican DS, McWilliam P, Tighe O, Parle-McDermott A, Croke DT (2002) Gene expression differences between the microsatellite instability (MIN) and chromosomal instability (CIN) phenotypes in colorectal cancer revealed by high-density cDNA array hybridization. Oncogene 21:3253–3257. doi: 10.1038/sj.onc.1205431

Esufali S, Charames GS, Pethe VV, Buongiorno P, Bapat B (2007) Activation of Tumor-Specific Splice Variant Rac1b by Dishevelled Promotes Canonical Wnt Signaling and Decreased Adhesion of Colorectal Cancer Cells. Cancer Res 67:2469–2479. doi: 10.1158/0008-5472.CAN-06-2843

Ewing I, Hurley JJ, Josephides E, Millar A (2014) The molecular genetics of colorectal cancer. Frontline Gastroenterol 5:26–30. doi: 10.1136/flgastro-2013-100329

Fagoonee S, Picco G, Orso F, Arrigoni A, Longo DL, Forni M, Scarfò I, Cassenti A, Piva R, Cassoni P, Silengo L, Tolosano E, Aime S, Taverna D, Pandolfi PP, Brancaccio M, Medico E, Altruda F (2017) The RNA-binding protein ESRP1 promotes human colorectal cancer progression. Oncotarget 8:6. doi: 10.18632/oncotarget.14318

Farrell RE (2017) RNA methodologies: laboratory guide for isolation and characterization, Fifth edition. Academic Press, an imprint of Elsevier

Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F (2015) Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012: Globocan 2012. Int J Cancer 136:E359–E386. doi: 10.1002/ijc.29210

Ferre F (1992) Quantitative or semi-quantitative PCR: reality versus myth. Genome Res 2:1–9. doi: 10.1101/gr.2.1.1

Fiegen D, Haeusler L-C, Blumenstein L, Herbrand U, Dvorsky R, Vetter IR, Ahmadian MR (2004) Alternative Splicing of Rac1 Generates Rac1b, a Self-activating GTPase. J Biol Chem 279:4743–4749. doi: 10.1074/jbc.M310281200

Forler D, Rabut G, Ciccarelli FD, Herold A, Köcher T, Niggeweg R, Bork P, Ellenberg J, Izaurralde E (2004) RanBP2/Nup358 provides a major binding site for NXF1-p15 dimers at the nuclear pore complex and functions in nuclear mRNA export. Mol Cell Biol 24:1155–1167

Fu X, Xie F, Gong F, Yang Z, Lv X, Li X, Jiao H, Wang Q, Liu X, Yan L, Xiao R (2018) Suppression of PTBP1 signaling is responsible for mesenchymal stem cell induced invasion of low malignancy cancer cells. Biochim Biophys Acta BBA - Mol Cell Res 1865:1552–1565. doi: 10.1016/j.bbamcr.2018.08.002

Goncalves V, Henriques A, Pereira J, Neves Costa A, Moyer MP, Moita LF, Gama-Carvalho M, Matos P, Jordan P (2014) Phosphorylation of SRSF1 by SRPK1 regulates alternative splicing of tumor-related Rac1b in colorectal cells. RNA 20:474–482. doi: 10.1261/rna.041376.113

Gonçalves V, Matos P, Jordan P (2009) Antagonistic SR proteins regulate alternative splicing of tumor-related Rac1b downstream of the PI3-kinase and Wnt pathways. Hum Mol Genet 18:3696–3707. doi: 10.1093/hmg/ddp317

Gonçalves V, Pereira JFS, Jordan P (2017) Signaling Pathways Driving Aberrant Splicing in Cancer Cells. Genes 9:1. doi: 10.3390/genes9010009

Gylfe AE, Kondelin J, Turunen M, Ristolainen H, Katainen R, Pitkänen E, Kaasinen E, Rantanen V, Tanskanen T, Varjosalo M, Lehtonen H, Palin K, Taipale M, Taipale J, Renkonen–Sinisalo L, Järvinen H, Böhm J, Mecklin J, Ristimäki A, Kilpivaara O, Tuupanen S, Karhu A, Vahteristo P, Aaltonen LA (2013) Identification of Candidate Oncogenes in Human Colorectal Cancers With Microsatellite Instability. Gastroenterology 145:540-543.e22. doi: 10.1053/j.gastro.2013.05.015

Hayakawa A, Saitoh M, Miyazawa K (2016) Dual Roles for Epithelial Splicing Regulatory Proteins 1 (ESRP1) and 2 (ESRP2) in Cancer Progression. In: Atassi MZ (ed) Protein Reviews. Springer, pp 33–40

He X, Arslan AD, Ho T-T, Yuan C, Stampfer MR, Beck WT (2014) Involvement of polypyrimidine tract-binding protein (PTBP1) in maintaining breast cancer cell growth and malignant properties. Oncogenesis 3:e84–e84 . doi: 10.1038/oncsis.2013.47

He X, Pool M, Darcy KM, Lim SB, Auersperg N, Coon JS, Beck WT (2007) Knockdown of polypyrimidine tract-binding protein suppresses ovarian tumor cell growth and invasiveness in vitro. Oncogene 26:4961–4968. doi: 10.1038/sj.onc.1210307

Henriques AFA, Barros P, Moyer MP, Matos P, Jordan P (2015) Expression of tumor-related Rac1b antagonizes B-Raf-induced senescence in colorectal cells. Cancer Lett 369:368–375. doi: 10.1016/j.canlet.2015.08.027

Hertel KJ (2008) Combinatorial Control of Exon Recognition. J Biol Chem 283:1211–1215. doi: 10.1074/jbc.R700035200

Hollander D, Donyo M, Atias N, Mekahel K, Melamed Z, Yannai S, Lev-Maor G, Shilo A, Schwartz S, Barshack I, Sharan R, Ast G (2016) A network-based analysis of colon cancer splicing changes reveals a tumorigenesis-favoring regulatory pathway emanating from ELK1. Genome Res 26:541–553. doi: 10.1101/gr.193169.115

House AE, Lynch KW (2008) Regulation of Alternative Splicing: More than Just the ABCs. J Biol Chem 283:1217–1221. doi: 10.1074/jbc.R700031200

Hutten S, Walde S, Spillner C, Hauber J, Kehlenbach RH (2009) The nuclear pore component Nup358 promotes transportin-dependent nuclear import. J Cell Sci 122:1100–1110. doi: 10.1242/jcs.040154

Ibarra A, Hetzer MW (2015) Nuclear pore proteins and the control of genome functions. Genes Dev 29:337–349. doi: 10.1101/gad.256495.114

Iñiguez LP, Hernández G (2017) The Evolutionary Relationship between Alternative Splicing and Gene Duplication. Front Genet 08:14. doi: 10.3389/fgene.2017.00014

Ishii H, Saitoh M, Sakamoto K, Kondo T, Katoh R, Tanaka S, Motizuki M, Masuyama K, Miyazawa K (2014) Epithelial Splicing Regulatory Proteins 1 (ESRP1) and 2 (ESRP2) Suppress Cancer Cell Motility via Different Mechanisms. J Biol Chem 289:27386–27399. doi: 10.1074/jbc.M114.589432

Jeong HM, Han J, Lee SH, Park H-J, Lee HJ, Choi J-S, Lee YM, Choi Y-L, Shin YK, Kwon MJ (2017) ESRP1 is overexpressed in ovarian cancer and promotes switching from mesenchymal to epithelial phenotype in ovarian cancer cells. Oncogenesis 6:e389 . doi: 10.1038/oncsis.2017.87

Jo YK, Roh SA, Lee H, Park NY, Choi ES, Oh J-H, Park SJ, Shin JH, Suh Y-A, Lee EK, Cho D-H, Kim JC (2017) Polypyrimidine tract-binding protein 1-mediated down-regulation of ATG10 facilitates metastasis of colorectal cancer cells. Cancer Lett 385:21–27. doi: 10.1016/j.canlet.2016.11.002

Jordan P, Brazão R, Boavida MG, Gespach C, Chastre E (1999) Cloning of a novel human Rac1b splice variant with increased expression in colorectal tumors. Oncogene 18:6835–6839. doi: 10.1038/sj.onc.1203233

Joseph J, Dasso M (2008) The nucleoporin Nup358 associates with and regulates interphase microtubules. FEBS Lett 582:190–196. doi: 10.1016/j.febslet.2007.11.087

Kalluri R, Weinberg RA (2009) The basics of epithelial-mesenchymal transition. J Clin Invest 119:1420–1428. doi: 10.1172/JCI39104

Kalyna M, Simpson CG, Syed NH, Lewandowska D, Marquez Y, Kusenda B, Marshall J, Fuller J, Cardle L, McNicol J, Dinh HQ, Barta A, Brown JWS (2012) Alternative splicing and nonsense-mediated decay modulate expression of important regulatory genes in Arabidopsis. Nucleic Acids Res 40:2454–2469. doi: 10.1093/nar/gkr932

Kelemen O, Convertini P, Zhang Z, Wen Y, Shen M, Falaleeva M, Stamm S (2013) Function of alternative splicing. Gene 514:1–30. doi: 10.1016/j.gene.2012.07.083

Kim TK (2015) T test as a parametric statistic. Korean J Anesthesiol 68:540. doi: 10.4097/kjae.2015.68.6.540

Kim TK, Eberwine JH (2010) Mammalian cell transfection: the present and the future. Anal Bioanal Chem 397:3173–3178. doi: 10.1007/s00216-010-3821-6

Kim Y-J, Kim H-S (2012) Alternative Splicing and Its Impact as a Cancer Diagnostic Marker. Genomics Inform 10:74. doi: 10.5808/GI.2012.10.2.74

Koscielny G, Texier VL, Gopalakrishnan C, Kumanduri V, Riethoven J-J, Nardone F, Stanley E, Fallsehr C, Hofmann O, Kull M, Harrington E, Boué S, Eyras E, Plass M, Lopez F, Ritchie W, Moucadel V, Ara T, Pospisil H, Herrmann A, G. Reich J, Guigó R, Bork P, Doeberitz M von K, Vilo J, Hide W, Apweiler R, Thanaraj TA, Gautheret D (2009) ASTD: The Alternative Splicing and Transcript Diversity database. Genomics 93:213–220. doi: 10.1016/j.ygeno.2008.11.003

Kosinski C, Li VSW, Chan ASY, Zhang J, Ho C, Tsui WY, Chan TL, Mifflin RC, Powell DW, Yuen ST, Leung SY, Chen X (2007) Gene expression patterns of human colon tops and basal crypts and BMP antagonists as intestinal stem cell niche factors. Proc Natl Acad Sci 104:15418–15423. doi: 10.1073/pnas.0707210104

Lareau LF, Brooks AN, Soergel DAW, Meng Q, Brenner SE (2007a) The coupling of alternative splicing and nonsense-mediated mRNA decay. Adv Exp Med Biol 623:190–211

Lareau LF, Inada M, Green RE, Wengrod JC, Brenner SE (2007b) Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements. Nature 446:926–929. doi: 10.1038/nature05676

Lee PY, Costumbrado J, Hsu C-Y, Kim YH (2012) Agarose Gel Electrophoresis for the Separation of DNA Fragments. J Vis Exp 62:e3923. doi: 10.3791/3923

Lewis BP, Green RE, Brenner SE (2003) Evidence for the widespread coupling of alternative splicing and nonsense-mediated mRNA decay in humans. Proc Natl Acad Sci 100:189–192. doi: 10.1073/pnas.0136770100

Lodish HF (2016) Molecular cell biology, Eighth edition. W.H. Freeman-Macmillan Learning, New York

Luco RF, Misteli T (2011) More than a splicing code: integrating the role of RNA, chromatin and non-coding RNA in alternative splicing regulation. Curr Opin Genet Dev 21:366–372. doi: 10.1016/j.gde.2011.03.004

Manley JL, Krainer AR (2010) A rational nomenclature for serine/arginine-rich protein splicing factors (SR proteins). Genes Dev 24:1073–1074. doi: 10.1101/gad.1934910

Marei H, Malliri A (2017) Rac1 in human diseases: The therapeutic potential of targeting Rac1 signaling regulatory mechanisms. Small GTPases 8:139–163. doi: 10.1080/21541248.2016.1211398

Matlin AJ, Clark F, Smith CWJ (2005) Understanding alternative splicing: towards a cellular code. Nat Rev Mol Cell Biol 6:386–398. doi: 10.1038/nrm1645

Matos P, Collard JG, Jordan P (2003) Tumor-related Alternatively Spliced Rac1b Is Not Regulated by Rho-GDP Dissociation Inhibitors and Exhibits Selective Downstream Signaling. J Biol Chem 278:50442–50448. doi: 10.1074/jbc.M308215200

Matos P, Gonçalves V, Jordan P (2016) Targeting the serrated pathway of colorectal cancer with mutation in BRAF. Biochim Biophys Acta BBA - Rev Cancer 1866:51–63. doi: 10.1016/j.bbcan.2016.06.003

Matos P, Jordan P (2008) Increased Rac1b Expression Sustains Colorectal Tumor Cell Survival. Mol Cancer Res 6:1178–1184. doi: 10.1158/1541-7786.MCR-08-0008

Matos P, Jordan P (2005) Expression of Rac1b stimulates NF-κB-mediated cell survival and G1/S progression. Exp Cell Res 305:292–299. doi: 10.1016/j.yexcr.2004.12.029

Matos P, Oliveira C, Velho S, Gonçalves V, da Costa LT, Moyer MP, Seruca R, Jordan P (2008) B-RafV600E Cooperates With Alternative Spliced Rac1b to Sustain Colorectal Cancer Cell Survival. Gastroenterology 135:899–906. doi: 10.1053/j.gastro.2008.05.052

Matos P, Skaug J, Marques B, Beck S, Veríssimo F, Gespach C, Boavida MG, Scherer SW, Jordan P (2000) Small GTPase Rac1: Structure, Localization, and Expression of the Human Gene. Biochem Biophys Res Commun 277:741–751. doi: 10.1006/bbrc.2000.3743

Matunis MJ, Wu J, Blobel G (1998) SUMO-1 modification and its role in targeting the Ran GTPase-activating protein, RanGAP1, to the nuclear pore complex. J Cell Biol 140:499–509.

McCutcheon IE, Hentschel SJ, Fuller GN, Jin W, Cote GJ (2004) Expression of the splicing regulator polypyrimidine tract-binding protein in normal and neoplastic brain. Neuro-Oncol 6:9–14. doi: 10.1215/S1152851703000279

Mehner C, Miller E, Khauv D, Nassar A, Oberg AL, Bamlet WR, Zhang L, Waldmann J, Radisky ES, Crawford HC, Radisky DC (2014) Tumor Cell–Derived MMP3 Orchestrates Rac1b and Tissue Alterations That Promote Pancreatic Adenocarcinoma. Mol Cancer Res 12:1430–1439. doi: 10.1158/1541-7786.MCR-13-0557-T

Miller CJ, Kassem HS, Pepper SD, Hey Y, Ward TH, Margison GP (2003) Mycoplasma infection significantly alters microarray gene expression profiles. BioTechniques 35:812-814.

Mitry RR, Hughes RD (eds) (2012) Human cell culture protocols, 3rd ed. Humana Press, New York.

Moyer MP, Manzano LA, Merriman RL, Stauffer JS, Tanzer LR (1996) NCM460, a normal human colon mucosal epithelial cell line. In Vitro Cell Dev Biol Anim 32:315–317.

Naftelberg S, Schor IE, Ast G, Kornblihtt AR (2015) Regulation of Alternative Splicing Through Coupling with Transcription and Chromatin Structure. Annu Rev Biochem 84:165–198. doi: 10.1146/annurev-biochem-060614-034242

Navarro MS, Bachant J (2008) RanBP2: A Tumor Suppressor with a New Twist on TopoII, SUMO, and Centromeres. Cancer Cell 13:293–295. doi: 10.1016/j.ccr.2008.03.011

Oberstrass FC (2005) Structure of PTB Bound to RNA: Specific Binding and Implications for Splicing Regulation. Science 309:2054–2057. doi: 10.1126/science.1114066

Odell ID, Cook D (2013) Immunofluorescence Techniques. J Invest Dermatol 133:1–4. doi: 10.1038/jid.2012.455

Orlichenko L, Geyer R, Yanagisawa M, Khauv D, Radisky ES, Anastasiadis PZ, Radisky DC (2010) The 19-Amino Acid Insertion in the Tumor-associated Splice Isoform Rac1b Confers Specific Binding to p120 Catenin. J Biol Chem 285:19153–19161. doi: 10.1074/jbc.M109.099382

Papasaikas P, Valcárcel J (2016) The Spliceosome: The Ultimate RNA Chaperone and Sculptor. Trends Biochem Sci 41:33–45. doi: 10.1016/j.tibs.2015.11.003

Peifer M (2002) Developmental biology: Colon construction. Nature 420:274–276. doi: 10.1038/420274a

Pestana E, Belak S, Diallo A, Crowther JR, Viljoen GJ (2010) Early, rapid and sensitive veterinary molecular diagnostics - real time PCR applications. Springer.

Phelan MC (1998) Basic Techniques for Mammalian Cell Tissue Culture. Curr Protoc Cell Biol 00:1.1.1-1.1.10. doi: 10.1002/0471143030.cb0101s00

Pichler A, Gast A, Seeler JS, Dejean A, Melchior F (2002) The nucleoporin RanBP2 has SUMO1 E3 ligase activity. Cell 108:109–120.

Poulos MG, Batra R, Charizanis K, Swanson MS (2011) Developments in RNA Splicing and Disease. Cold Spring Harb Perspect Biol 3:a000778–a000778. doi: 10.1101/cshperspect.a000778

Radisky DC, Levy DD, Littlepage LE, Liu H, Nelson CM, Fata JE, Leake D, Godden EL, Albertson DG, Angela Nieto M, Werb Z, Bissell MJ (2005) Rac1b and reactive oxygen species mediate MMP-3-induced EMT and genomic instability. Nature 436:123–127. doi: 10.1038/nature03688

Raices M, D'Angelo MA (2012) Nuclear pore complex composition: a new regulator of tissue-specific and developmental functions. Nat Rev Mol Cell Biol 13:687–699. doi: 10.1038/nrm3461

Renner O, Fominaya J, Alonso S, Blanco-Aparicio C, Leal JFM, Carnero A (2007) Mst1, RanBP2 and eIF4G are new markers for in vivo PI3K activation in murine and human prostate. Carcinogenesis 28:1418–1425. doi: 10.1093/carcin/bgm059

Reverter D, Lima CD (2005) Insights into E3 ligase activity revealed by a SUMO–RanGAP1–Ubc9–Nup358 complex. Nature 435:687–692. doi: 10.1038/nature03588

Roberts RJ (2005) How restriction enzymes became the workhorses of molecular biology. Proc Natl Acad Sci 102:5905–5908. doi: 10.1073/pnas.0500923102

Rossbach O, Hung L-H, Schreiner S, Grishina I, Heiner M, Hui J, Bindereif A (2009) Auto- and Cross-Regulation of the hnRNP L Proteins by Alternative Splicing. Mol Cell Biol 29:1442–1451. doi: 10.1128/MCB.01689-08

Rye C, Wise RR, Jurukovski V, Desaix J-F, Choi JH, Avissar Y (2017) Biology. OpenStax, Rice University.

Saitoh N, Sakamoto C, Hagiwara M, Agredano-Moreno LT, Jiménez-García LF, Nakao M (2012) The distribution of phosphorylated SR proteins and alternative splicing are regulated by RANBP2. Mol Biol Cell 23:1115–1128. doi: 10.1091/mbc.e11-09-0783

Salina D, Enarson P, Rattner JB, Burke B (2003) Nup358 integrates nuclear envelope breakdown with kinetochore assembly. J Cell Biol 162:991–1001. doi: 10.1083/jcb.200304080

Saraswathy N, Ramalingam P (2011) Introduction to proteomics. In: Concepts and Techniques in Genomics and Proteomics. Elsevier, pp 147–158.

Schnelzer A, Prechtel D, Knaus U, Dehne K, Gerhard M, Graeff H, Harbeck N, Schmitt M, Lengyel E (2000) Rac1 in human breast cancer: overexpression, mutation analysis and characterization of a new isoform, Rac1b. Oncogene 19:3013–3020. doi: 10.1038/sj.onc.1203621

Seruca R, Velho S, Oliveira C, Leite M, Matos P, Jordan P (2009) Unmasking the role of KRAS and BRAFpathways in MSI colorectal tumors. Expert Rev Gastroenterol Hepatol 3:5–9. doi: 10.1586/17474124.3.1.5

Shin C, Manley JL (2004) Cell signalling and the control of pre-mRNA splicing. Nat Rev Mol Cell Biol 5:727–738. doi: 10.1038/nrm1467

Shitashige M, Satow R, Honda K, Ono M, Hirohashi S, Yamada T (2008) Regulation of Wnt Signaling by the Nuclear Pore Complex. Gastroenterology 134:1961-1971.e4. doi: 10.1053/j.gastro.2008.03.010

Shkreta L, Bell B, Revil T, Venables JP, Prinos P, Elela SA, Chabot B (2013) Cancer-Associated Perturbations in Alternative Pre-messenger RNA Splicing. Cancer Treat Res 158:41–94. doi: 10.1007/978-3-642-31659-3_3

Shukla S (2005) Competition of PTB with TIA proteins for binding to a U-rich cis-element determines tissue-specific splicing of the myosin phosphatase targeting subunit 1. RNA 11:1725–1736. doi: 10.1261/rna.7176605

Silva AL, Carmo F, Bugalho MJ (2013) RAC1b overexpression in papillary thyroid carcinoma: a role to unravel. Eur J Endocrinol 168:795–804. doi: 10.1530/EJE-12-0960

Singh A, Karnoub AE, Palmby TR, Lengyel E, Sondek J, Der CJ (2004) Rac1b, a tumor associated, constitutively active Rac1 splice variant, promotes cellular transformation. Oncogene 23:9369–9380. doi: 10.1038/sj.onc.1208182

Smith CW, Valcárcel J (2000) Alternative pre-mRNA splicing: the logic of combinatorial control. Trends Biochem Sci 25:381–388.

Stallings-Mann ML, Waldmann J, Zhang Y, Miller E, Gauthier ML, Visscher DW, Downey GP, Radisky ES, Fields AP, Radisky DC (2012) Matrix metalloproteinase induction of Rac1b, a key effector of lung cancer progression. Sci Transl Med 4:142ra95. doi: 10.1126/scitranslmed.3004062

Stamm S, Ben-Ari S, Rafalska I, Tang Y, Zhang Z, Toiber D, Thanaraj TA, Soreq H (2005) Function of alternative splicing. Gene 344:1–20. doi: 10.1016/j.gene.2004.10.022

Sugiyama T, Taniguchi K, Matsuhashi N, Tajirika T, Futamura M, Takai T, Akao Y, Yoshida K (2016) MiR-133b inhibits growth of human gastric cancer cells by silencing pyruvate kinase muscle-splicer polypyrimidine tract-binding protein 1. Cancer Sci 107:1767–1775. doi: 10.1111/cas.13091

Symons M, Settleman J (2000) Rho family GTPases: more than simple switches. Trends Cell Biol 10:415–419.

Takahashi H, Nishimura J, Kagawa Y, Kano Y, Takahashi Y, Wu X, Hiraki M, Hamabe A, Konno M, Haraguchi N, Takemasa I, Mizushima T, Ishii M, Mimori K, Ishii H, Doki Y, Mori M, Yamamoto H (2015) Significance of Polypyrimidine Tract-Binding Protein 1 Expression in Colorectal Cancer. Mol Cancer Ther 14:1705–1716. doi: 10.1158/1535-7163.MCT-14-0142

Thorpe GH, Kricka LJ (1986) Enhanced chemiluminescent reactions catalyzed by horseradish peroxidase. Methods Enzymol 133:331–353.

Vecchione L, Gambino V, Raaijmakers J, Schlicker A, Fumagalli A, Russo M, Villanueva A, Beerling E, Bartolini A, Mollevi DG, El-Murr N, Chiron M, Calvet L, Nicolazzi C, Combeau C, Henry C, Simon IM, Tian S, in 't Veld S, D'ario G, Mainardi S, Beijersbergen RL, Lieftink C, Linn S, Rumpf-Kienzl C, Delorenzi M, Wessels L, Salazar R, Di Nicolantonio F, Bardelli A, van Rheenen J, Medema RH, Tejpar S, Bernards R (2016) A Vulnerability of a Subset of Colon Cancers with Potential Clinical Utility. Cell 165:317–330. doi: 10.1016/j.cell.2016.02.059

Vogelstein B, Fearon ER, Hamilton SR, Kern SE, Preisinger AC, Leppert M, Smits AMM, Bos JL (1988) Genetic Alterations during Colorectal-Tumor Development. N Engl J Med 319:525–532. doi: 10.1056/NEJM198809013190901

Wang Y, Liu J, Huang B, Xu Y-M, Li J, Huang L-F, Lin J, Zhang J, Min Q-H, Yang W-M, Wang X-Z (2015) Mechanism of alternative splicing and its regulation. Biomed Rep 3:152–158. doi: 10.3892/br.2014.407

Wang Z, Liu D, Yin B, Ju W, Qiu H, Xiao Y, Chen Y, Khanna C, Lu C (2017) High expression of PTBP1 promote invasion of colorectal cancer by alternative splicing of cortactin. Oncotarget 8:22. doi: 10.18632/oncotarget.15873

Wennerberg K (2005) The Ras superfamily at a glance. J Cell Sci 118:843–846. doi: 10.1242/jcs.01660

Wilson K, Walker JM (eds) (2009) Principles and techniques of biochemistry and molecular biology, 7th ed. Cambridge University Press, Cambridge

Yang P-C, Mahmood T (2012) Western blot: Technique, theory, and trouble shooting. North Am J Med Sci 4:429. doi: 10.4103/1947-2714.100998

Yang Y, Carstens RP (2017) Alternative splicing regulates distinct subcellular localization of Epithelial splicing regulatory protein 1 (Esrp1) isoforms. Sci Rep 7:1. doi: 10.1038/s41598-017-03180-3

Zhang J, Manley JL (2013) Misregulation of Pre-mRNA Alternative Splicing in Cancer. Cancer Discov 3:1228–1237. doi: 10.1158/2159-8290.CD-13-0253

Zhou H-L, Hinman MN, Barron VA, Geng C, Zhou G, Luo G, Siegel RE, Lou H (2011) Hu proteins regulate alternative splicing by inducing localized histone hyperacetylation in an RNA-dependent manner. Proc Natl Acad Sci 108:E627–E635. doi: 10.1073/pnas.1103344108