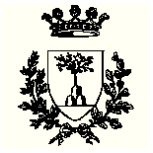


Università degli Studi di Ferrara

DOTTORATO DI RICERCA IN
BIOCHIMICA, BIOLOGIA MOLECOLARE E BIOTECNOLOGIE

CICLO XXVIII

COORDINATORE Prof. Francesco Bernardi

Functional association between the Microprocessor complex and the Spliceosome in the processing of Splice site Overlapping microRNAs

Settore Scientifico Disciplinare BIO/11

Dottoranda
Dott. Pianigiani Giulia

Tutore
Prof. Pagani Franco

Anni 2012/2015

Contents

List of Figures

Abbreviations

The standard abbreviations used in this thesis follow IUPAC rules. All the abbreviations are explained also in the text when introduced for the first time.

A	adenine
aa	amino acid
AGO	argonaute
AS	alternative splicing
bp	base pair
BP	branch Point
BBP	branch point binding protein
BPRS	branch point recognition sequence
C	cytosine
CED	central domain
CLL	chronic lymphocytic leukemia
COL7A1	collagen, type VII, alpha 1
COL17A1	collagen, type XVII, alpha 1
CTT	C-terminal tail
DGCR8	DiGeorge syndrome critical region 8
dsRNA	double-stranded RNA
dsRBD	double-stranded RNA binding domain
EdU	5-ethinyl-2'-desoxyuridin
ESE	exonic splicing enhancer
ESS	exonic splicing silencer
Exp5	exportin-5
FC	fold change
FDR	false discovery rate
FSTL1	follistatin-like 1

G	guanine
HD	hemidesmosome
HITS-CLIP	high-trhroughput sequencing of RNA isolated by crossLinking immunoprecipitation
hnRNP	heterogeneous nuclear ribonucleoprotein
hsa	homo sapiens
ISE	intronic splicing enhancer
ISS	intronic splicing silencer
kb	kilobase
kDa	kilodalton
KSRP	K-homology splicing regulatory protein
LAMB3	laminin subunit Beta 3
m^7G	7-methyl-guanosine
M	A or C
mRNA	messenger RNA
miRNA	microRNA
MPC	Microprocessor complex
N	any nucleotide
NE	nuclear extract
NHEK	normal human epidermal keratinocyte
NLS	nuclear localization signal
nc	non coding
nt	nucleotide
ORF	open reading frame
P	proline
PACT	protein activator of PKR
PCR	polymerase chain reaction
Poly(A)	polyadenilation
pre-mRNA	precursor mRNA
pre-miRNA	precursor miRNA
pri-miRNA	primary miRNA
PPT / $(Y)_n$	polypyrimidine tract
PTB	PPT-binding protein
qRT-PCR	quantitative reverse transcription PCR

R	purine (A or G)
RISC	RNA-induced silencing complex
RIIID	RNase III domain
RLC	RISC loading complex
RNA Pol II	RNA polymerase II
RNA Pol III	RNA polymerase III
RNAi	RNA interference
RRM	RNA recognition motif
RS domain	arginine/serine-rich domain
RT	reverse transcriptase
RT	room temperature
SELEX	systematic evolution of ligands by exponential enrichment
SF3b1	splicing factor 3b subunit 1
siRNA	small interfering RNA
small RNA-seq	small RNA sequencing
snRNA	small nuclear RNA
snRNP	small nuclear ribonucleoprotein
SO-miRNA	Splice site Overlapping miRNA
SR	serine-arginin
SRSF	serine/arginine-rich splicing factor
ss	splice site
ssRNA	single-stranded RNA
T	thymine
TRBP	trans-activation response RNA-binding protein
TSS	transcription start site
TU	transcription unit
U	uracil
U2AF	U2 auxiliary factor
UTR	untranslated region
UV	ultraviolet
WT	wild type
Y	pyrimidine (T or C)

Chapter 1

Introduction

1.1 RNA processing and gene expression

Eukaryotic gene expression is a complex stepwise process: the production of mature, functional RNA molecules requires a series of fine-regulated reactions in order to obtain protein-coding RNAs (messenger RNAs - mRNAs) or non-coding RNAs (ncRNA), participating in transcription, RNA processing, translation or gene regulation. Precursor mRNAs (pre-mRNAs) are transcribed by the RNA polymerase II (Pol II) in the nucleus, afterwards the nascent pre-mRNA passes through a complex network of dynamic interactions and co-transcriptional events that comprise 5'-end capping, introns removal and exons splicing and the 3'-end processing and polyadenylation. The mature mRNA is then released from the site of transcription and exported to the cytoplasm for translation [1]. Distinct and highly complex cellular machineries can carry out each of these steps in the gene expression process, even if more and more evidences suggest that these processes are physically and functionally coupled to one another and occur co-transcriptionally (fig.1.1) [2–4].

1.1.1 5'-end capping

The capping event occurs co-transcriptionally and comprises the addition of 7-methyl-guanosine (m^7G) to the 5'-end of the nascent pre-mRNA. This is fundamental for its stability since it prevents degradation by 5'-3' exonucleases, mediates export from the nucleus to the cytoplasm and promotes translation [5–7].

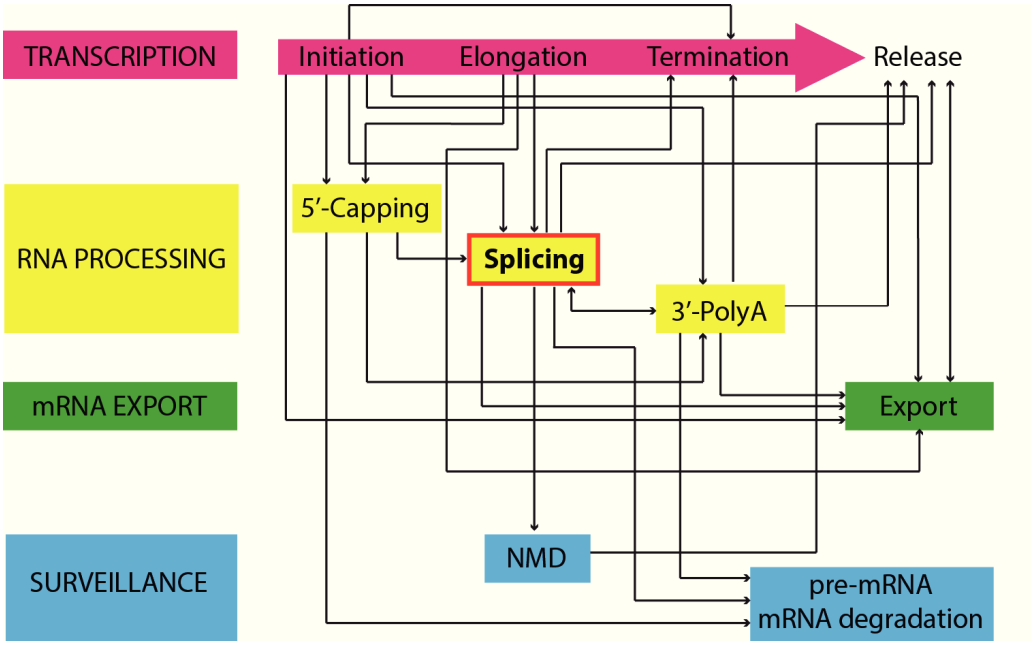


Figure 1.1: **A complex network of coupled interactions in gene expression**
The major events of gene expression are shown on the left (red, yellow, green and blue boxes) and each transcriptional step is shown along the top in the red arrow. "Release" indicates release of the mature mRNA from the site of transcription. The processes within each of the steps are shown below the arrow (colored boxes). The highlighted splicing step makes part of the RNA processing but is also linked with other steps of the gene expression. The black arrows indicate physical and/or functional interactions between components of gene expression machine. Adapted from [1].

1.1.2 3'-end processing and polyadenylation

3'-end processing is a two step reaction that comprises pre-mRNA cleavage and subsequent polyadenylation at the 3' end of the transcript. The cleavage and polyadenylation specificity factor (CPSF) binds specifically to the AAUAAA consensus sequence [8], promoting the cleavage at a 5'-CA-3' dinucleotide located 10-35 bases downstream its binding site. After the cleavage, poly(A) polymerase adds a poly(A) tail composed of 200-250 adenine (A) residues. During this process the intervention of polyadenylate-binding protein 2 (PABP-2) that binds to the nascent poly(A) tail stabilizes and protects pre-mRNA from the exonucleases digestion [9].

1.2 pre-mRNA splicing

A typical mammalian gene is composed of several relatively short coding sequences (exons) that are interrupted by much longer non-coding regions (introns). The coding segments represent a minor portion of the genetic information, in fact account for no more 1% of the entire genome. The translation machinery cannot discriminate introns from exons, which means that the introns have to be removed from the transcript before translation.

The first evidences of the exonic-intronic organization of a pre-mRNA and the discontinuity of genes and, consequently, the first system used to study splicing, was obtained during studies on adenovirus infection [10, 11]. These studies have been soon after expanded to the entire eukaryotic world: Jeffreys and Flavell demonstrated that the rabbit β-globin gene contains a large insert in the coding sequence [12], while a year after, Chambon and collaborators studying the ovalbumin gene structure found that the sequences at the exon-intron boundaries carry some common features suggesting that common signals on the pre-mRNA have a role in the excision-ligation processes during splicing [13, 14].

Splicing is a well conserved pre-mRNA processing mechanism, found from unicellular eukaryotic organism as *S. cerevisiae* to metazoans. It displays increasing levels of regulation and complexity as the number and the length of introns in multicellular eukaryotes increases [15]. The basic consensus sequences (see below), that reside in proximity to the exon-intron boundaries, are also conserved and are found in yeast, plants and vertebrates [16].

To generate correct, mature mRNAs, the introns must be removed from the pre-mRNA and the exons have to be joined together: it occurs in two transesterification reactions performed in the nucleus by a large protein machinery, called the spliceosome, which localizes in the nucleus [1]: its components have a dynamic and particular localization pattern, called the "speckled pattern". These sub-nuclear structures play a role in the organization of gene-expression machineries and they are enriched in pre-mRNA splicing factors that cycle continuously between speckles and other nuclear localizations, such as transcription sites [17, 18]. Splicing factors are recruited from speckles to sites of transcription and when transcription is inhibited splicing factors accumulate in enlarged, rounded speckles: this aspect supports the idea of nuclear speckles as regulators of the pool of factors that are accessible to pre-mRNA splicing machinery [19].

Although splicing occurs in the nucleus, some studies suggest that in particular cell types it can occur in other compartments [20], such as in rat neuronal dendrites [21] or in human anucleate platelets [22].
To distinguish exons from introns, the spliceosomal machinery has to recognize several motifs that reside in proximity of the boundaries: these sequences, which are known as 5' and 3' splice sites (ss), exhibit a variable degree of conservation, therefore the presence of some auxiliary factors, which act in *cis* and *trans* are also required for regulating the splicing reaction [23].

1.2.1 The chemistry of splicing reaction

From a chemical point of view RNA splicing is a quite simple reaction, consisting of two consecutive transesterification steps and involving three reactive groups of pre-mRNA (fig. 1.2) [24]. The initial signals necessary to start the splicing reactions are very conserved motifs at the intron-exon borders and their proximity: the GU dinucleotide at the 5'-splice site (5'ss), the AG dinucleotide at the 3'-splice site (3'ss), the A residue at the branch point (BP) and the polypyrimidine tract ($(Y)_n$ or PPT), both located upstream of the 3'ss.
First, the 2'-hydroxyl group of the A residue at BP attacks the phosphodiester bond at the 5'ss, leading to the cleavage of the 5'-exon from the intron and the concerted ligation of the intron 5'-end to the BP 2'-hydroxyl group in a lariat form: two reaction intermediates are then produced, a detached 5'-exon 1 and an intron-3' exon 2 fragment in a lariat configuration.
The second transesterification step is the attack on the phosphate at 3'-end of the intron by the 3'-hydroxyl of the detached exon: this step results in the ligation of two exons via a phosphodiester bond and the release of the intron, still in the form of a lariat. The lariat is then debranched to give a linear excised intron, that is rapidly degraded [24].

1.2.2 Splicing machinery: the spliceosome

The spliceosome is a highly dynamic macromolecular complex that is assembled on the nascent mRNA transcript, recognizes the exon-intron junctions, catalyzes the intron removal and joins the exons and finally is released after the reaction has been completed [26]. The core of the spliceosome consists of five uridine-rich (U-rich) small nuclear ribonucleoproteins (snRNPs) U1, U2, U4/U6, and U5 along with a large number (from 150 to 300) of non-snRNP associated proteins, such as heterogeneous nuclear RNPs (hnRNPs) and serine-arginine rich (SR) proteins.

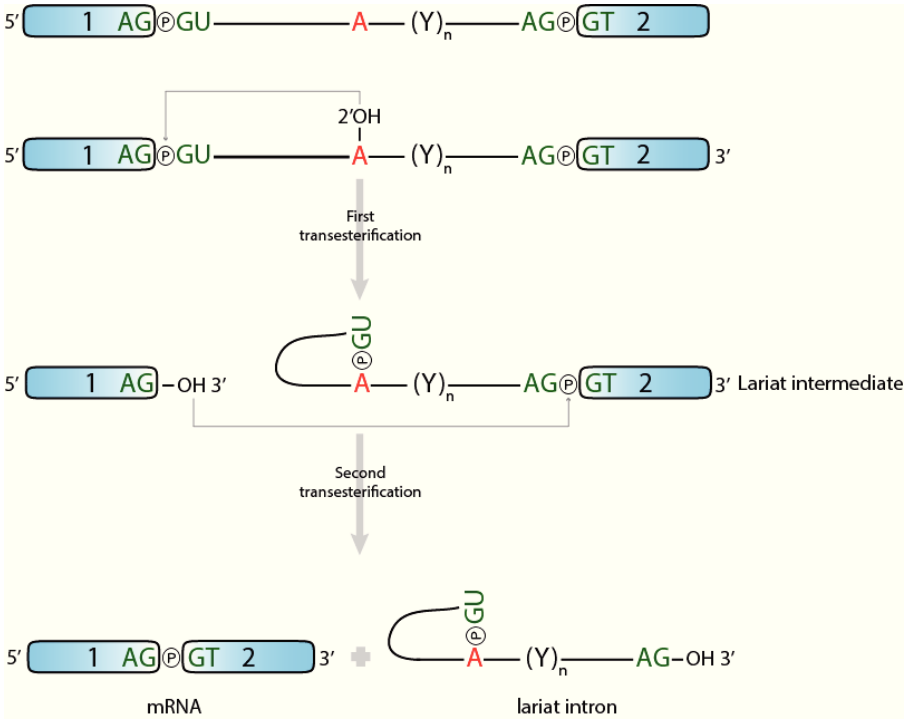


Figure 1.2: **Splicing occurs in two transesterification reactions**
Schematic representation of the pre-mRNA with essential signals for splicing reaction. Exons are light blue boxes, intron is a black line. The GU and AG dinucleotides at the 5' and 3'ss respectively are indicated in green, the red A residue of the branch point (BP) and the polypyrimidine tract $(Y)_n$ are indicated. In the first transesterification step, the A residue of the BP attacks the phosphodiester bond at the 5'ss of exon 1 generating two splicing intermediates: free exon 1 and lariat-exon2 (lariat intermediate). During the second transesterification reaction, the 3'-OH group of the detached exon attacks the phosphate at the AG 3'-end of the intron. Final products of splicing reaction are two exons joined together through phosphodiester bond (mRNA) and the intron in a lariat form. Phosphates are shown as circled P's [25].

Each snRNP consists of a small stable RNA bound by a specific set of seven proteins (Smith (Sm) proteins), plus numerous other less stably associated splicing factors [27]: the number of these associated proteins vary greatly among the U snRNPs, from 3 only for U1 snRNP to more than 20 for U2 snRNP. The snRNPs have a critical role in the recognition of correct splice sites within a multitude of similar sequences. The production of a spliced, mature mRNA requires extensive specific and dynamic interactions of different nature, such as RNA-RNA base pairing, RNA-protein and protein-protein binding and a lot of structural changes [28, 29].

Some metazoan species and plants contain a second, minor spliceosome responsible for the excision of a rare class of introns, that is composed of structurally distinct but functionally analogous U11/U12 and U4atac/U6atac snRNPs, with the U5 snRNP which is common for the

two spliceosomal machineries [30].

The spliceosome assembly occurs in a stepwise manner, involving assembly/disassembly of different snRNPs and non-snRNP splicing factors on the pre-mRNA (fig. 1.3). The assembly of the spliceosome (early (E) complex or commitment complex) starts with the recognition of the 5'ss by the U1 snRNP through its 5'-end, in a ATP-independent manner. U1 snRNP associated proteins (U1-70k and U1-C) stabilize this transient interaction [31]. The other important event that occurs during the E complex formation is the recognition of the 3'ss: the U2 auxiliary factor (U2AF) recognizes with the U2AF65 subunit the polypyrimidine tract [32], the AG dinucleotide at the 3'ss interacts with the U2AF35 subunit [33], whereas the branch point binding protein (BBP) binds to the branch point [34]. Subsequent to E complex formation, the A complex is built: U2 snRNP replace the BBP factor and binds to BP in an ATP-dependent manner. The base pairing interaction is further stabilized through SF3a and SF3b subunits of the U2 snRNP [35]. The transition from A to B complex are marked by the ATP-dependent addition of the U4/U6 and U5 snRNPs as a pre-assembled tri-snRNP. At this step all snRNPs are present, but the spliceosome is catalytically inactive and requires a conformational and compositional rearrangement to became active and promoting the first transesterification step of splicing. In order to activate the spliceosome the following events occur: the displacement of U1 at the 5'ss by U6, the disruption of the U4/U6 base pairing interaction and the formation of extensive interaction between U2 and U6 snRNPs [36]. These conformational changes provide the structural basis to juxtapose the branch site and the donor splice site, promoting the formation of the activated B complex (B*). The first step of the two splicing transesterification reactions occurs in the B* complex. This event is followed by the formation of the C complex in which the second transesterification reaction takes place. The U5, U2 and U6 snRNPs promote alignment of the exons for the second catalytic reaction, that results in a post-spliceosomal complex that contains the lariat intron and spliced exons. Finally, the snRNPs are released and recycled for additional rounds of splicing. The release of the spliced product from the spliceosome is catalyzed by the DExD/H helicase Prp22 [37, 38].

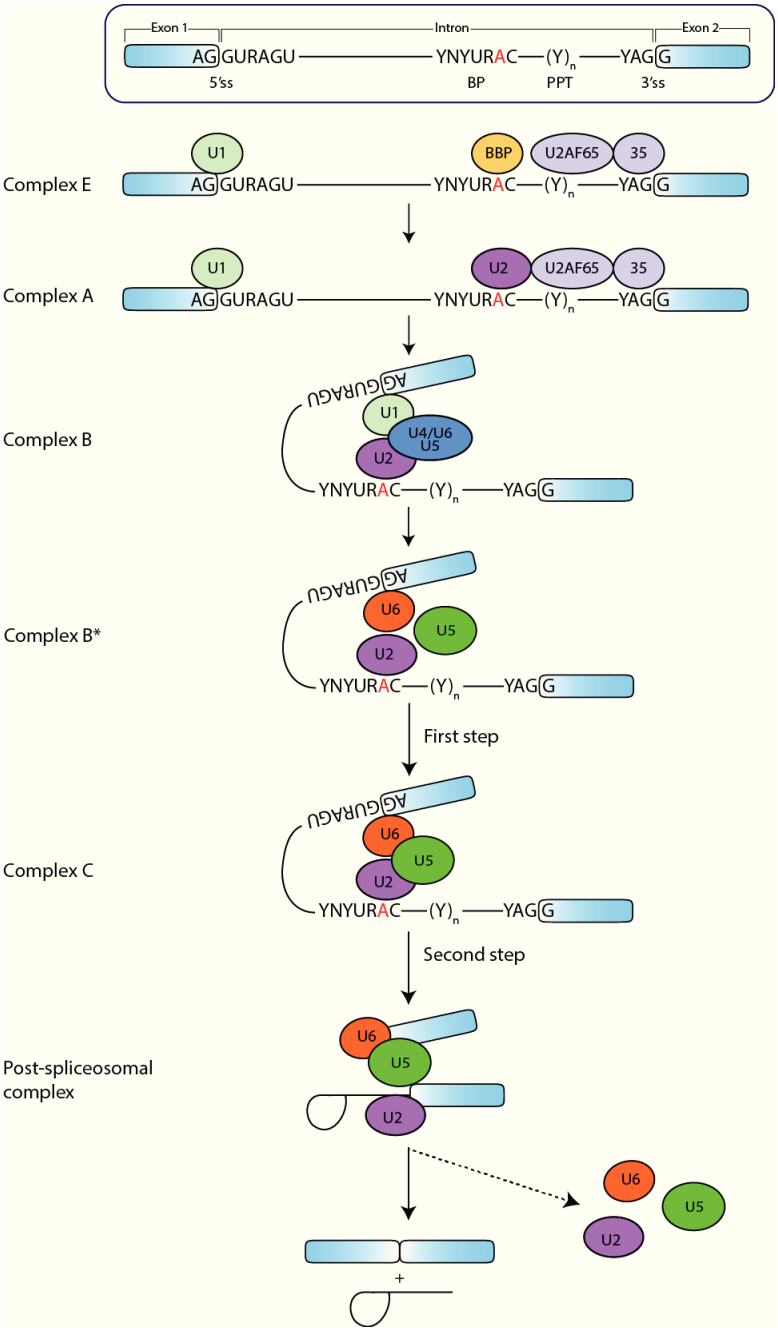


Figure 1.3: **Spliceosome assembly**

Schematic representation of pre-mRNA splicing and the spliceosome assembly pathway. Exons are light blue boxes and intron is a black line; consensus nucleotides are indicated above the line and the A of the BP is in red. The complex E contains U1 snRNP bound to the 5'ss, BBP bound to the BP, and U2AF65 and U2AF35 bound to the PPT and 3'ss, respectively. In the complex A, BBP is replaced by U2 snRNP at the BP. The U4/U6.U5 tri-snRNP is recruited to form the B complex. After a series of rearrangements the complex B is catalytically active (B* complex) and carries out the first catalytic step of splicing, gnerating complex C, which undergoes additional rearrangements to carried out the second catalytic step, resulting in a post-spliceosomal complex that contains the lariat intron and spliced exons. The U2, U5 and U6 snRNPs are released and recycled (dashed arrow) and the spliced product is released from the spliceosome by the Prp22 helicase.

1.2.3 U2 small nuclear RNP (U2 snRNP)

In humans the U2 snRNP consists of a 186nt long RNA molecule, that forms five stem-loops, associated with several proteins (fig. 1.4): seven Sm proteins, common to all the U snRNPs (Sm B, D3, G, E, F, D2 and D1), two U2-specific proteins (U2-A' and U2-B") and two heteromeric splicing factors SF3a and SF3b, which consist of three (SF3a120, SF3a66, SF3a60) and seven proteins (SF3b155 (SF3b1), SF3b145, SF3b130, SF3b49, SF3b14a, SF3b14b (p14), SF3b10), respectively and are named according to their molecular weight [39].

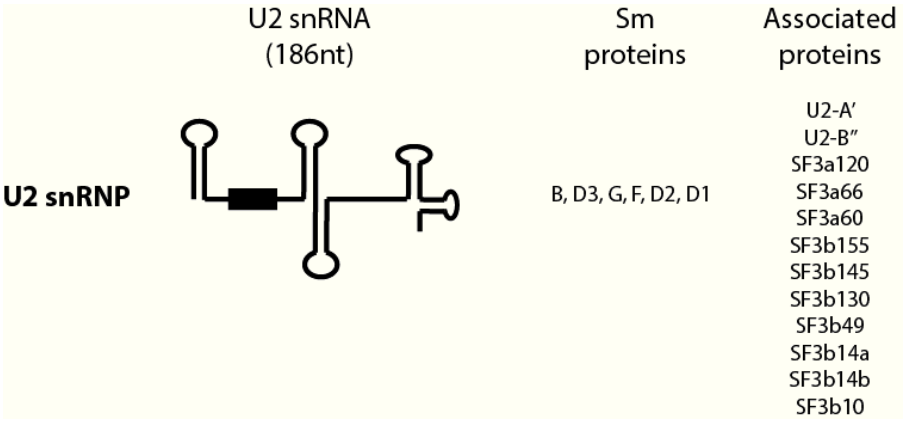


Figure 1.4: **Composition of U2 snRNP**

Schematic representation of the U2 snRNP composed by the U2 snRNA and several associated proteins. snRNA secondary structure formed by five stem-loop and the branch point recognition sequence (BPRS), represented as a black box, are shown in the first column. The nucleotide length of the RNA molecule is indicated. The second and third columns report the associated Sm proteins and other associated proteins.

The U2 snRNP plays an essential role in pre-mRNA splicing. In the early stage of spliceosome assembly, the BP sequence is recognized by the BBP factor, while the U265 and U2AF35 recognize the PPT and the 3'AG, respectively. Subsequently, in an ATP-dependent manner, the U2 snRNP replaces the BBP/SF1 factor, inducing the A complex formation (fig. 1.5). Stable binding of U2 snRNP to the BP requires an initial base-pairing interaction between the BP site and a region of U2 snRNA known as the branch point recognition sequence (BPRS) [40]. This base-pairing involves the entire BP sequence, except for the A residue, that is forces to flip out and be exposed. To further stabilize this interaction the SF3b1 subunit interacts with BP both upstream and downstream sequences [41], highlighting the role for this factor in tethering U2 snRNP in the proximity of the BP, while the BP adenosine, which is bulged out from the base-paired region, is directly bound by the p14 factor [42].

The central function for SF3b1 in U2 snRNP recruitment and splicing catalysis is also suggested by its interaction with U2AF65. In addition, U2 snRNP binding serves as a platform for additional RNA–RNA and protein interactions. This will lead to recruitment of the U4/U6.U5 tri-snRNP and formation of mature spliceosome complexes, within which numerous RNA rearrangements and changes in protein composition facilitate splicing catalysis.

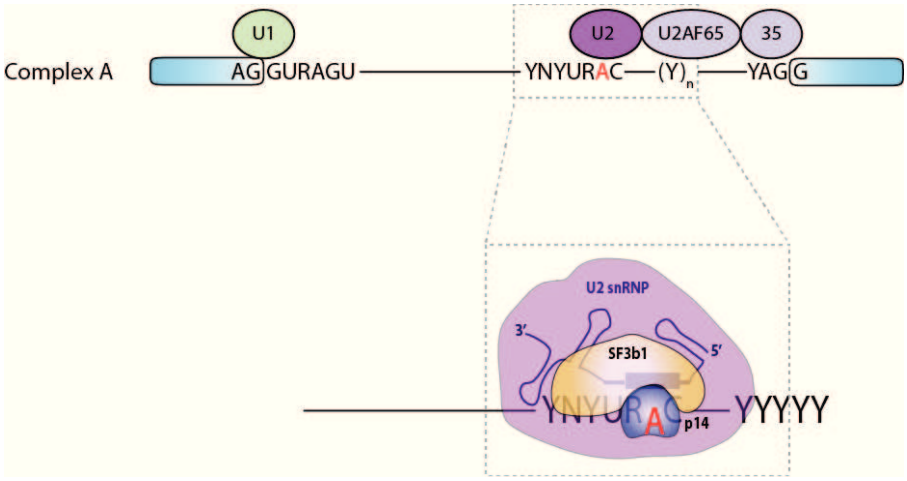


Figure 1.5: **U2 snRNP assembly on branch point**
U2 snRNP binds BP, inducing complex A formation. U2 snRNP assembly involves base-pairing interactions between the BPRS (black box) in U2 snRNA and nucleotides flanking the BP adenosine, with the A (red) bulged out from the base-paired region. SF3b1 (orange) interacts with both 5' and 3' of the BP, while p14 (blue) contacts the BP adenosine.

Somatic mutation of SF3b1 factor were observed in patients with different hematological malignancies, such as myelodysplastic syndromes [43,44], uveal melanoma [45,46], skin melanoma, breast, endometrial and bladder cancers, pancreatic ductal adenocarcinoma and adenoid cystic carcinomas [47]. SF3b1 mutations have also ben reported, with high mutation rates, among chronic lymphocytic leukemia (CLL) patients and are predictive of a poor outcome, with earlier need for treatment and shorter survival [48, 49]. Almost all SF3b1 mutations localize to its C-terminal, which comprises 22 HEAT repeats. In particular, the mutations have been detected between the fifth to the eighth HEAT repeats (encoded by exons 14-16). K700, G742, K666, E622, R625 and H662 are the most frequently mutated sites [50]. In line with the role of SF3b1 in splicing, mutations of this factor result in aberrant splicing of a sub-population of exons, but, despite these evidences, the contribution of these mutation to cancer development in non completely understood.

1.2.4 Recognition of the exon: the canonical *cis*-acting elements

Several motifs in the nucleotide sequences near the exon/intron boundaries are required for proper exon definition (fig. 1.6): these elements are short consensus sequences surrounding the 3' and 5'-end of the introns, which are known as 3'ss and 5'ss, respectively [51].

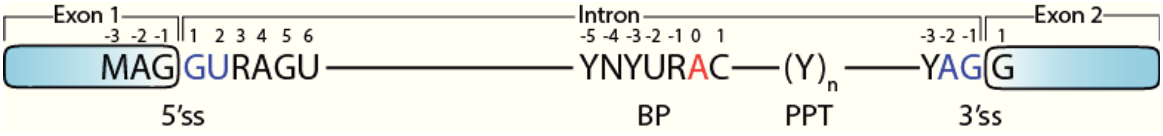


Figure 1.6: **Overview of canonical *cis*-acting splicing signals**
Consensus GU at the 5'ss (blue), the A residue of the BP (red), the PPT $(Y)_n$ preceding the AG at the 3'ss (blue) are shown in a two-exon pre-mRNA. The sequence motifs that surround these conserved nucleotides are shown. Numbers indicate the position of the nucleotides in the consensus sequence.

1.2.4.1 5' splice site (5'ss) or donor site

The 5'ss marks the exon/intron junction at the 5'-end of the intron. Its consensus sequence consists of 9 partially conserved nucleotides, MAG|GURAGU (M indicates A or C, R indicates purines A or G and the | the exon/intron boundary), located on both sides of the exon/intron boundary: the sequence, indeed, spans from position -3 to +6. The underlined GU dinucleotide is almost universally conserved as it is found in more of 98% of human donor splice site and mutations in this dinucleotide lead to splicing defects [52].
A minority of 5'ss (<1%) has a GC dinucleotide at the exon/intron boundary. Nevertheless the entire consensus donor splice site determines the 5' cleavage site, not only the invariant GU dinucleotide [53].

1.2.4.2 3' splice site (3'ss)

The intronic element that identifies the 3'ss usually appears several thousand bases downstream the 5'ss. It is composed of three different elements: the branch point (BP), the polypyrimidine tract (PPT) and the terminal AG dinucleotide (or acceptor site) [54].

The **branch point** is characterized by the presence of a conserved A (underlined), surrounded by a highly degenerated motif YNYURAC (Y indicates pyrimidines, R indicates purine and

N any nucleotide) and is the site involved in first transesterification reaction. It is commonly found 18-40 nucleotides upstream of the acceptor site [55].

The **polypyrimidine tract** is composed of a stretch of pyrimidines, in particular uridines, and is located between the BP and the terminal AG at the intron/exon junction. The PPT is essential for promoting splicing reaction: it has been shown that progressive deletion of the PPT impairs splicing, while elongating its lenght can improve its efficiency [56].

The **AG dinucleotide** is located just downstream the PPT and defines the 3' border of the intron. This site is characterized by the short Y<u>AG</u>|G sequence (Y is a pyrimidine and | is the intron/exon boundary), where the AG (underlined) is highly conserved and it is fundamental for the second transesterification step of splicing reaction.

1.2.5 Auxiliary *cis*-acting elements

The degenerate nature of the 5' and 3'ss allows for the existence of a wide range of natural splice sites within intron sequences: these pseudo splice sites largely outnumber real splice sites and, in many cases, their strength can surpass that of correct splice sites [57]. They match the consensus but are not selected during the process. Thus, the splicing machinery is able to recognize the real splice sites, although they are weak, distant from each other and surrounded by several pseudo 5' and 3'ss: this suggest that the consensus elements are not sufficient to define exon/intron junctions and recognition of correct splice sites is the result of a combinatorial regulatory mechanism that involved other *cis*-acting elements [23].
These additional elements are located both in introns and in exons (fig. 1.7) and, depending on their position and function, are identified as exonic or intronic splicing enhancers (ESEs, ISEs) and silencers (ESSs, ISSs) [23]. Even if they are conserved among species, their sequences are highly degenerated and their functions may overlap [58].

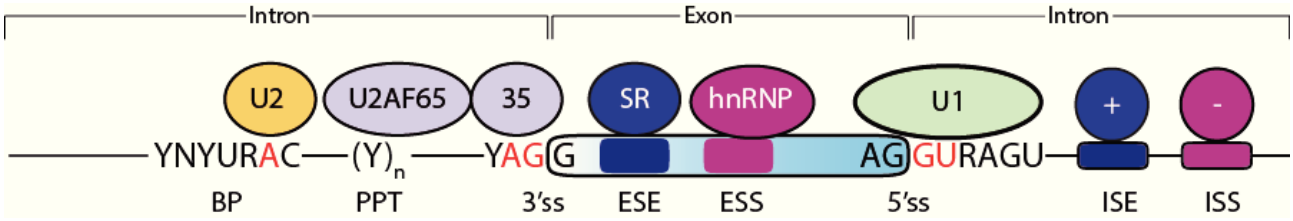


Figure 1.7: **Auxiliary *cis*-acting elements in pre-mRNA splicing**
Schematic representation of possible distribution of canonical and additional splicing *cis*-acting elements, along with *trans*-acting factors. The canonical splicing signals that define exon boundaries are relatively short and contain poorly conserved sequences. Only the GU at the 5'-end and the AG at the 3'-end together with the BP adenosine, all highlighted in red, are strongly required and invariant. A PPT of variable length $(Y)_n$ is reported upstream of the 3'ss. The basal components of the splicing machinery bind to the consensus sequences and promote assembly of the splicing complex. The U1 snRNP binds to the 5'ss, the U2 snRNP binds the BP and the U2AF, through its two components U2AF65 and U2AF35, interacts with and recognizes the PPT and the AG at the 3'ss, respectively. Auxiliary *cis*-acting enhancer and silencer elements in the exons (ESE, ESS) and/or introns (ISE, ISS) allow the correct splice site to be distinguished from the many pseudo splice sites. *Trans*-acting splicing factors interact with the regulatory sequences and, accordingly to their general functions, are divided into SR proteins family with enhancer activity and hnRNP protein family with inhibitory activity.

In addition, it has been reported an overlapping functions of enhancer and silencer for some regulatory sequences, as in the case of composite exonic regulatory elements of splicing (CERES) identified in the cystic fibrosis transmembrane conductance regulator gene (CFTR) exons 9 and 12 [58]. Regulatory elements, if present in multiple copies, can operate in additive manner: they can increase the affinity of the associated *trans*-acting factors [59,60]. Different regulatory elements may cooperate to promote the recognition of an exon [61]. This fine-tuned balance of splicing factors is of extreme importance for the alternative splicing mechanism that, in turn, contributes in a significant way to gene expression regulation [62].

1.2.5.1 Splicing enhancers

Splicing enhancers are regulatory elements that reside in both exonic and intronic regions and stimulate exon inclusion.

Exonic splicing enhancers (ESEs) are short (6-8 nucleotides), degenerate and partially overlapping sequences, that are typically recognized by members of serine/arginine-rich (SR) protein family [63]: they act by directly recruiting the splicing factors and/or by antagonizing

the action of nearby silencer elements [23]. These two models of splicing enhancement are not necessarily mutually exclusive, as they might reflect different requirements in the context of different exons.

There are more than twenty different SR proteins and SR-like proteins in humans. SR proteins have an RNA recognition motif (RRM) at their N-terminus, which allows them to bind to specific RNA sequences, and a C-terminal domain that is highly enriched in arginine/serine dipeptides (RS domain) that is involved in protein-protein interactions.

Several functional systematic evolution of ligands by exponential enrichment (SELEX) experiments done *in vitro* [64, 65] or *in vivo* [66] have shown the existence of several types of ESEs: these motifs have been classified as purine-rich enhancers, AC-rich enhancers or pyrimidine-rich enhancers. In addition, many different ESEs with unique sequences have been also described. Using the sequences that resulted from the functional selection procedure, Cartegni and collaborators [67] calculated the frequencies of individual nucleotides at each position and these data were used to create matrices in order to predict the location of putative SR protein-specific ESEs. The implementation of the motif-scoring matrices was inserted in a web-based program called ESEfinder (release 3.0; http://rulai.cshl.edu/tools/ESE) which allows scanning of nucleotide sequences to predict putative ESEs responsive to the human SR proteins SRSF1 (SF2/ASF), SRSF2 (SC35), SRSF5 (SRp40) or SRSF6 (SRp55). How the SR proteins bound to ESE facilitates exon definition and correct selection of splice sites is described below in the paragraph that consider these *trans*-acting factors.

Fewer large-scale studies [68] have been conducted for **intronic splicing enhancers (ISEs)** and many more intronic elements are expected to be identified in future studies. Most of the ISEs have been identified by the study of disease-causing point mutations: some of them occur 20-40bp downstream of the 5'ss, causing exon skipping [69]. These sequences can be fundamental for constitutive exon recognition or can be regulated in a tissue or developmental specific manner [70].

1.2.5.2 Splicing silencers

Splicing silencers are regulatory elements located in introns (**intronic splicing silencers, ISSs**) or in exons (**exonic splicing silencers, ESSs**), that are responsible for splicing inhi-

bition. About one-third of randomly selected short, human DNA fragments showed splicing inhibitory activity *in vivo*, when inserted into the middle exon of a three-exon minigene [71]: these data suggested that silencers are very frequent, even though only limited number has been characterized by mutational analysis.

Splicing silencers can be purine or pyrimidine-rich and bind a diverse array of proteins. Silencers usually work by interacting with negative regulators, which often belong to the heterogeneous nuclear ribonucleoprotein (hnRNP) family, a class of diverse RNA-binding proteins that associate with nascent pre-mRNAs [72,73]. Similar to SR proteins, hnRNP proteins have a modular structure, which consists of one or more RNA-binding domains associated with an auxiliary domain that is often involved in protein-protein interactions.

Not only hnRNPs have been implied in the silencing but also other factors such as the PPT-binding protein (PTB) [74, 75] or, in peculiar cases, also the SR proteins [76]. Silencers can regulate splicing in different ways: by antagonizing the function of a nearby ESE or by recruiting factors that interfere with the splicing machinery by direct binding, exon looping or by nucleation and cooperative binding [23].

1.2.5.3 RNA secondary structure

RNA molecules have a natural tendency to form highly stable secondary structures: these structures might act as *cis*-acting elements, influencing splicing process [77].

RNA secondary structure may hinder the accessibility of splicing factors to functional sequences, either splice sites or other regulatory sequences, present in the pre-mRNA by sequestering them in stems or looping them out [78].

RNA secondary structure can also affect the relative distance between splicing elements determining a considerable variation in the splice site recognition [77] or, for example, splicing process can be stimulate when silencer elements are trapped within a RNA structure.

1.2.6 *Trans*-acting factors

Proteins involved in the splicing reaction can be divided in two major groups: the snRNPs and the non-snRNPs. These splicing factors, independently of their functional characteristics, share some similar structural features such as the RRM and/or protein binding domains and usually target short sequence elements adjacent to site of regulation.

As described above, the **U1, U2, U4, U5 and U6 snRNPs** are the main components of the spliceosome and catalyze the splicing reaction [79]. Each snRNP particle is composed of U-rich snRNA molecule, seven Sm or Sm-like proteins and several splicing factors [80]. The snRNAs are characterized by their small size, stability and show high level of sequence conservation [81]. RNA Pol II transcribes all U snRNAs with the exception of U6 snRNA that is transcribed by RNA Pol III [80]. Upon transcription, the assembly of U6 snRNP probably takes place in the nucleus, whereas the other snRNAs are transported to the cytoplasm where snRNP assembly initiates [82]. Following their export to the cytoplasm, the snRNA precursors bind seven Sm proteins in order to form the structural core of snRNPs. A properly assembled Sm core, together with cap hypermethylation and 3'-end processing of the snRNA, are required for nuclear import [82]. Apart from the Sm and Sm-like proteins, that are required for U6 snRNP assembly, snRNPs also contain a certain number of other particle-specific proteins.

For instance, the mammalian U1 snRNP has three proteins that are specific for this particle: U1-70K and U1-A interact directly with the RNA and are important for splice site selection, U1-C establishes interaction with U1-70K and Sm proteins, thus contributing to the stabilization of the U1 snRNP particles [83].

The **non-snRNPs** proteins are involved in the regulation of both general and tissue-specific splicing events. In particular, two families of RNA binding proteins, SR proteins and hnRNPs, have been found as components of distinct regulatory complexes with functional specificity in splicing [84].

The **SR proteins** are a family of highly conserved nuclear phosphoproteins that play multiple roles in splicing and in general cell metabolism [85]. All SR proteins have a modular structure and contain one or two N-terminal RRMs that interacts with the pre-mRNA and a C-terminal RS domain rich in arginines and serines [86], that mediates protein-protein interaction with members of spliceosome [87]. Two non-exclusive models have been proposed to explain the role of SR proteins in pre-mRNA splicing: SR proteins that are bound to ESEs act by directly recruiting the splicing machinery through their RS domain (RS-domain dependent mechanism) and/or by antagonizing the action of nearby silencer elements (RS-domain independent mechanism) [63].

Sequence specificity is essential for SR protein binding, but the consensus sequence is highly degenerated [88]: for this reason SR proteins can bind the target sequence of some other family member, thus contributing to functional redundancy [89], even if is not a rule, because some members cannot be replaced [90]. The activity of SR proteins is regulated through phosphorylation/dephosphorylation cycles, that influence the sub-cellular localization of the protein and its activity [18]. The RS domain phosphorylation is determinant for the localization of SR proteins, that is preferentially nuclear, but can rapidly shuttle between the nucleus and the cytoplasm [91]. Moreover phosphorylation increases the RNA-binding specificity of SR proteins [92].

The **hnRNP protein** family is a class of several RNA-binding proteins that associate with nascent pre-mRNA until its processing is completed. Although these factors are predominantly localized in the nucleus, a subset of these proteins shuttles continuously between nucleus and cytoplasm: this indicates a role of these proteins in mRNA export from nucleus to cytoplasm and in other cytoplasm processes [93,94]. This family includes more than twenty members, each of theme with multiple splicing isoforms and different post-translational modifications, such as phosphorylation, arginine methylation and SUMOylation [72]. These proteins are expressed in all tissues but the relative amount of different hnRNPs vary among cell types and show stage-specific expression pattern: some hnRNPs are extremely abundant (about 100 million copies per nucleus), while others are present in a lower amount [95].
The structure of hnRNPs contains one or more RRMs associated with auxiliary domains that have been shown mediate protein-protein interactions [72]. The hnRNP proteins usually mediate splicing inhibition, particularly through the interaction with ESS elements or by sterical interference with other splicing factors [23], antagonizing directly or indirectly SR proteins. Nevertheless, depending on the position of the splicing regulatory elements hnRNPs can also associate with enhancer elements to help exon inclusion [96], indicating that hnRNPs can have various roles in pre-mRNA splicing.
However, assemblies of tissue-specific factors and proteins of the SR and hnRNP families can have positive or negative roles depending on their precise location, composition and state of modification of their components [97].

1.2.7 Alternative splicing

The proteome diversity is guaranteed by several cellular events that allow the generation of different polypeptides from a single gene. These mechanisms include the usage of multiple transcription start sites, alternative pre-mRNA splicing, polyadenylation, RNA editing and post-translational modifications. Alternative splicing (AS) splicing is considered one of the most important source of proteome diversity. It has been described in nearly all metazoans and almost all genes are subjected to at least one alternative splicing event [98, 99]. This process generates a large number of mRNAs that encode proteins with different or even opposite biological functions with crucial consequences on cells metabolism [100]. Non-coding transcripts also show alternative splicing patterns [101]. AS is a highly regulated process that plays a key role many cellular processes, such as sex determination, cell differentiation, cell transformation or apoptosis [62, 102]. Many different alternative splicing events are present in nature (fig. 1.8). A single cassette exon can reside between two different constitutively spliced exons and can be either spliced out from the mature mRNA (exon skipping) or retained (exon inclusion). Alternative 3' and 5'ss can be used and therefore exon length and sequence can vary. Moreover, multiple cassette exons can be located between the two constitutive exons and the splicing machinery chooses between them (mutually exclusive exons). In addition, intronic sequences can be included in the mature mRNA and, in turn, translated into proteins whose functionality depends on the maintenance of the correct protein reading frame (intron retention). Different 5' starting point can be used by the selection of different promoters (alternative promoters), as well as alternative polyadenylation signals (alternative poly(A) sites) can regulate the usage of the 3'-terminal exon. All these patterns can be combined in a single transcript to produce a complex array of splice isoforms [62, 97].

The mechanisms that promote which splice site will be used and/or which exon will be included in mature mRNA in different cell types or developmental stages have been studied in recent years, although it is clear that further studies are required to fully understand the molecular mechanisms at the basis of this phenomena. Most alternatively spliced exons are thought to be controlled by multiple auxiliary *cis*-acting elements and *trans*-acting factors in a combinatorial manner [62]. Moreover, in the last years it has became more clear that splicing decisions are tightly coupled to epigenetic factors, such as RNA Pol II elongation rate, nucleosome

positioning and chromatin remodeling. All these events can have an impact on exon definition and fate [103–105]. Thus, the emerging picture of AS events underlines its extreme complexity and its precise regulation, what makes it a central event in the regulation of gene expression and cell fate.

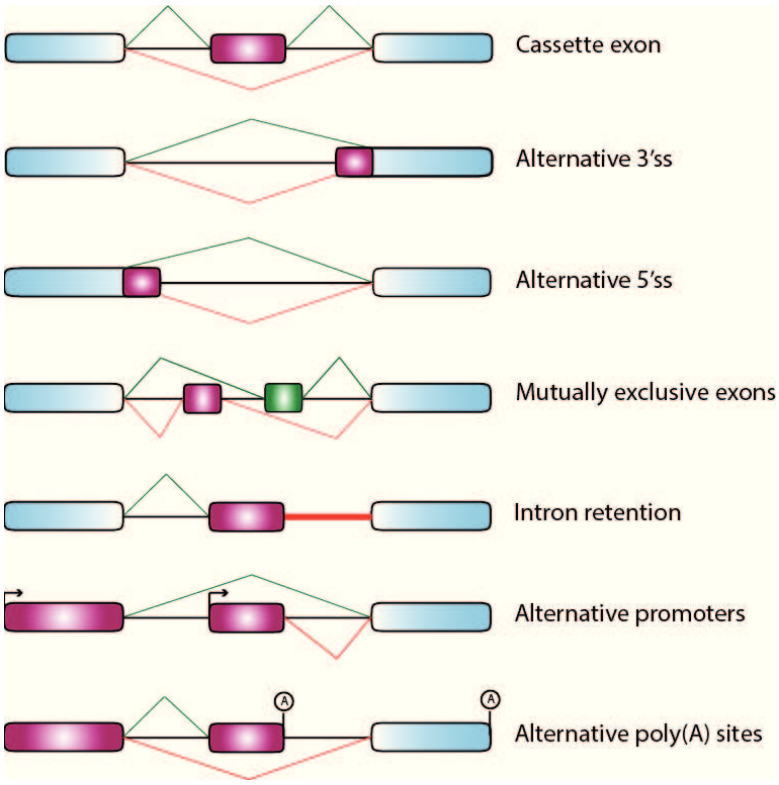


Figure 1.8: **Schematic representation of alternative splicing patterns.**
The light blue boxes represent constitutive exons, the purple ones alternative exons and introns are black lines. The lines above and below the boxes indicate the normal (green) and the alternative splicing events (red). Arrows indicate promoters, while circled A's alternative poly(A) sites. AS of internal exons includes the cassette exon, alternative 3'ss, alternative 5'ss, mutually exclusive exons and intron retention. The usage of alternative promoters leads to the selection of one of the multiple first exons, while the usage of alternative terminal exons generates mature transcripts with different poly(A) sites.

1.3 miRNA biogenesis, localization, function and regulation

microRNAs (miRNAs) are a family of ∼22 nt non-coding, small, single-stranded RNAs (ssRNAs) generated from endogenous transcripts that contain a local hairpin structure [106]. miRNAs comprise one of the more abundant classes of gene regulatory molecules in multicellular organisms and likely influence the output of many protein-coding genes.

28645 precursor miRNAs, expressing 35828 mature miRNA sequences, found in 223 different species, have been reported in the miRBase database, including 1881 precursor and 2588 mature entries in human (http://mirbase.org; Release 21: June 2014), although the functional importance of many of these miRNA annotations remains to be determined [107]. With the number of identified miRNA increasing rapidly, rules of annotation have been suggested to designate individual miRNAs [106], such as hsa-miR-34, with "hsa" standing for the species, homo sapiens, "miR" for microRNA, and the number indicating the order of being discovered. miRNAs with similar sequence add a letter after the number (miR-34a, miR-34b, miR-34c), while miRNAs with identical mature sequence, coming from different precursors add another number after the letter (miR-92a-1, miR-92a-2). Two mature miRNAs originate from the same precursor: miRNA-3p and miRNA-5p arise from the 3' arm and the 5' arm, respectively.
They negatively regulate gene expression at the post-transcriptional level: miRNAs function as guide molecules in post-transcriptional gene silencing by base pairing with target mRNAs, which leads to mRNA degradation or translational repression. However, recent findings indicate that miRNA-mediated repression can be reversed, prevented or even act as translational activators [108] and also an epigenetic mechanism of miRNA-directed transcriptional gene silencing, through the heterochromatin formation, has been described [109].
Several studies have revealed that miRNAs have key roles in various cellular activities, such as control of developmental timing, cell differentiation, cell proliferation, cell fate determination and apoptosis. Additionally, miRNAs have been shown to participate in multiple biological processes including cardiogenesis, skeletal muscle proliferation and differentiation, brain morphogenesis, hematopoietic lineage differentiation and tumorigenesis [110].
Improvements in the characterization of miRNAs revealed their roles not only in various cellular processes, but also abnormal pattern of miRNA expression in various diseases.

1.3.1 Discovery of miRNAs

miRNAs have been first described in 1993: the groups of Ambros and Ruvkun [111, 112], that conducted studies in *C. elegans*, discovered that lin-4 did not encode for a protein but, instead, produced short RNA transcripts each regulating the timing of larval development by translational repression of lin-14, which encodes a nuclear protein. They postulated that this regulation was due in part to sequence complementarity between lin-4 and unique repeats within the 3' untranslated regions (3'UTR) of the lin-14 mRNA, downregulation of which at the end of the first larval stage initiates developmental progression into the second larval stage.
The discovery of lin-4 and its target-specific translational inhibition hinted at a new mechanism of gene regulation during development. In 2000, almost seven years after the initial identification of lin-4, the second miRNA, let-7, was discovered, in worms. Similar to lin-4, let-7 performs its function by binding to the 3'UTR of lin-41 and hbl-1 (lin-57), and inhibiting their translation [113]. The identification of let-7 not only provided another vivid example of developmental regulation by small RNAs, but also raised the possibility that such RNAs might be present in species other than nematodes. Since the discovery of let-7, several miRNAs have been identified in other organisms. The researchers shown that let-7 is evolutionarily conserved throughout metazoans, with homologues that were detected in flies, mice and humans [114]. Almost all human miRNAs are conserved in mouse [115] and one-third of *C. elegans* miRNAs have vertebrate homologs: this conservation indicates a more general role of small RNAs in developmental regulation [116].

1.3.2 Canonical miRNA biogenesis

1.3.2.1 miRNA genes transcription

The transcription of most miRNA genes is mediated by RNA Pol II (fig. 1.12): these non-coding primary transcripts, referred as pri-miRNAs, have a 5' m^7G cap, and a 3' poly(A) tail, similar to mRNAs, and are usually several kilobases long and contain a local hairpin structure [117, 118]. However, miRNA that are part of a repetitive elements such as Alus, are usually transcribed by RNA Pol III [119].
A typical metazoan pri-miRNA has a stem-loop secondary structure of ∼80 bp, with a stem of ∼33 bp, a terminal loop and flanking ssRNA segments.

1.3.2.2 Nuclear processing by Microprocessor complex

Human Microprocessor complex (MPC) is a heterotrimeric complex of ∼364 kDa composed of one Drosha and two DGCR8 molecules (fig. 1.9 - left panel) [120]. MPC is the machinery involved in the first step of miRNAs maturation: it recognizes and crops the pri-miRNA hairpin into precursor miRNA (pre-miRNA). **Drosha** is a 159 kDa nuclear type III RNase protein, it contains a proline-rich (P-rich) and a RS-rich N-terminal domains, a central domain (CED), two RNase III domains (RIIIDa and RIIIDb) and a double-stranded RNA binding domain (dsRBD) at the C-terminus. Its cofactor **DGCR8** (DiGeorge syndrome critical region 8) is a 86 kDa protein, which contains a nuclear localization signal (NLS) located in the N-terminal region, a central RNA-binding domain (Rhed), two dsRBDs (dsRBD1 and dsRBD2) and the C-terminal tail region (CTT) (fig. 1.9 - right panel) [120]. Drosha recognizes the ssRNA-dsRNA junction (basal junction) and interacts with the UG motif here located, while the DGCR8 dimer contacts the UGU motif at the apical junction. Therefore, the UG and UGU motifs are fundamental to assure the correct orientation of the complex. Drosha functions as a ruler, cutting 11 bp from the basal juction to generate a stem-loop precursor of ∼70 nt with a 2 nt 3'-overhang (fig. 1.12) [121]. DGCR8 binds and stabilizes Drosha, enhancing its processing efficiency [120].

1.3.2.3 Nuclear export by Exportin-5

The pre-miRNA is exported from the nucleus into the cytoplasm by **Exportin-5 (Exp5)** (fig. 1.12), a member of the nuclear transport receptor family: in particular, Exp5 is a RanGTP-dependent nuclear export receptor that recognizes a dsRNA stem of >16 bp and a short 3'-overhang [122]. It binds cooperatively to the pre-miRNA and the GTP-bound form of the cofactor Ran in the nucleus, and releases its cargo in the cytoplasm, following the hydrolysis of GTP to GDP [123]. Exp5 has been first characterized as a minor export transporter for tRNAs [124], because it can transport tRNAs when the primary export factor is depleted or overloaded, but the major substrate for its activity are pre-miRNAs [123].

When the cells are depleted of Exp5, the pre-miRNA level and the mature miRNA level are reduced in the cytoplasm [122]. Notably, pre-miRNA does not accumulate in the nucleus subsequent to the depletion of Exp5. This indicates that pre-miRNA might be relatively unstable and also that pre-miRNA might be stabilized through its interaction with Exp5.

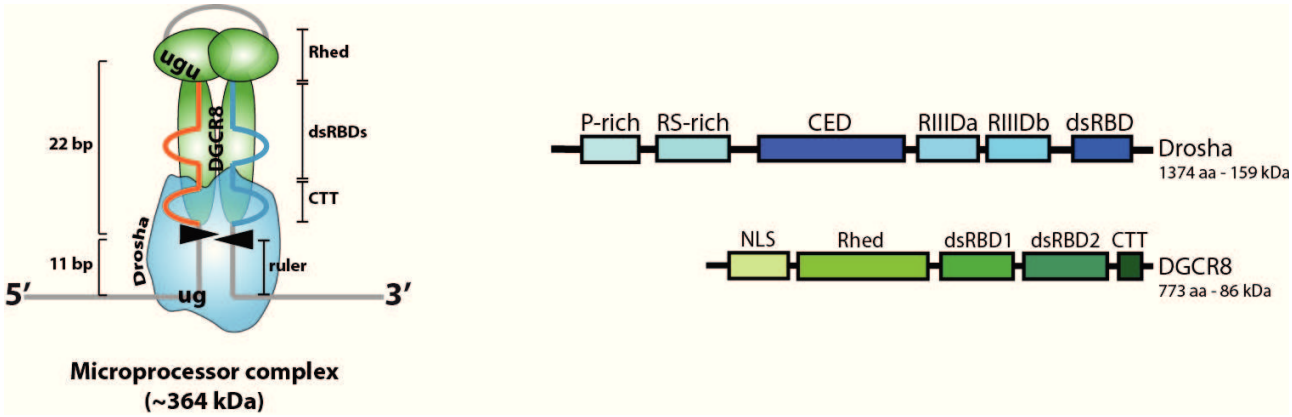


Figure 1.9: **Human Microprocessor complex: Drosha and DGCR8 domain organization**
[Left panel] Functional human Microprocessor complex on a pri-miRNA molecule (shown as a grey line, with the mature miRNA and miRNA* highlighted in orange and blue, respectively). Drosha (light blue) recognizes the ssRNA-dsRNA junction, interacts with the UG motif and functions as ruler, measuring 11 nt from the basal junction. DGCR8 dimer (green) binds and stabilizes Drosha with CTT, interacts with stem via its dsRBDs and recognizes the apical junction with the UGU motif through the Rhed domain. [Right panel] Drosha contains a P-rich and a RS-rich N-terminal domains, a central domain (CED), two RNase III domains (RIIIDa and RIIIDb) and a double-stranded RNA binding domain (dsRBD) at the C-terminus. DGCR8 contains a NLS in the N-terminal region, a central RNA-binding domain (Rhed), two dsRBDs (dsRBD1 and dsRBD2) and the C-terminal tail region (CTT).

1.3.2.4 Cytoplasmatic processing by Dicer

In the cytoplasm, pre-miRNA are cleaved near the terminal loop and processed into ∼22 nt miRNA duplexes by the cytoplasmatic RNase III Dicer [125] (fig. 1.12). Dicer's domain structure comprises a long N-terminal segment that contains a large helicase domain (which itself is composed of three predicted globular domains (HEL1, HEL2i and HEL2)), that contributes to pre-miRNA binding, followed by DUF283, a small domain of unknown function. Next there are a structure known as Platform and the PAZ domain, that binds to the 3' protruding end of the substrate [125]. The PAZ domain is separated in space from the catalytic core of the enzyme, composed of two tandem RIIIDs (RIIIDa and RIIIDb) and a dsRBD, located at C-terminal end. This region acts as a ruler for the enzyme to produce small RNAs of given size. (fig. 1.10) [126]. Like Drosha, Dicer interacts with two closely related proteins, the trans-activation response RNA-binding protein (TRBP) and the protein activator of PKR (PACT): these partners are not necessary for efficient cleavage, but they seem to contribute in the formation of the RNA-induced silencing complex (RISC), the effector complex for miRNA function [127].

The result of the sequential Drosha and Dicer cleavage is a short, imperfect miRNA duplex (miRNA:miRNA*) that contains both the mature miRNA strand (guide strand) and its complementary strand (passenger strand or miRNA*).

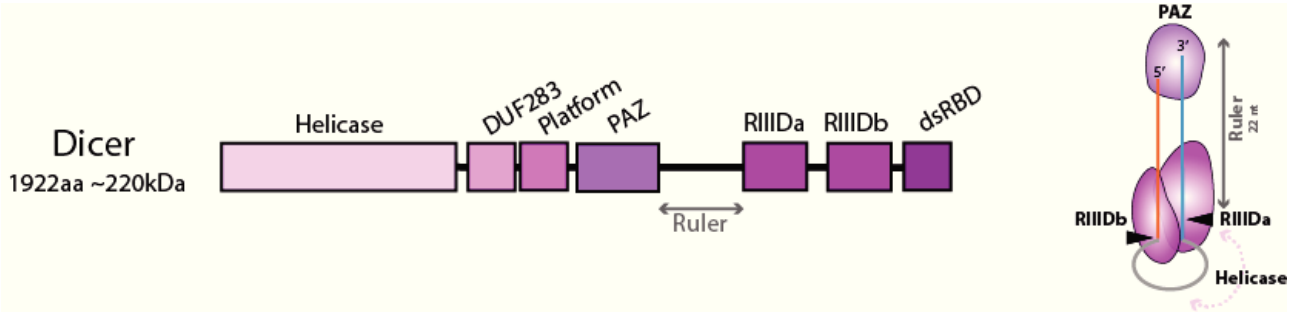


Figure 1.10: **Dicer domain organization and mechanism of action**
[Left panel] Dicer contains a large helicase domain, a small domain DUF283, followed by a Platform structure and the PAZ domain. Ruler (dark grey arrow) is followed by two tandem RIIIDs (RIIIDa and RIIIDb) and a dsRBD. [Right panel] pre-miRNA processing by Dicer: PAZ domain (light fuchsia) binds the 2 nt 3'-overhang. The separation distance (Ruler) between the PAZ and the RIIIDs domains correspond to the size of substrate produced. The helicase domain (light pink dotted arrow) is adjacent to the catalytic core.

1.3.2.5 RISC assembly and the AGO proteins

Strand selection is not a very precise process and mature miRNAs from both strands can be produced at the same frequency: the relative thermodynamic stability of the two ends of the duplex determines which strand is to be selected. The 5'-end that is less stable is usually kept as the miRNA [128, 129]. Mature miRNA is incorporated into effector complex (fig. 1.12), containing Argonaute proteins (AGO), that is known as miRNA-containing RNA-induced silencing complex (mi-RISC) [130]. Dicer, TRBP, PACT and AGO proteins form the RISC loading complex (RLC). The stable strand of the RNA duplex is bound to TRBP, whereas the other one interacts with the AGO protein [127,131]. An RNA helicase is thought to mediate the unwinding and removal of the unselected strand of the miRNA duplex.
The details of strand selection and RLC formation have been explicated through RNA interference (RNAi) experiments in *D. melanogaster* [132]: Dicer 2 and R2D2, which has two dsRBDs and is a sensor for thermodynamic asymmetry, form a stable heteroduplex and bind to the more stable end of the siRNA duplex and orientates AGO2 on the RNA duplex, that is responsible for the removal of passenger strand of siRNA duplex.

The Argonaute (AGO) family of proteins, that are expressed in all higher eukaryotes, in the cytoplasm of somatic cells, provide numerous possibilities for RNA-protein interactions that might underlie the proposed determinants of small RNA strand sorting. The interaction between AGOs and small RNAs occurs through several contact points in three characteristic domains of the protein: the PAZ, MID and PIWI domains (fig. 1.11) [133]:

- the PAZ domain has the capacity to bind the 2 nt 3'-overhang of the miRNA [125];
- the PIWI domain, which can fold into a structure similar to that of RNase H [134], harbours the residues required for catalytic activity: it is responsible for the endonucleolytic (slicing) activity of some AGO proteins (in humans only AGO2 possesses this characteristics);
- the MID domain, localized between PAZ and PIWI domains, forms a binding pocket that anchors the 5' phosphate of the terminal nucleotide of the miRNA.

In humans there are four AGO proteins, AGO1-4, that binds to miRNAs with only few differences, which suggests that human AGO1-4 might not have significantly differentiated functions.

Figure 1.11: **Domain organization of an AGO protein**
The interaction between AGOs and small RNAs occurs through several contact points in three characteristic domains of the protein: the PAZ, MID and PIWI domains. The PAZ domain (yellow box) hosts the 3'-end of the small RNA, whereas the MID domain (orange box) forms a binding pocket that anchors the 5' phosphate of the terminal nucleotide of the small RNA. The PIWI domain (red box) harbors the residues required for catalytic activity. Thus, cleavage-competent AGO proteins carry out endonucleolytic cleavage of target transcripts through their PIWI domain.

Now the miRNA is in its mature form, it binds to the target mRNA and carries out its functions as guide molecules in translational control or cleavage of certain mRNAs (fig. 1.12).

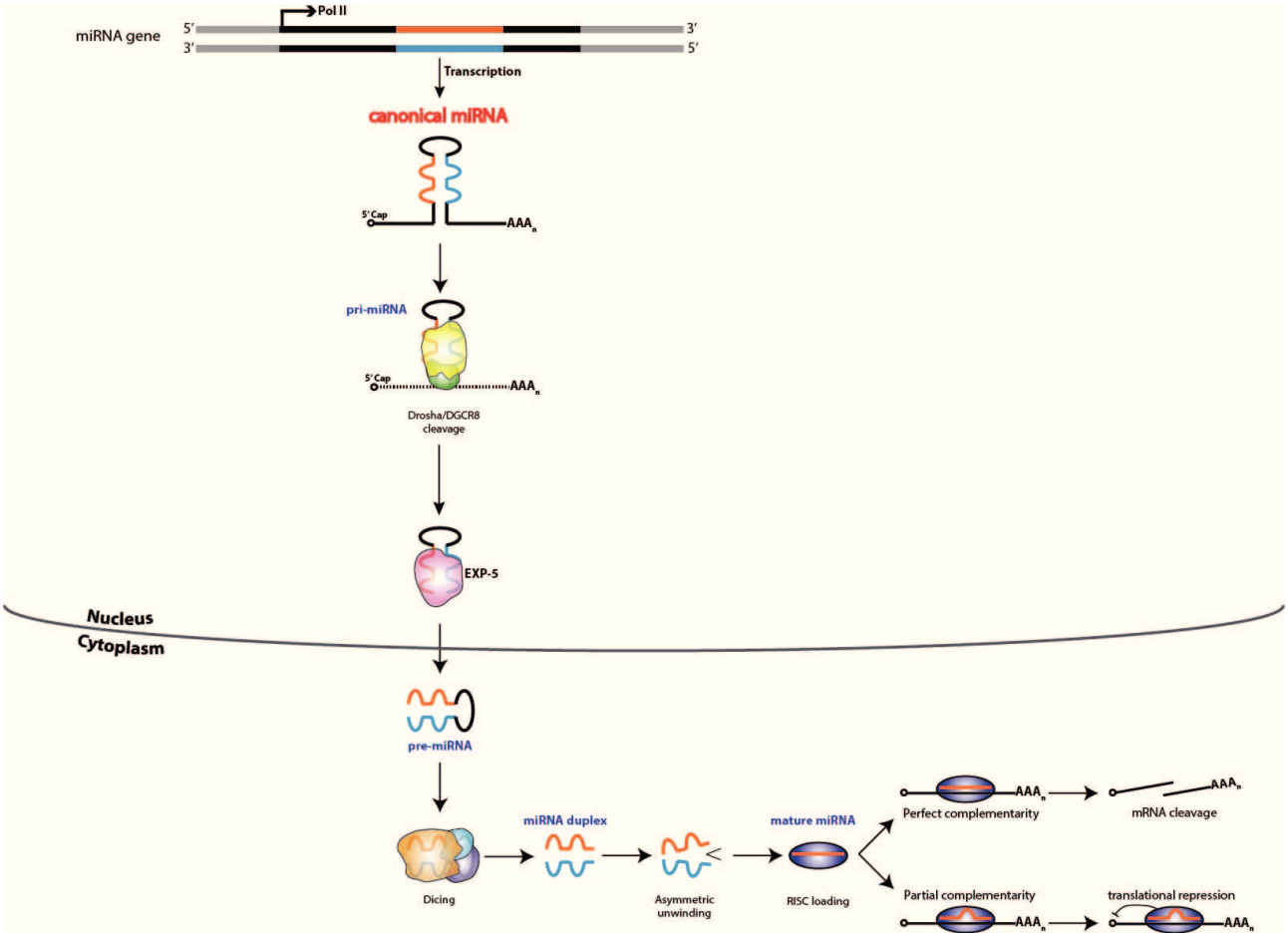


Figure 1.12: **Overview of canonical miRNA biogenesis pathway**

miRNA genes are transcribed by RNA Pol II. The initiation step -cropping- by Drosha/DGCR8 results in ∼70 nt pre-miRNAs. pre-miRNAs are exported by Exp5. Following export, Dicer catalyses the second processing step -dicing- to produce miRNA duplexes. The duplex is separated and one strand is selected as mature miRNA, while the other is degraded. The final products are incorporated into miRISC complex to function as guide molecules in translational control or cleavage of certain mRNAs.

1.3.3 Transcriptional and post-transcriptional regulation of miRNA maturation

Due to the fact that miRNAs are important regulators of gene expression, their biogenesis need to be finely controlled. This control can be realized both at the transcriptional and post-transcriptional level of miRNA processing.
In general, the transcription efficiency of the genes where the miRNAs are embedded determines the level of their expression and this is typical for miRNAs located in introns of coding genes. In addition, there are examples of miRNA-containing transcription units activated/regulated by transcription factors in response to specific stimuli or developmental changes. Several RNA Pol II associated transcription factors control miRNA gene transcription [117]. For example, myogenin and myoblast determination factor 1 (MYOD1) bind upstream miR-1, miR-133 and induce their transcription during myogenesis [135, 136]. Two antagonistic transcription factors, CREB and REST, function in an opposite and temporally regulated manner to control the expression of miR-9-2. In proliferating cells REST is active on the miR-9-2 promoter and prevents transcriptional activity. When neuronal differentiation occurs, CREB is phosphorylated and it promotes transcription [137]. Similarly some miRNAs are under the control of tumor-suppressive or oncogenic transcription factors. Indeed, members of the miR-34 family were originally identified as direct transcriptional targets of the tumor suppressor p53. A p53 responsive element is present in the promoter of miR-34a and in the putative promoter of miR-34b and miR-34c (see below for details). In this case the proposed mechanism is that following a DNA damage, there is an increase in the expression of p53 and, as a consequence, of the miR-34 family, that downregulates genes of the cell cycle, promoting apoptosis and arrest in the G1-phase. However, in some tumors, it has been noticed a decrease in the expression levels of these miRNAs [138–140]. Another example is the miR-17-92 cluster that, activated by the oncogenic transcription factor MYC, promotes cell growth and angiogenesis [141].
Further level of miRNA transcription regulation may occur through epigenetic control. If CpG-rich sequences undergo DNA methylation, they may inhibit miRNA expression [142].

Post-transcriptional regulation is another important and not completely understood way of regulation: frequently in a single transcription unit there are two or more miRNAs, that, even transcribed at the same time, they do not have the same expression levels.

Several proteins, activators and/or repressors regulate processing either by interacting with Drosha or Dicer or by binding to miRNA precursors [143]. For example p53 facilitates Drosha-mediated pri-miRNA processing of miR-16-1, miR-143 and miR-145 miRNAs, interacting with p68 and p72 helicases, that are components of the Microprocessor complex [144]. lin-28 binds to the terminal loop of pri-let-7 and inhibits pri and pre-miRNA processing, interfering with the cleavage by Drosha and blocking the processing by Dicer because induces the uridynilation of the pre-miRNA [145]. SMAD proteins, interacting with Drosha/pri-miRNA complex through protein cofactors or recognizing consensus site in pri-miRNA, facilitates Drosha-mediated pri-miR-21 processing [146].

Also several splicing factors function as miRNA processing regulators: SF2/ASF binds to pri-miR-7 and promotes its cleavage by Drosha [147], hnRNP A1 binds to the loop regions of pri-miR-18a and facilitates its Drosha-mediated processing [148]. Similarly, the splicing protein KSRP (K-homology splicing regulatory protein) regulates the processing of a subset of miRNAs binding to a specific G-rich motif present in the terminal loop and stabilizing the interaction with Drosha [149].

In addition both chemical modifications and editing, modifying the sequence of the miRNAs can affect the availability of the miRNAs. Uridylation of pre-let-7 blocks Dicer processing and facilitates the degradation of the miRNA precursor [145]. Editing of pri-miRNAs or pre-miRNAs by adenosine deaminases that act on RNA (ADAR1 and ADAR2) affects accumulation of mature miRNA and might also influence miRNA target specificity. However editing can also enhance Drosha processing [150, 151]. Turnover of miRNAs also contributes to miRNA regulation. Certain miRNAs, such as miR-29b, might be degraded much more rapidly than other miRNAs [152], suggesting a specific recognition of miRNA sequences by nucleases.

A feedback loop between Drosha and DGCR8 is also observed: Drosha downregulates DGCR8 cleaving DGCR8 mRNA, while DGCR8 stabilizes Drosha through protein-protein interactions [153].

1.3.4 miRNA targets and their biological functions

Most mammalian miRNA genes have multiple isoforms, these different paralogues are probably the result of gene duplication during evolution. Generally, the paralogues have identical sequence in the 5'-end at nucleotide position 2-8 (fig. 1.13): this sequence, known as seed

sequence, are crucial for the interaction of the miRNAs with their mRNA targets via base pairing, and, for this reason, it seems that paralogues can act redundantly. It was originally thought that a perfect base pairing between the seed region and the target was necessary for the function of a miRNA, however there are functional miRNA sites that contain mismatches or bulged nucleotides in the seed region, such as the let-7 miRNA that act on lin-41 mRNA in *C. elegans* [154]. The 3' half sequence (position 13-16) of the miRNAs also contribute to target binding, stabilizing the interaction particularly when the seed matching is suboptimal. The miRNA-mRNA duplex contains, instead, mismatches and bulges in the central region (position 10-12) to prevent endonucleolytic cleavage of mRNA by an RNAi mechanism.

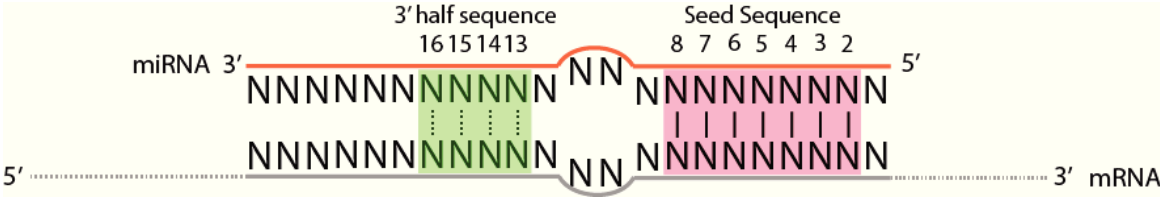


Figure 1.13: **miRNA-mRNA interactions**
miRNAs interact with their mRNA targets by base pairing. A perfect base pairing (solid lines) between miRNA and its target is present at nucleotide position 2-8 (seed sequence , shown in red). Bulges or mismatches are present in the central region of the duplex. A good base pairing (dotted lines) between miRNA 3' half sequence and target (green) also stabilize the interaction. Nucleotide position is indicated above the miRNA sequence.

Moreover, the mature sequence of a miRNA is not constant, the 5' and 3' ends can be heterogeneous and the situation is further complicated by the evidence that some functional target sites do not contain the seed matching sequence, but instead exhibit 11-12 continuous base pairs in the central region of the miRNA [155].

The majority of miRNA binding sites are located in the 3'UTR of the target mRNA, even if there are evidences for target sequences in the 5'UTR and coding region of a gene [156]. Multiple sites for the same or different miRNAs are frequently found on the same target. They are required for effective repression and they tend to act cooperatively [157].

The exact mechanisms of gene repression are still being elucidated, miRNA can either inhibit translation, at the initiation or post-initiation step, of target mRNAs or induce mRNA degradation through deadenylation and/or destabilization of the target mRNA.

1.3.4.1 Translation repression

1.3.4.2 Repression at the initiation step

mRNAs without a functional 5'-cap structure, or mRNAs containing internal ribosome entry sites (IRESs), which translation is cap-independent are refractory to miRNA-mediated translational repression [158]. miRNAs interfere with eIF4E function or its recruitment to the 5'-cap. One mechanism by which miRNAs impede the cap recognition step of translation initiation has been proposed by Kiriakidou et al. [159]: they observed that the AGO2 MID domain contains a structure similar to the cap-binding protein eIF4E, suggesting that AGO2 can binds directly to the cap structure and competes with eIF4E. Moreover Kinch and Grishin [160] indicated that AGO2 shares extremely limited structural similarity to eIF4E; therefore remain to be understood if AGO proteins bind directly to the cap structure.
Other hypotheses suggest that the miRISC inhibits the formation of ribosome initiation complex by interfering with eIF4F-cap recognition and 40S small ribosomal subunit recruitment [161], by antagonizing 60S subunit joining [127] and preventing 80S ribosomal complex formation [162].

1.3.4.3 Repression at post-initiation steps

Several studies try to demonstrate that miRNAs inhibit translation at post-initiation steps. The first observation originate from polysomal sedimentation analyses: the researchers showed that in *C. elegans* lin-14 and lin-28 mRNAs, which are targets of lin-4 miRNA, remain associated with polysomes during larval development in spite of reduced protein levels [163] and similar result were seen also in mammalian cells. Some reports indicate that IRES-driven mRNA translation is sensitive to miRNAs and in particular Sharp et al. [164] proposes that miRNAs can antagonize elongation by causing premature termination and ribosome drop-off.
Another hypothesis is that miRNA machinery recruits proteolytic enzymes to polysomes, that can degrade the nascent polypeptides [165], but this idea does not seem to be real because endoplasmic reticulum is thought to protect polypeptides from proteolytic cleavage.

The initiation and post-initiation mechanisms are not mutually exclusive: it is possible that initiation is always inhibited, but when the elongation step is also repressed, ribosomes would queue on the mRNA, thereby masking the effect of an initiation block.

1.3.4.4 mRNA deadenylation

mRNA degradation initiates with the shortening of the poly(A) tail, that requires a 3'→5' exoribonucleases and the CCR4-NOT1 complex or the 5'-cap is first removed by the decapping DCP1-DCP2 complex and then the RNA is degraded by XRN2, a 5'→3' exonuclease. The miRNA-mediated deadenylation require the AGO and GW182 proteins [166]. The GW182-AGO interaction is mediated via GW repeats in the GW182 N-terminus through binding to the AGO MID/PIWI domains. GW182 is able to mediate repression of the target even in the absence of AGO proteins, meaning that AGO is a scaffold to recruit GW1182 to mRNA. GW182 can be tethered to the 3'UTR of the target mRNA and recruits the CCR4-NOT1 complex to promote deadenylation [166]. Translation inhibition and deadenylation are the two main mechanisms that have emerged from experimental research. It is possible that there are other events that participate in miRNA function and that they can overlap in a cell or developmental dependent manner. It is also possible that the experimental design can favor one mechanism over another. Much evidence exists indicating that many components of the miRNA machinery and the repression process itself may not be localized in the cytosol but that they occur in association with different cellular organelles or structures. miRNA can accumulate in discrete cytoplasmic foci, such as P-bodies. Some studies have demonstrated that AGO1-2 are localized to mammalian P-bodies, indicating a functional link between P-bodies and a miRNA-induced repression [167].

1.3.4.5 Translation activation

miRNAs binding to 3'UTR or 5'UTR of target mRNAs, in specific situations, activate rather than repress translation [168]. Although "RNA activation" mechanism remains to be elucidated, the process may require the AGO2 protein and could be associated with histone changes linked to gene activation. The importance of faithful miRNA expression has been implicated in numerous biological and cellular events. miRNAs can potentially regulate every aspect of cellular activity and they are related to the molecular mechanisms of various clinical diseases and they were also well documented in cancer. miRNA signatures have diagnostic and prognostic value, and may become valuable clinical tools in cancer therapy. They could have therapeutic implications, in which disease-related miRNAs could be antagonized or functional miRNAs restored.

1.3.5 Genomic localization of pri-miRNA and pri-miRNA-like hairpins

miRNA genes are either intergenic, located in non-coding regions >1 kb away from annotated genes, or intragenic, located, in sense or antisense orientation, within protein coding genes or in non transcriptional units (TUs). The position of miRNA hairpins in the "transcriptome" is mostly derived from bioinfomatics analysis with gene and miRNA annotations and revealed that most of them are located in introns [169]. Approximately 62% of miRNAs are located in the intronic region of both protein-coding (43%) and non-coding TUs (19%), 9% in exonic regions or UTRs and 2% in alternatively spliced exons [169]. A small fraction of miRNA hairpins (1-2%), named Splice site Overlapping (SO)-miRNAs are also juxtaposed either to donor or acceptor splice sites (fig. 1.14) [170,171]. In addition, according to recent CLIP data in HEK293T cells, 45% of DGCR8 targets are mapped to intergenic regions, 43% to protein-coding genes and 5% to long non-conding RNAs. Most of these DGCR8 targets are real pri-miRNA hairpins but a a significant amount does not have a corresponding miRNA in database. These pri-miRNA-like hairpins have a major role in regulating the fate or quality of the transcript [172]. miRNAs hairpins may be situated alone or grouped into clusters. In protein-coding genes, 26% of the clusters were in introns, followed by exons (12%). In some cases the same miRNA, depending from the promoter used, can originate from different TUs, but few data are available regarding this aspect. miRNAs have also been mapped in repetitive elements, such as Alu repeats or long interspersed nuclear elements [173].

1.3.6 Intronic hairpins

The majority of pri-miRNA hairpins are located in larger introns (62%), far away from splice sites, to avoid interference between the MPC and the spliceosome [169, 174]: miRNA cleavage from introns, that occurs co-transcriptionally before intron splicing, does not alter, in general, the levels of mature mRNAs [169].

According to the exon-tethering model, the C-terminal domain of RNA Pol II can tether an upstream exon before reaching the next one, even if the intervening sequences are cleaved by the MPC [175]. In this case, the MPC can excise the pre-miRNA co-transcriptionally from the intron, without affecting the recognition of the flanking splice sites by the spliceosome.

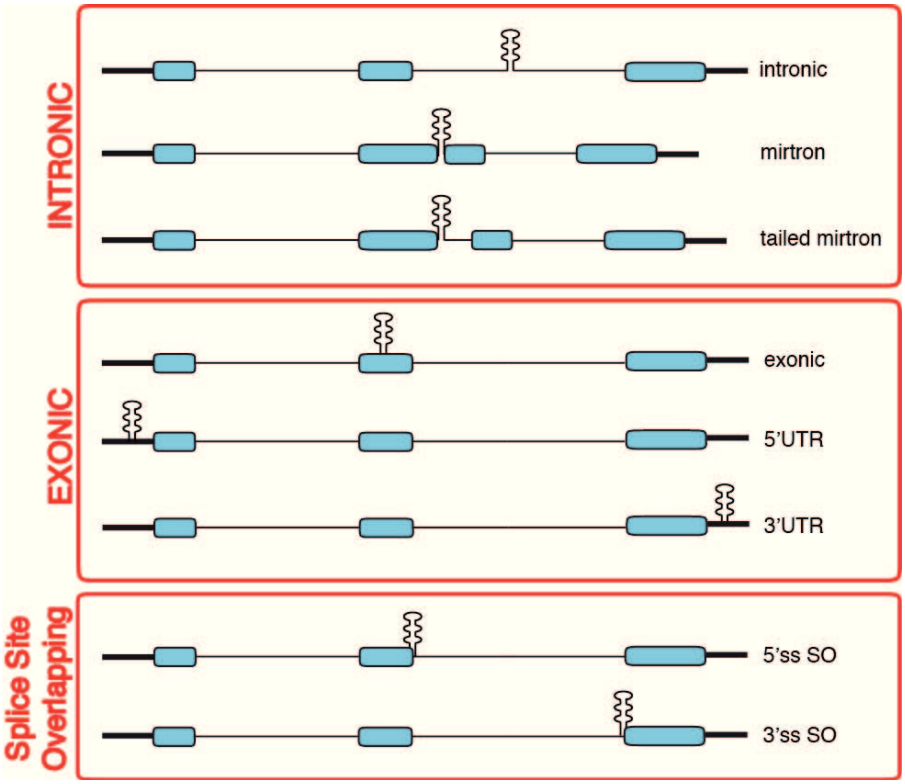


Figure 1.14: **Genomic location of pri-miRNA and pri-miRNA-like hairpins**
pri-miRNA and pri-miRNA-like hairpins can be divided into three groups according to their position in the nascent transcript. The intronic group contains the intronic hairpins, the mirtrons/simtrons and the tailed mirtrons. The exonic group includes hairpins located in central exons and in the 5' and 3'UTRs (or in corresponding first and terminal exons for non-coding transcripts). In the Splice site Overlapping (SO) group, the hairpins overlap with the 5' or the 3' splice sites. Blue boxes represent the exons, thin lines the introns and the thick black lines the 5' or 3'UTRs.

This interesting model might be important for longer introns before the polymerase reaches the next exon, whereas in shorter introns miRNA cropping might occur when the two flanking exons are already engaged in splicing or even later on the excised lariat. In any case, as a common precursor generates both the miRNA and the mRNA, one pre-mRNA molecule will produce one miRNA precursor and one mRNA. However, in experimental systems, where the miRNA hairpins are relatively distant from splice sites, MPC-dependent cropping was shown to have a small influence on splicing, either in a positive [169, 176] or negative [174] manner. Furthermore, spliceosome assembly can facilitate cropping [176, 177]. In the case of the melanoma invasion suppressor miR-211, expressed from intron 6 of melastatin, mutation of the 5'ss was shown to reduce the biosynthesis of the miRNA and the recruitment of Drosha to the

pri-miRNA hairpin. Moreover, knockdown of the U1 snRNP specific factor U1-70K globally altered the production of intronic miRNAs [176], suggesting that the 5'ss has a positive influence on cropping of intronic miRNAs.

1.3.6.1 Alternative intronic miRNA biogenesis pathways: mirtrons and simtrons

Apart the canonical miRNA biogenesis, several alternative miRNA biogenesis pathway have emerged in the last few years. Different type of miRNAs are processed by non-canonical miRNA biogenesis pathways. These comprehend mirtron and the recently reported simtron pathways (fig. 1.15).

Mirtrons, that were originally recognized in flies and worms [178, 179] and recently validated in mammals [180], are small RNAs that are located in the introns of the mRNA encoding host genes. The small RNA-generating hairpin can be defined either by the entire length of the intron in which it is located, starting and ending precisely with splice donor and splice acceptor sites (canonical mirtron) or can be resides at one end of the intron: the 3' tailed mirtron, in which the 5' end of this mirtron hairpin locates at the 5'-end of the intron, while the 3'-end is produced from the middle of the intron; and the 5' tailed mirtron, where the 3'-end of this hairpin corresponds to the 3'-end of the intron.

The first processing step is performed by the splicing machinery and not by Drosha: the pre-miRNA excised by splicing is initially in the form of an intron lariat, in which the 3' branch point is ligated to the 5'-end of the intron; the structure is subsequently linearized by the debranching enzyme (DBR1 in humans) and the intron, adopting a pre-miRNA fold, can be exported to the cytoplasm by Exp5, and then recognized and cleaved by Dicer to form a mature miRNA [179]. However, splicing inhibition is compensated by canonical MPC processing.

In the case of 3' tailed mirtrons (such as miR-1017 in *D. melanogaster*), after splicing and debranching, the 3' tail following the hairpin is trimmed by the RNA exosome, the major eukaryotic 3'→5' exonuclease complex. The trimming reaction occurs in the nucleus and is prerequisite to shorten the tail sufficiently to serve as an Exp5 substrate [181]. No 3' tailed mirtrons have yet been identified in vertebrate species; instead some 5' tailed mirtrons have been found in various mammals. The 5' tail can be trimmed by XRN1/2, the major 5'→3' exonucleases in eukaryotes, although this remains to be tested (fig. 1.15).

In 2012 Havens et al. identified some splicing-independent mirtron-like miRNAs, the **simtrons** [182]. They showed that two miRNAs, miR-1225 and miR-1228, originally predicted to be mirtrons, due to their predicted hairpin structure and genomic location spanning short introns, are not processed by the mirtron-processing pathway (they do not require splicing for their biogenesis) or by the canonical miRNA pathway: their maturation involves Drosha (and an unknown binding partner), but does not require the others miRNA biogenesis components DGCR8, Exp5 and Dicer, even if they enter the RISC complex with any of four human AGO proteins and are functional in targeted gene silencing (fig. 1.15). However, Drosha processing may not be a strict requirement. In particular, simtrons seem to be sensitive to Drosha activity when splicing is inactive, indicating that simtrons can be excised by either the mirtron processing pathway or by a pathway involving Drosha.

1.3.7 Exonic hairpins

Exonic pri-miRNA or pri-miRNA-like hairpins constitute a small part (9%) of the known human miRNAs. Most of them are located in non-coding genes or in UTRs of coding transcripts, possibly to avoid the interference with the protein-coding sequences and splicing regulatory elements. However, there are some cases in which the exonic position of the hairpin influences the quality of the mRNA or its function, as for example the pri-miRNA-like hairpin located in the 5'UTR of DGCR8 gene and the pri-miR-198 in the 3'UTR of FSTL1 gene.
The pri-miRNA-like hairpin in the 5'UTR of DGCR8 has an important effect on the fate of the transcript. This stem-loop structure is a substrate for the MPC and, once it is cropped, the transcript cannot be processed further, with a decrease in both the DGCR8 mRNA and protein [153, 183]. The presence of a second hairpin in the first exon also contributes to the DGCR8 regulation. Therefore, the DGCR8 5'UTR hairpin mainly promotes mRNA degradation and is a useful system to finely tune the levels of DGCR8. Moreover, it constitutes a feedback loop mechanism to control the levels of the MPC itself, because DGCR8 stabilizes Drosha through protein–protein interactions [153].

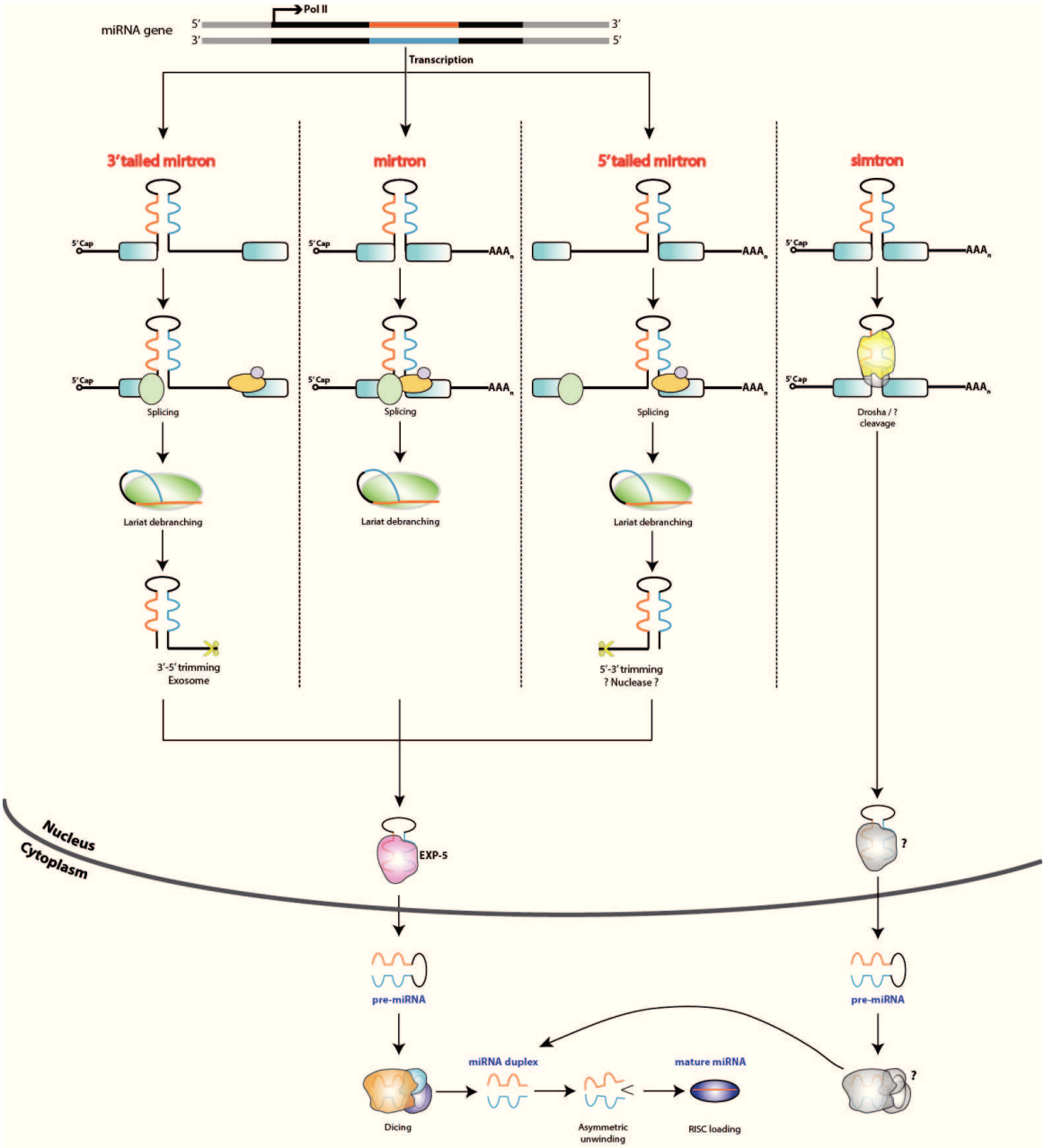


Figure 1.15: **The mirtron and simtron pathways**

Mirtrons are pre-miRNA-like hairpin introns that are directly produced by splicing and debranched, exported from the nucleus by Exp5, cleaved by Dicer and enter the RISC complex. Tailed mirtrons also undergo splicing and debranching, after which the tails on the resulting hairpins are trimmed back. Simtrons are splicing-independent mirtron-like miRNAs, that are processed by Drosha and an unknown binding partner. They are further processed by unknown factors (?) and enter the RISC complex, producing functional mature miRNAs.

Another example of MPC-dependent transcription regulation occurs on the follistatin-like1 (FSTL1) gene and its 3'UTR pri-miR-198 hairpin, where it switches a protein-coding toward a miRNA-producing transcript in wound healing [184]. In normal conditions, human skin keratinocytes express miR-198 but not FSTL1, indeed the transcript is retained in the nucleus where it is used to produce only the miRNA. When skin is wounded, cells start to migrate, produce FSTL1, and progressively reduce miR-198 expression: after injury, the mRNA is not cleaved by the MPC and goes to the cytoplasm where it is translated into protein. As FSTL1 and miR-198 have anti-migratory and pro-migratory effects, respectively, the switch between the two alternative gene products contributes to wound re-epithelialization. A major player in regulating this switch is the RNA-binding protein KSRP. This factor was previously shown to regulate cropping efficiency of a set of miRNAs through its interaction with G-rich sequences in the stem–loop [149]. KSRP binds to a GUG motifs within the terminal loop of pri-miR-198 and promotes its cropping. As the 3'UTR MPC-dependent hairpin found in FSTL1 is proximal to the polyadenylation site, it is not clear if the miRNA cropping occurs before or after the polyadenylation process. However, it is possible that in particular conditions the miRNA cropping can function as an alternative 3'-end processing mechanism in competition with the classical one, as Drosha silencing was shown to increase the levels of FSTL1 protein.
This peculiar type of post-transcriptional regulation might be present in other genes that contain miRNA hairpins located on 3'UTR. For example, the coding DCP1A, HOXA7, SNX12 and INO80E genes contain pri-miRNA-like secondary structures in their 3'UTR, which are DGCR8-binding sites according to DGCR8 HITS-CLIP data [172]. In addition, Cáceres and collaborators identified also some hairpin structures located in alternatively spliced exons of coding genes. Detailed analysis of four cases showed that silencing of either Drosha or DGCR8 changes their pattern of splicing, suggesting that the MPC specifically cleaves and destabilizes the mRNA isoforms that contain the hairpins, regulating alternative splicing [172].

An additional example is represented by the pri-miR-133b, located in the terminal exon of the long noncoding linc-MD1 gene, whose biosynthesis is regulated by the splicing factor HuR [185]. linc-MD1 acts in the cytoplasm repressing target genes that promote myogenesis and this effect is partially mediated by its 'sponge' activity on its own producing miR-133b. If Drosha crops pri-miR-133b in the nucleus, the pre-mRNA linc-MD1 will be used to generate the miRNA;

alternatively, if linc-MD1 is not processed by the MPC, it will act as a sponge in the cytoplasm for miR-133b and probably for other miRNAs of the same family [101, 185]. Thus, the decision to cleave a pri-miRNA hairpin located in a terminal exon switches between two alternative and antagonistic gene products: a cytoplasmic sponge RNA and a miRNA.

1.3.8 Splice site Overlapping miRNAs

Splice site Overlapping (SO)-miRNAs are characterized by the presence of overlapping miRNA hairpins and splice sites on the nascent transcripts [170, 171]. The hairpins can be juxtaposed either with the donor or the acceptor splice sites and, accordingly, they are classified as 5' SO-miRNAs or 3' SO-miRNAs, respectively. SO-miRNAs are present in vertebrates and non vertebrates and represent a small but significant fraction of annotated miRNAs (1–2%). They are very heterogeneous in respect to the position of the splice site in the hairpin, that can be located either near the ssRNA-dsRNA junction at the base of the hairpin or more internally in the stems. Interestingly, the majority of SO-miRNAs are embedded in protein-coding transcripts, some are encoded by putative open reading frames (ORFs) and others are within annotated non-coding transcripts. The majority of SO-miRNAs are poorly described in literature, even if for some of them there are clearly correlations with physio-pathological conditions. For example miR-205 is described both as a tumor suppressor and an oncogene. It can inhibit or favor cell proliferation and invasion, depending on the tumor considered and the target genes [186], and has a protective role against oxidative stress and endoplasmic reticulum stress in renal tubular cells [187]. miR-133a-2 is described as a tumor suppressor deregulated in colorectal and bladder cancers [188, 189] but its main functions are in maintaining the structure and biogenesis of skeletal muscle [190] and heart [191].

1.3.9 miR-34 family

In mammalians, the miR-34 family comprises three processed miRNAs: miR-34a, miR-34b and miR-34c that originate from two distinct genomic location: miR-34a is located on chromosome 1 and originates from the EF609116 transcript, whereas miR-34b and miR-34c share a common transcript BC021736 from chromosome 11. The three miRNAs are 80% homologous with a perfect sequence identity in their seed regions.

BC021736 is a non-coding transcript with two exons: the 3' SO-miR-34b overlap with the

junction between intron 1 and exon 2, while miR-34c is located in exon 2. The transcript is poorly conserved in vertebrates. The only conserved regions are the p53-responsive promoter [192], the 5'ss of exon 1, a CpG island upstream of the miRNAs, the 3'ss of exon 2, the miRNAs themselves and the polyadenylation site. According to Lujambo et al. [193], the CpG island can also function as a promoter and therefore there are two possible transcription start sites (TSSs), one located at the 5' side of exon 1 and the other upstream of exon 2 (fig. 1.16).

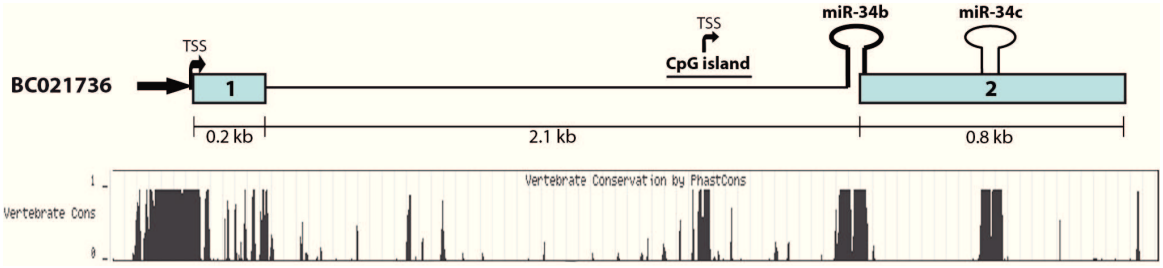


Figure 1.16: **miR-34b/c transcript and vertebrate conservation**
[Upper panel] Schematic representation of the BC021736 non-coding transcript. Light blue boxes represent exons, thin line the intron and hairpin the miRNAs. The black arrow is the p53-responsive promoter, the curved black arrows the TSSs and the thick black line a CpG island. miR-34b and miR-34c hairpins are indicated. The kb lenght of exons and introns is reported under the picture. [Lower panel] Conservation profile of the transcript in vertebrates. The transcript is poorly conserved, except for the promoter region, the splice sites, the CpG island, the miRNAs themselves and the poly(A) signal.

The predicted miR-34b hairpin has a typical pri-miRNA secondary structure with three stems (A, B and C stems) and a terminal loop. The AG dinucleotide of the 3'ss is located near the end of the hairpin in stem A, four nucleotide above the ssRNA-dsRNA junction. According to miRBase database, the predominant mature miRNA is the miR-34b-3p, which is expressed from the 3' arm (fig. 1.17).
The miR-34 family is extensively described in literature and involved in several physio-pathological conditions. miR-34b has been associated with p53 and cell proliferation control. It was observed that miR-34b was strongly downregulated in primary mouse ovarian surface epithelium cells subjected to acute inactivation of p53 [194] or in p53-deficient mouse embryonic fibroblasts [192]. Several other reports have then confirmed these evidences, and have linked miR-34 family to different types of tumors such as pancreatic cancer [195], prostate cancer [196], hepatocellular carcinoma [197], breast cancer [198], neuroblastoma [199], osteosarcoma [200] and acute myeloid leukemia [201].

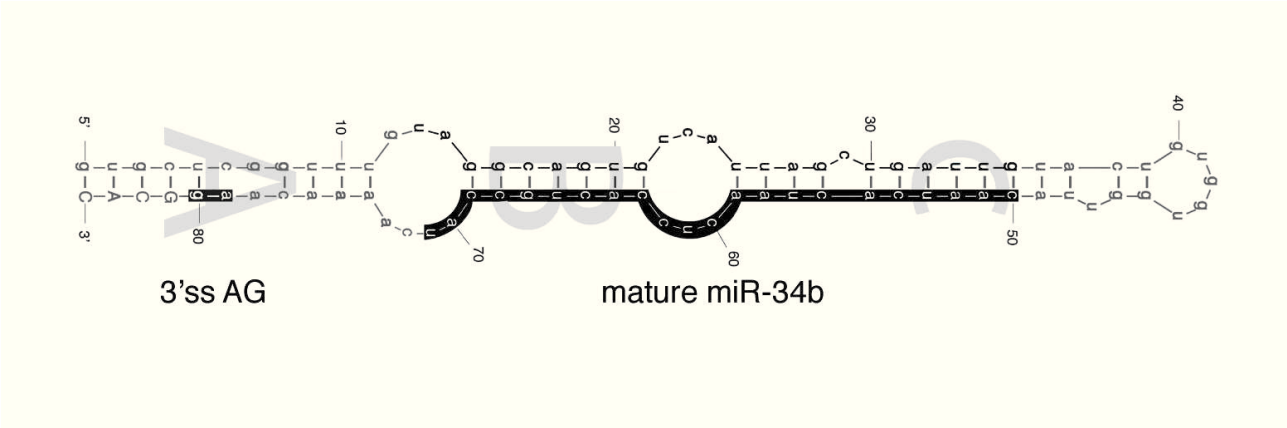


Figure 1.17: **Predicted RNA secondary structure of miR-34b hairpin**
The secondary structure of miR-34b hairpin has been calculated through the Mfold server. The A, B and C stems, the AG dinucleotide of the 3' ss and the predominant mature miR-34b-3p are indicated.

Other reports focused on the relation between miR-34s and epigenetic changes. Epigenetic silencing of miR-34b and miR-34c was found in various malignant tumors including colorectal cancer [202], ovarian cancer [203], lymph node metastatic cancer [193], small-cell lung cancer [142], malignant pleural mesothelioma [204], gastric cancer [205], oral squamous cell carcinoma [206], causing tumor cell proliferation and invasiveness, in multiple myeloma [207] and in malignant melanoma in which it correlated with metastatic potential [193, 208].
Other than cancers, miR-34 family has a role in spermatogenesis [209], heart function and pathological cardiac remodeling [210], osteoblast proliferation and bone development [211, 212]. Another role of miR-34s occurs in the central nervous system being associated with ageing and neurodegeneration [167], central stress response [213], neuronal differentiation [214], and neurodegeneration associated to Alzheimer's disease [215]. miR-34b has been specifically involved in Parkinson [216] and Huntington's diseases [217].

1.4 Molecular architecture and biology of epidermal keratinocytes

1.4.1 Structure of the skin

The skin is the largest organ system of the body, which consists of epithelial layer (epidermis), connective tissue layer (dermis) and adipose layer (hypodermis). Its main function is to form a barrier, that protects the underlying tissues against environmental damage by microbial and other harmful substances, prevents the dehydration of the organism and function to regulate temperature, produce hormones and vitamins, such as vitamin D, and responds to environmental factors including ultraviolet (UV) radiation [218].

Hypodermis is the undermost layer of the skin, it is composed of fat and connective tissue and is fundamental for the body insulation.

Dermis is comprised mainly of connective tissue, collagen bundles and elastic fiber, and is responsible for the structural strength of the skin. In the layer are also present blood vessels, nerves, sweat and sebaceous glands and hair follicles. It does not only protect body from the mechanical injury but also binds water, participates in thermal regulation and contains sensory receptors.

Epidermis is the outer layer of the skin, it is a stratified, squamous, continually renewing ephitelium, that forms the protective covering of the skin. Keratinocytes are the predominant cells type in the epidermis and represent the 90-95% of cells expressed in this layer.

1.4.2 Epidermal growth and differentiation

In the epidermis keratinocytes are sorted and organized into four histologically distinct cellular layers (fig. 1.18 - upper panel), that correspond to progressive stages of differentiation [219].

The basal layer (**stratum basale**) is the deepest layer of the epidermis, made up of a single layer of basal, proliferative and undifferentiated keratinocytes. These cuboidal-shaped cells are the epidermal stem cells, precursors of all the keratinocytes, that have an high clonogenic potential and divide to produce one cell that remains in the basal layer and conserved the proliferative potential and one that is destined to differentiate and is pushed into the other strata [220]. Basal keratinocytes are structurally and functionally associated with components of the underlying basal membrane via hemidesmosomes. This association contributes to epidermal structural integrity and provides regulatory signals to control proliferation, migration and differentiation.

The spinous layer (**stratum spinosum**) is spiny, as the name suggests, due to the histological appearance of the resident keratinocytes, that are join and strength one to each other by the desmosomes: the formation of 'spines', indeed, is a staining process artifact due to shrinking of the microfilaments between desmosomes. The stratum is composed of different layers of keratinocytes, that have a polyhedral shape in the lower layer and become larger and more flattened in the upper layers. As new keratinocytes are produced atop the stratum basale, the keratinocytes of the stratum spinosum are pushed into the stratum granulosum. At this stage, the keratinocytes start to synthesize involucrin a soluble, transglutaminase substrate protein, precursor of the cornified envelope.

The granular layer (**stratum granulosum**) is a flatter and thin layer of cells with a grainy appearance, that produce a large amounts of keratin and keratohyalin proteins, which accumulates as lamellar granules within the cells. The granular cells synthesize, modify and cross-link proteins involved in keratinization and, at the same time produce different cellular organelles degradative enzymes, that are involved in their own programmed destruction.

The (**stratum corneum**) is the most superficial layer of the epidermis. It forms a barrier to the outside environment and contributes to mechanical protection. It is composed of multiple layers of polyhedral, anucleated, post-mitotic, terminally differentiated keratinocytes (corneocytes) embedded within a highly hydrophobic lipid matrix. The cell membrane is replaced by a layer of ceramides which become covalently linked to a rigid cornified envelope. Cells in this layer are shed periodically during the desquamation process and are replaced by cells pushed up from the deeper strata: this fine balance between keratinocyte proliferation and differentiation contribute to the maintenance of epidermal homeostasis [221].

Keratinocytes differentiation from the basal to the cornified layer is marked by several morphological and biochemical changes. Changes in gene expression profile *in vivo*, such as increase involucrin expression from the stratum spinosum to the stratum corneum, are, therefore, used as marker of keratinocytes differentiation in culture [222].

The epidermis is separated from the underlying dermis by the basement membrane. It consists of two layers: a clear layer, called **lamina lucida** and a dense layer, the **lamina densa**. The stable attachment between the epidermis, the basement membrane and the dermis, that is of fundamental importance for maintaining skin integrity and epidermal homeostasis, is mediated by the epithelial cell adhesion complex containing hemidesmosomes (HDs), anchoring fibrils and

anchoring filaments (fig. 1.18 - lower panel). HDs, located in the stratum basale, are connect to the basement membrane through $\alpha 6\beta 4$ integrin. The extracellular part of the integrin $\alpha 6\beta 4$ binds laminin 5, which, spanning the lamina lucida, forms a bridge to the lamina densa and interacts with the collagen VII, forming a continuous structural link between the different compartments. This connection is further stabilized by anchoring fibrils of collagen XVII, which play a critical role in maintaining the linkage between the keratinocytes in the stratum basale and the extracellular structural elements involved in epidermal adhesion [223, 224].
Laminin is a family basement membrane proteins, composed of three subunits (alpha, beta, and gamma) of different types: for example, laminin subunit beta 3 (**LAMB3**) serves as the beta chain in laminin 5. The type VII collagen fibril are composed of three identical alpha collagen chains. **COL7A1** encodes the alpha chain of type VII collagen, which interacts with laminin 5. Collagen XVII transmembrane protein is a homotrimer of three collagen XVII alpha 1 chains (**COL17A1**). Mutations in LAMB3 and COL7A1 and COL17A1 genes are associated with various types of epidermolysis bullosa [225].

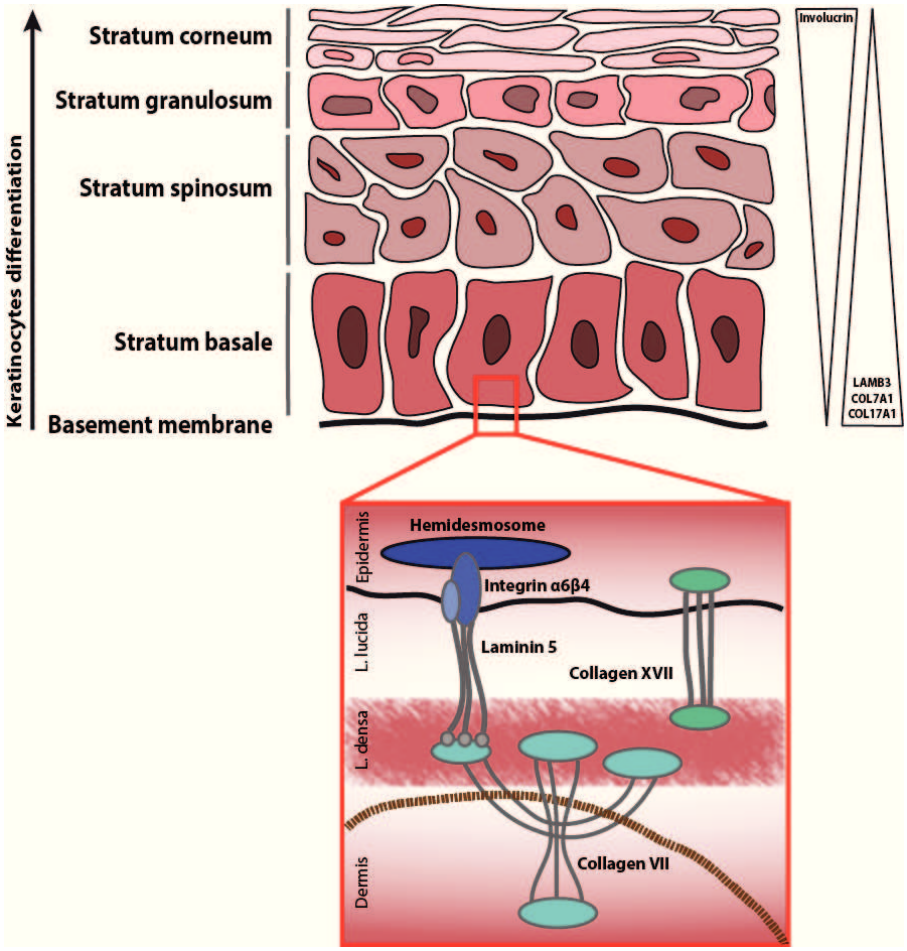


Figure 1.18: **Epidermal keratinocytes differentiation**
[Upper panel] Keratinocytes proliferate within the basal cell layer (stratum basale) and, when the differentiation program is activated, migrate into the into the spinous (stratum spinosum) and granular (stratum granulosum) layers, becoming anucleated and increasingly compacted in size in the stratum corneum. Involucrin protein, produced in the stratum spinosum, increase its expression during differentiation, while the basement membrane proteins decrease [Lower panel] The adhesion of the epidermis to the underlying basement membrane and dermis is mediated by the epithelial cell adhesion complex containing hemidesmosomes, connected to with the underlying connective tissue through integrin $\alpha 6\beta 4$, laminin 5 and collagen VII, anchoring fibrils of collagen XVII and anchoring fibers.

Chapter 2

Materials and methods

2.1 Chemical reagents

General chemicals were purchased from Sigma-Aldrich, Merck, Gibco BRL, Boehringer Mannheim, Invitrogen, Fluka and Qiagen.

2.2 Standard solutions

All solutions are identified in the text except for the following:

- 5X TBE: 53g Tris-HCl, 27.5g Boric acid, 20ml 0.5M EDTA, pH 8.0 in 1l

- 6X DNA sample buffer: 0.25%w/v bromophenol blue, 0.25%w/v xylene cyanol FF, 30% v/v glycerol in H_2O

- 5X Running buffer: 30g Tris-HCl, 144g glycine, 5 g SDS in 1l

- 10X Blotting buffer: 30g Tris-HCl, 144g glycine, 20% methanol in 1l

- 10X Protein sample buffer: 20%w/v SDS, 1M DTT, 0.63M Tris-HCl (pH 7.0), 0.2%w/v bromophenol blue, 20%v/v glycerol, 10mM EDTA (pH 7.0)

- 1X PBS: 137mM NaCl, 2.7mM KCl, 10mM Na_2HPO_4, 1.8mM KH_2PO_4, pH 7.4

2.3 Bacterial culture

The *E. coli* K_{12} strain DH5α was transformed with the plasmids described in this study and used for their amplification. Plasmids were maintained in the short term as single colonies on agar plates at 4°C. Bacteria were amplified by an overnight incubation in Luria-Bertani medium (LB medium: 10g Difco Bactotryptone, 5g Oxoid yeast extract, 10g NaCl, pH 7.5 in 1l). Bacterial growth media were sterilized before use by autoclaving. Then ampicillin (Sigma-Aldrich) was added to the media at a final concentration of 100µg/ml.

2.4 Preparation of bacterial competent cells

Bacterial competent cells were prepared following the method described by Chung et al. [226]. *E. coli* strains were grown overnight in 10ml of LB at 37°C. The following day, 150ml of fresh LB were added and the cells were grown in the shaker at room temperature (RT) until the OD_{600} was 0.3-0.4. Then the cells were put on ice to stop the growth and centrifuged at 4°C, 1000g for 1 hour. The pellet was resuspended in 1/10 volume of cold 1X TSS solution (10% w/v PEG 4000, 5% v/v DMSO, 35mM $MgCl_2$, pH 6.5 in LB medium). The cells were aliquoted, rapidly frozen in liquid nitrogen and stored at -80°C. Competence was determined by transformation with 0.1ng of pUC19 control DNA plasmid and was deemed satisfactory if this procedure resulted in more than 100 colonies.

2.5 Transformation of bacteria

Transformation of ligation reaction were performed using half of the initial reaction volume (10µl). Transformation of positive clones was carried out using 1ng of the DNA plasmid. The DNA was incubated with 65µl of competent cells for 30 minutes on ice, followed by a heat shock at 42°C for 30 seconds and finally the cells were placed again on ice for 1 minute and then spread onto agarose plates containing the appropriate antibiotic concentration (100µg/ml of ampicillin). The plates were then incubated for 12-15 hours at 37°C.
When DNA inserts were cloned into β-galactosidase-based virgin plasmid (pUC19), 30µl of IPTG 100mM and 30µl of X-Gal (4% v/v in dimethylformamide) were spread onto the surface of the agarose before plating to facilitate screening of positive clones (white colonies) through

identification of β-galactosidase activity (blue colonies).

2.6 DNA preparation

2.6.1 Small scale preparation of plasmid DNA from bacterial cultures

Single bacterial colony was picked and transferred into 6ml of LB medium containing 100µg/ml of ampicillin. The culture was incubated overnight at 37°C in a shaking incubator. The Wizard Plus SV Minipreps DNA Purification System (Promega) was used according to the manufacturer's instructions in order to obtain small amounts of pure plasmidic DNA. The final pellet was resuspended in 100µl of Nuclease-Free water. The quality of extracted plasmid DNA was verified on 0.8% agarose gels. The DNA was stored at -20°C. Routinely 3µl of such DNA preparation were taken for the restriction enzyme digestion and 20µl were taken for sequence analysis (Macrogen).

2.6.2 Medium scale preparation of plasmid DNA from bacterial cultures

For medium-scale preparations (Midiprep) of plasmid DNA that was necessary for the transfection experiments, JETSTAR Plasmid Purification Kit (Genomed) was used according to the manufacturer's instructions. In order to get a good amount of plasmid DNA, an overnight bacterial culture of 50ml of LB medium was used.

2.7 Enzymatic modification of DNA

2.7.1 Restriction enzymes

Restriction enzymes were purchased from New England Biolabs Inc. (NEB). All buffers were also supplied by the same company and were used according with the manufacturer's instructions. In alternative we also used 0.5X, 1X or 2X concentration of 10X OPA buffer (100mM Tris-acetate, pH 7.5, 100mM magnesium acetate, 500mM potassium acetate). For analytical digests 300ng of DNA were digested in a volume of 20µl containing the appropriate units (Us) of the restriction enzyme per µg of DNA. The digest was incubated 2 hours at the optimal temperature required by the enzyme used.

Preparative digestions of vectors and inserts were made of 3µg DNA using the appropriate condition needed by the restriction enzyme in 50µl reaction volume. The enzymatic activity was stopped by heat inactivation.

2.7.2 DNA Polymerase I, Large (Klenow) Fragment T4 Polynucleotide Kinase

These enzymes, provided from New England BioLabs Inc., were used to treat PCR products for blunt-end ligation during construction of recombinant plasmids.
The Large Fragment of DNA Polymerase I (Klenow) is a proteolytic product of *E. coli* DNA Polymerase I, which retains polymerization and $3' \rightarrow 5'$ exonuclease activity, but has lost $5' \rightarrow 3'$ exonuclease activity. This was useful for digesting specific residues added by *Taq* DNA Polymerase at the 3' terminus to create a compatible end for ligation. Briefly the DNA was dissolved in 1X NEBuffer 2.1 and supplemented with 25µM dNTPs. 1U of Klenow per µg DNA was added and the mixture (final volume 20µl) was incubated 10 minutes at RT. The reaction was inactivated by heating at 70°C for 20 minutes.
T4 Polynucleotide Kinase catalyzes the transfer and exchange of phosphate from ATP to the 5'-hydroxyl terminus of ds and ssDNA and RNA. It was useful for the addition of 5'-phosphate to PCR products to allow subsequent ligation. The proper Us of Kinase, its reaction buffer (70mM Tris-HCl, 10mM $MgCl_2$, 5mM Dithiothreitol (DTT), pH 7.6) and ATP 10mM were added to the DNA and incubated at 37°C for 30 minutes. The enzyme was inactivated by incubation at 65°C for 20 minutes.

2.7.3 T4 DNA Ligase

T4 DNA Ligase, provided from Promega, catalyzes the joining of two strands of DNA between the 5'-phosphate and the 3'-hydroxyl groups of adjacent nucleotides in either a cohesive-ended or blunt-ended configuration. 20-40ng of linearized vector were ligated with a 5-10 fold molar excess of insert in a total volume of 20µl containing 1X ligase buffer and 1U of enzyme. Reaction was carried out at RT for 3 hours.

2.7.4 Alkaline Phosphatase, Calf Intestinal (CIP)

Calf Intestinal Phosphatase (CIP), provided from New England BioLabs Inc., catalyzes the removal of 5'-phosphate groups from DNA and RNA. Since CIP-treated fragments lack the 5'-phosphoryl termini required by ligases, they cannot self-ligate. This property can be used to decrease the vector background in cloning strategies. The standard reaction was carried out in a final volume of 50µl using 0.5U of enzyme per 1µg DNA at 37°C for 30 minutes. The enzyme was inactivated by incubation at 85°C for 15 minutes.

2.8 Agarose gel electrophoresis of nucleic acids

DNA sample were size fractionated by electrophoresis in agarose gels ranging in concentration from 0.8% w/v (large fragments) to 3% w/v (small fragments). The gels contained ethidium bromide (0.5µg/ml) in 1X TBE solution. Sample containing 1X DNA Loading Buffer (0.4% bromophenol blue, 60% glycerol, water) were loaded into submerged wells. Horizontal gels were used for fast analysis of DNA restriction enzyme digestions, estimation of DNA concentration, or DNA fragment separation prior to elution from the gel.
A fast analysis of RNA samples was obtained by running samples on 0.8% agarose gels.
The gels were electrophoresed at 50-90mA in 1X TBE running buffer for a time depending on the expected fragment length and gel concentration. Sample was visualized by UV transillumination and the result recorded by digital photography.

2.9 Elution and purification of DNA fragments from agarose gels

The following protocol was applied for purification of both vectors and inserts of DNA for cloning strategies. The DNA samples were electrophoresed onto an agarose gel as previously described. The DNA was visualized with UV light and the DNA fragment of interest was excised from the gel and the QIAquick Gel Extraction Kit (Qiagen) was used for extraction according to the manufacturer's instruction. 30-50µl of DNA was obtained: the amount of recovered DNA was estimated by UV fluorescence of intercalated ethidium bromide in agarose gel electrophoresis.

2.10 Amplification of selected DNA fragments

The polymerase chain reaction (PCR) was performed using both genomic and plasmid DNA as templates and following the basic protocol of Roche or NEB Taq DNA polymerase. The volume of the reaction was 50µl and comprised: 1X Taq buffer, dNTPs mix (100µM each), oligonucleotide primers (100nM each), Taq DNA polymerase (2.5U) and 100ng of genomic DNA or 0.1ng of plasmid DNA. The synthetic DNA oligonucleotides used for PCR amplification were purchased from Sigma-Aldrich and from Integrated DNA Technologies.
The standard amplification conditions were the following: 94°C for 5' (initial denaturation), 94°C for 45" (denaturation), 56°C for 45" (annealing), 72°C for 45" (extension) for 28-35 cycles and 72°C for 10' (final extension). PCRs were optimized to be in the exponential phase of amplification and products were fractionated on agarose gel. The PCR reactions were performed on a Gene Amp PCR System (Applied Biosystem).

2.11 Sequence analysis for cloning purpose

Sequence analyses of plasmid DNA were performed in Macrogen Europe (Amsterdam, Netherlands). 20µl of DNA samples (100ng/µl) in a 1.5ml tubes were shipped to Macrogen. The primers used for sequencing are either primers specific for plasmid vectors already available in the Macrogen Europe facility collection (universal primers) or primers binding on the cloned sequence, that were shipped with DNA samples (2µl of primer at concentration of 5pmol/µl for each sample). The result of sequencing were analyzed using the FinchTV and DNA Strider programs.

2.12 Hybrid minigene constructs

Hybrid minigenes are useful construct tools to study *cis* and *trans*-acting elements that influence splicing process, to understand the cell-specific splicing pattern, to identify exonic and intronic elements that enhance or silence splicing and to determine whether a specific mutation can compromise the splicing and to establish the role of the splice sites in the exon recognition [227]. This system, after transient transfection and RNA analysis, allows us to study the splicing outcome.

2.12.1 pcDNA3pY7 miR-34b hybrid minigenes

A DNA fragment of 196 bp containing the human pri-miR-34b hairpin, located on the junction between intron 1 and exon 2, with 18 bp of upstream intronic flanking region and 97 bp of exon 2 was cloned in the unique XhoI-XbaI cassette of the pcDNA3pY7 minigene [228] to generate pcDNA3pY7 miR34b construct, that is the wild type (WT) construct.
The creation of minigene has been done through PCR amplification of the genomic region of interest, treatment with Klenow-Kinase enzymes as described previously, and cloning in pUC19 vector into the restriction site SmaI. Following sequencing, in order to check the nucleotide sequence to be correct, the subcloning in the final minigene was performed using the enzyme restriction sites XhoI and XbaI.
The oligonucleotides used to create this new hybrid minigene are the following:

2310 XhoI dir	5'-TATACTCGAGCCGCGGGTGCCCGGTGCT-3'
2505 XbaI rev	5'-TATCTAGACCACGCCGACGCCGCGCT-3'

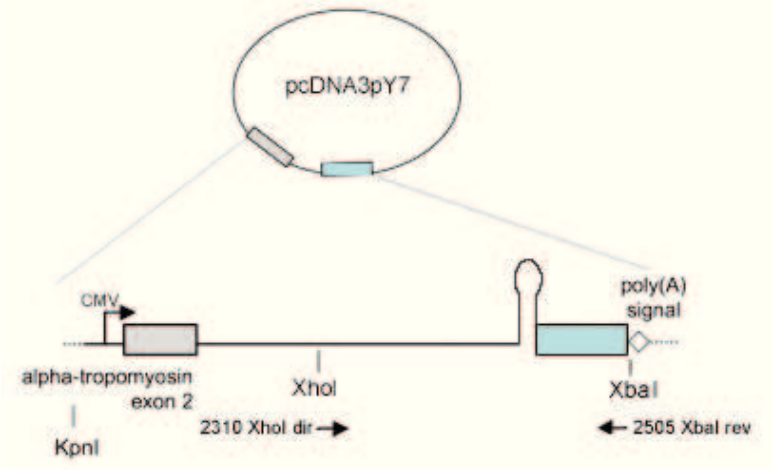


Figure 2.1: **Schematic representation of pcDNA3pY7 miR-34b minigene**
A DNA fragment of 196 bp containing the human pri-miR-34b hairpin, located on the junction between intron 1 and exon 2, with intronic flanking regions (thin line) and 97 bp of exon 2 (orange), was cloned in the unique XhoI-XbaI cassette of the pcDNA3pY7 minigene. The CMV promoter, the poly(A) site (rhombus), the primers used for amplification (black arrows) and the cloning restriction sites are indicated.

2.12.1.1 pcDNA3pY7 miR-34b mutated constructs

To generate the mutated constructs, the sequence of interest was amplified, using standard PCR conditions, with mutated primers built either for direct or overlapping PCR, according to the position of the mutation. The pcDNA3pY7 miR-34b as PCR template. The oligonucleotides used for generation of the mutated constructs are the following:

3'ss mut dir	5'-CTGCCATCAAAACAAACCACGGCATCAC-3'
3'ss mut rev	5'-GTGATGCCGTGGTTTGTTTTGATGGCAG-3'
ΔC2 dir	5'-GTACTGTGGTGACTCCACTGCCATCAAAACAA-3'
ΔC2 rev	5'-GGCAGTGGAGTCACCACAGTACAATCAGCTAA-3'
-2U>A dir	5'-GCAGTGTCATTAGCTGATTGTACTGTGGTGGTTACAATCACAAACTCCACTGCCATCAAAACAAGG-3'
-2U>A rev	5'-GCAGTGGAGTTAGTGATTGTAACCACCACAGTACAATCAGCTAATGACACTGCCTACAAACCG-3'
Δ2457-2482 dir	5'-AGAAGACGCCGGCTCGGGCATGAGAAGCGCGGCGTCGGCGTGG-3'
Δ2457-2482 rev	5'-CCACGCCGACGCCGCGCTTCTCATGCCCGAGCCGGCGTCTTCT-3'
2482 XbaI rev	5'-TATCTAGAGGCATCTTCTCTCGAAGGCTG-3'
ESEmut dir	5'-GCCCGCAGCCTTCGCGCGTATATGCCTGAGAAGCGCGGCGT-3'
ESEmut rev	5'-ACGCCGCGCTTCTCAGGCATATACGCGCGAAGGCTGCGGGC-3'

2.12.2 pFAN-COL17A1 and pFAN-LAMB3 hybrid minigenes

A DNA fragment of 436 bp containing 36 bp of COL17A1 exon 29, with 200 bp of upstream and downstream intronic flanking regions was cloned in the unique NdeI site of the pFAN minigene [229] to generate pFAN-COL17A1 wild type construct. For the generation of the pFAN-LAMB3 wild type construct a DNA fragment of 505 bp containing 221 bp of LAMB3 exon 15, with 84 bp of upstream intron and 200 bp of downstream intronic region was cloned in the unique NdeI site of the pFAN minigene. The oligonucleotides used to create this new hybrid minigenes are the following:

COL17A1 int28 NdeI dir	5'-TACATATGAAAAGAGACAATTGCTGGTT-3'
COL17A1 int29 NdeI rev	5'-TACATATGCACATCTGAACTGATAGTTT-3'
LAMB3 int14 NdeI dir	5'-TACATATGGCATGGGCACTGCCGTGGTGGGC-3'
LAMB3 int14 NdeI rev	5'-TACATATGCTCCATGAGAGCTAAGGACCAG-3'

The oligonucleotides used for generation of the mutated pFAN-COL17A1 5'ss mut and pFAN-LAMB3 3'ss mut constructs are the following:

COL17A1 5'ss mut dir	5'-ACCAAGGCCCAAGAGTTTGGTCACT-3'
COL17A1 5'ss mut rev	5'-AGTGACCAAACTCTTGGGCCTTGGT-3'
LAMB3 3'ss mut dir	5'-GCTTCTCGCCACAACGGCCTTCCGGATGCTGTGCACTGCCTAC-3'
LAMB3 3'ss mut rev	5'-GTAGGCAGTGCACAGCATCCGGAAGGCCGTTGTGGCGAGAAGC-3'

2.12.3 pCI-neo-PISD and pcDNA3pY7-KRT15 minigenes for small RNA-seq data validation

A DNA fragment of 890 bp containing 139 bp of PISD exon 4, with upstream and downstream intronic and exonic flanking regions was cloned in the unique XhoI-ApaI cassette of the pcDNA3pY7 minigene. A DNA fragment of 558 bp containing 83 bp of KRT15 exon 2, with upstream and downstream intronic and exonic flanking regions was cloned in the unique XhoI-NotI cassette of the pCI-neo minigene. The oligonucleotides used to create this new hybrid minigenes are the following:

KRT15 int1 XhoI dir	5'-TACTCGAGTGGGACAAAACAAGAAATGG-3'
KRT15 ex3 ApaI rev	5'-TAGGGCCCCTCTTCGTGGTTCTTCTTCA-3'
PISD ex3 XhoI dir	5'-GGCTCGAGGTGGCTTTGTACAAGTCAGT-3'
PISD ex5 NotI rev	5'-GGGCGGCCGCCTGGGAAGTGGCGCCGGTGG-3'

2.13 Eukaryotic cell lines

The eukaryotic cell lines used for the experiments are: HeLa (human cervical carcinoma), HEK293T (human embryonic kidney), Hep3B (human hepatocellular carcinoma), CFPAC-1 (human cystic fibrosis pancreatic adenocarcinoma), NHEK (normal human epidermal keratinocytes), HaCaT (immortal human keratinocyte), BC-1 (human B-cell lymphoma), and MEC-1 (human B-chronic lymphocytic leukemia) cells.

2.13.0.1 Maintenance of cells in culture

HeLa and **HEK293T cells** were grown in Dulbecco's Modified Eagle Medium (DMEM with glutamine, sodium pyruvate, pyriodoxine and 4.5g/l glucose) supplemented with 10% heat-inactivated fetal bovine serum (FBS - EuroClone) and antibiotic antimycotic (Sigma) according to the manufacturer's instruction. A standard 100mm dish containing a confluent monolayer

of cells was washed twice with 1X PBS solution, in order to remove all the medium residues as well as the dead cells, treated with 1ml Trypsin (PBS containing 0.045mM EDTA and 0,1% trypsin) and incubated at 37°C for 2 minutes or until cells were completely detached. After adding 5ml of medium, to block trypsin, cells were precipitated by centrifugation (1000 rpm for 5 minutes) and resuspended in pre-warmed medium. A sub-cultivation ratio of 1:6 to 1:8 was used. 1ml of cell dilution was added to 9ml fresh medium and plated in a new 100mm dish.

CFPAC-1 cells were grown Iscove's Modified Dulbecco's Medium (IMDM) supplemented with 10% FBS and antibiotic antimycotic.

Hep3B cells were grown in minimal essential medium (MEM) supplemented with 10% FBS and antibiotic antimycotic.

HaCaT cells were grown in DMEM medium supplemented with 10% FBS and antibiotic antimycotic into T25 flasks. Sub-culturing of the cells was made as described above, but for this cell type the trypsinization times is much longer (10 minutes).

MEC-1 (kindly provided by Dott. Dimitar Efremov - ICGEB, Monterotondo - Rome) and **BC-1 cells** are cells growing in suspension, that were cultured in RPMI-1640 medium supplemented with 10% FBS and antibiotic antimycotic. Cell suspension was taked out from the flask, putted in a 15ml falcon and precipitated by centrifugation. The pellet was washed twice with 1X PBS solution and resuspended in pre-warmed medium. A sub-cultivation ratio of 1:2 to 1:3 was used. 1ml of this cell dilution was added to 5ml fresh medium and plated in a new T25 flask.

2.13.0.2 NHEK growth and differentiation

Normal human primary epidermal keratinocytes (kindly provided by Dott. Daniele Castiglia - Istituto Dermopatico dell'Immacolata, Rome) were isolated and cultured as described [230]. Keratinocytes were seeded at a concentration of 1-1.5x10^4 cells/cm^2 in serum-free Keratinocyte Growth Medium (KGM, Invitrogen) containing 0.15mM Ca^{2+}, cultured till confluence, and then induced to differentiate by culturing in KGM containing 1.2mM Ca^{2+} for 5 days.

2.14 Transfection of recombinant DNA

3x10^5 HeLa cells were plated as described above into 6-well cell culture dishes in order to reach a final confluence of 40-70%. The plasmid DNA used for transfection was prepared with

JETSTAR Purification Kit as previously described.

Transfection of HeLa cells was performed by applying Effectene Transfection Reagent (Qiagen) according to manufacturer's instructions. For 6-well cell culture plates, 500ng of plasmid DNA were first mixed with 150µl of EC buffer and 4µl of Enhancer and the mixture was incubated at RT for 5 minutes to allow the condensation of DNA. 5µl of Effectene Reagent were then added to the mixture and incubated for 10 minutes to allow Effectene-DNA complexes to form. After the addition of 500µl of complete growth medium, the mixture was added to the cells in 1.5ml of the same medium and incubated at 37°C. After 24 hours, cells were harvested and subjected to further analyses. Each transfection experiment was repeated at least three times in double.

2.15 Co-transfection of splicing factors

For the different splicing factors co-transfection experiments, 500ng of the minigenes have been transfected with 500ng of the expressing plasmid. The splicing factors used for the co-transfections are SRSF1 (SF2/ASF), SRSF3 (SRp20), SRSF4 (SRp75), SRSF7 (9G8), Tra2β, TIA1, PTB1, U2AF35, U2AF65, hnRNP A1, hnRNP A2, and hnRNP H coming from pCG expressing plasmid.

2.16 RNA preparation from cultured cells

Cultured cells were washed twice with 1X PBS and then 750µl of RNA TRI Reagent (Ambion) was added. After chloroform extraction, the RNA-containing supernatant was collected and precipitated by addition of 1 volume of isopropanol. The pellet was rinsed in 70% ethanol. The final pellet was resuspended in H_2O and stored at -80°C. The RNA quality was checked by electrophoresis on a 0.8% agarose gel.

In the case of DNA contamination, the extracted RNA was treated with RNase-free DNase (Promega) in 1X DNase Buffer, according to manufacturer's instructions. The DNA digestion was performed at 37°C for 30 minutes and the reaction was stopped by adding DNase Stop Solution and incubating at 65°C for 10 minutes. RNA was then purified using RNeasy Mini Kit - RNA Cleanup Protocol (Qiagen) according to the manufacturer's instructions.

2.17 Estimation of nucleic acid concentration

As both DNA and RNA molecules absorb UV light, this feature is used for measuring the concentration of nucleic acids with NanoDrop spectrophotometer. The nitrogenous bases in nucleotides have an absorption peak at 260nm, so an optical density of 1.0 at this wavelength is usually taken to be equivalent to a concentration of 50μg/ml for dsDNA, 40μg/ml for ssDNA and RNA and approximately 20μg/ml for ss oligonucleotide sample. The ratio of the absorbance at 260nm and 280nm is a measure of the purity of a sample: this ratio should be around 1.8 for pure DNA sample and 2.0 for RNA sample. If the ratio is appreciably lower, it may indicate the presence of protein, phenol or other contaminants that absorb strongly at or near 280nm.

2.18 The mRNA functional splicing analysis

2.18.1 cDNA synthesis

The first-strand cDNA synthesis was performed with the M-MLV Reverse Transcriptase Kit (Invitrogen) following manufacturer's instructions. 3-5μg of total RNA extracted from cells were mixed with 2μl of random primers (100ng/μl, Invitrogen) and diluited in water to the final volume of 12μl. To denature the RNA, the mixture was put at 94°C for 2 minutes and quick chilled on ice. After denaturation 6μl 5X First-Strand Buffer(250mM Tris-HCl (pH 8.3 at RT), 375mM KCl, 15mM $MgCl_2$), 3μl 0.1M DTT, 3μl dNTPs, 0.5μl M-MLV RT were added to the reaction. The final mixture was then incubated at 37°C for 90 minutes. 3μl of the reaction were used as a template for the PCR analysis.

2.18.2 PCR analysis

PCR reactions were performed using previously described protocol in a final volume of 25μl. Oligonucleotides used for PCR analysis, specific for the minigene system used or for the endogenous genes of interest, are listed below. The results of all the transfections are the representative of at least three independent experiments. PCR products were resolved by agarose gel electrophoresis. Quantification of exon inclusion, exon skipping and intron retention was performed using ImageJ 1.44o (http://imagej.nih.gov/ij).

pY7ex2 dir	5'-TACAAGGCTTGTCGAGGAGGACATC-3'
2505 XbaI rev	5'-TATCTAGAGTCCTGCGCAGCCTGC-3'
2482 XbaI rev	5'-TATCTAGAGGCATCTTCTCTCGAAGGCTG-3'
ASH1L ex15 dir	5'-GATGTTTGTCGTCTACGAAAGGC-3'
ASH1L ex17 rev	5'-GAAGCAGCAAGTCATCTCGGAGCA-3'
DNM2 ex1 dir	5'-CATGGGCAACCGCGGGATGGAAG-3'
DNM2 ex3 rev	5'-GATGCCTTTGTTGGTCCCCGTGAC-3'
SFRS2 ex1 dir	5'-GGACCGCTACACCAAGGAGT-3'
SFRS2 ex3 rev	5'-CTTCGATGGACTATGTGGTC-3'
NOP56 ex1 dir	5'-GAGCGCGCTAGCCGCATTGC-3'
NOP56 ex5 rev	5'-TCACCAGATTGTGGAAGTGC -3'
WHSC2 ex3 dir	5'-AGTGTGAAGCGTCTGCCATGCTGCCACTG-3'
WHSC2 ex7 rev	5'-GCGTGGACACCAGGGCCGGCTGCGTAGT-3'
CIT ex31 dir	5'-CTGCCCACCGCAAAGCAACG-3'
CIT ex35 rev	5'-CTGCTTTGGCTGTATTTGCG-3'
PYROXD2 ex5 dir	5'-AGAAGCAGATCGCCCAGTTCTCCCAG-3'
PYROXD2 ex9 rev	5'-TCCATGTGTGGTGGCTGAGCTTGCGA-3'
PISD ex3 dir	5'-CAACCTCAGCGAGTTCTTCC-3'
PISD ex5 rev	5'-CTGGGAAGTGGCGCCGGTGG-3'
KRT ex1 dir	5'-AGCCCAGAATGCGACTACAG-3'
KRT ex3 rev	5'-GTCAGCTCATCCAGGACTCG-3'
NRD1 ex2 dir	5'-GATGACGATGAAGAGGGTTTTGAT-3'
NRD1 ex6 rev	5'-CTTGAAGTACTTCCTCTGGACATC-3'
PISD ex3 XhoI dir	5'-GGCTCGAGGTGGCTTTGTACAAGTCAGT-3'
pCI rev	5'-GTATCTTATCATGTCTGCTCG-3'
KRT15 ex3 ApaI rev	5'-TAGGGCCCCTCTTCGTGGTTCTTCTTCA-3'
DUSP11 ex5 dir	5'-GATTTGCCAGAAACTGTTCCTTA-3'
DUSP11 ex7 rev	5'-TCTGAAGGTCTTCAATGTAGTTTTGT-3'
RBM5 ex15 dir	5'-CGGCTGTAGTGTCCCAGAGT-3'
RBM5 ex17 rev	5'-TTGCGAGTTGGGGTCATAAT-3'
COL17A1 ex28 dir	5'-CGAGCCTGGCATGAGAGGTTTGCCT-3'
COL17A1 ex31 rev	5'-CCTGAGGTCCCTGGGGTCCTGTGAGAC-3'
LAMB3 ex14 dir	5'-GTTGTCCCTTCCGAGAGACCTGGAG-3'
LAMB3 ex16 rev	5'-CGCTGGAGCTGGGCATTGAAGC-3'
α2,3 dir	5'-CAACTTCAAGCTCCTAAGCCACTGC-3'
pFAN 904 rev	5'-ATCAGCATTCTCTGAAGTACATCAAC-3'

2.19 Real Time Quantitative PCR analysis

2.19.1 SYBR Green

SYBR Green-based quantitative PCR (qPCR) was performed using the iQ SYBR Green Supermix (Bio-Rad) in a CFX96 Real-Time PCR system (Bio-Rad), following manufacturer's instructions. Briefly, 100ng of cDNA where mixed with the master mix and the forward and reverse primers and filled with water to a final volume of 15µl. Primers, designed to target an amplicon size of 100-200 bp, are listed below. Three replicates per RT reaction were performed. The thermal cycling protocol was: 95°C for 10' (initial denaturation), 95°C for 15" (denaturation), 60°C for 60" (annealing/extension + plate read) for 40 cycles and 65-95°C, 0.5°C increment 1"/step (Melt Curve analysis).

U6 dir	5'-GCTTCGGCAGCACATATACTAA-3'
U6 rev	5'-AAAATATGGAACGCTTCACG-3'
GAPDH dir	5'-GACAGTCAGCCGCATCTTCT-3'
GAPDH rev	5'-TTAAAAGCAGCCCTGGTGAC-3'
Drosha dir	5'-CATGTCACAGAATGTCGTTCCA-3'
Drosha rev	5'-GGGTGAAGCAGCCTCAGATTT-3'
Involucrin dir	5'-GAAGCACCACAAAGGGAGAAGTATTGC-3'
Involucrin rev	5'-CCACTGCACCTCCTGCTTCT-3'
RPL13A dir	5'-CTCAAGGTCGTGCGTCTGAA-3'
RPL13A rev	5'-TGGCTGTCACTGCCTGGTACT-3'
TBP dir	5'-TCAAACCCAGAATTGTTCTCCTTAT-3'
TBP rev	5'-CCTGAATCCCTTTAGAATAGGGTAGA-3'
RPLP0 dir	5'-CAGATTGGCTACCCAACTGTT-3'
RPLP0 rev	5'-GGGAAGGTGTAATCCGTCTCC-3'
LAMB3 ex15 dir	5'-TGAGGCTGGAGATGTCTTCGTTGC-3'
LAMB3 ex16 rev	5'-GCCACAGGCTGTGCCATTGTCTTG-3'
COL17A1 ex28 dir	5'-CGAGCCTGGCATGAGAGGTTTGCCT-3'
COL17A1 ex31 rev	5'-CCTGAGGTCCCTGGGGTCCTGTGAGAC-3'

2.19.2 TaqMan microRNA assay

2.19.2.1 TaqMan miRNA Reverse Transcription

Total RNA was reverse transcribed using M-MLV Reverse Transcriptase. For each 15µl RT reaction 1-10ng of total RNA were mixed with 7µl of RT master mix and 3µl of RT primer (Applied Biosystems) specific for each miRNA. The RT master mix was composed of 1.50µl dNTPs 10mM, 3µl 5X First-Strand Buffer, 0.25µl M-MLV RT 200U/µl, 0.19 RNase Inhibitor, 20U/µl and 2.06µl of water. The final mixture was incubated for 5' on ice, followed by 30' at 16°C, 30' at 42°C and 5' at 85°C.

2.19.2.2 TaqMan miRNA qPCR

PCR reaction was performed in triplicate on a CFX96 Real-Time PCR system. The final volume of the reaction was 20µl: 10µl of TaqMan 2X Universal PCR Master Mix, No AmpErase UNG and 7.67µl of water were mixed with 1.33µl RT reaction product (dilution 1:10) and 1µl 20X TaqMan MicroRNA assay, specific for each miRNA. Thermal cycling parameters are the following: 95°C for 10' (AmpliTaq Gold enzyme activation), 95°C for 15" (denaturation) and 60°C for 60" (annealing/extension + plate read) for 40 cycles. The predesigned TaqMan assays (Life Technologies) are the following:

miRNA	**Assay ID**	**miRNA**	**Assay ID**
miR-943	002188	miR-936	002179
miR-1287	002828	miR-34b	002102
miR-1178	002777	miR-636	002088
miR-6510	477465_mat	miR-1292	002824
miR-7109	466424_mat	miR-4260	244117_mat
miR-761	002030	miR-555	001523
miR-638	001582	miR-711	241090_mat
miR-26b	000407		

Results of SYBR Green and TaqMan assays were normalized to different endogenous controls. Fold changes were determined using the $2^{-\Delta\Delta Ct}$ method.

2.20 Small interfering RNA (siRNA) transfection

The sense strand of RNAi oligos (Dharmacon), which were used for silencing the expression of different target genes, were the following:

Drosha	5'-CGAGUAGGCUUCGUGACUU-3'
DGCR8	5'-CAUCGGACAAGAGUGUGAU-3'
SF3b1	5'-CGCCAAGACUCACGAAGAU-3'
UBC	5'-GUGAAGACCCUGACUGGUA-3'
LUCIFERASE	5'-UAAGGCUAUGAAGAGAUAC-3'
Non-Targeting siRNA #1	5'-UGGUUUACAUGUCGACUAA-3'

$2.5x10^5$ cells were plated in 60mm plates to achieve 40-50% confluence the following day when siRNA transfection was carried out preparing two reaction mixtures. The first comprised 10µl of Oligofectamine Reagent (Invitrogen) combined with 20µl of Opti-MEM medium (Invitrogen), whereas the second one consisted of 10µl of 40µM siRNA duplex oligos diluted in a final volume of 370µl of Opti-MEM. The two mixtures were combined, left at RT for 20 minutes and finally added to the cells, which were maintained in 1.6ml of Opti-MEM. After 24 hours the second treatment with siRNA was performed as described above. 6-8 hours later Opti-MEM was exchanged with DMEM and the cells were transfected with the minigene of interest using Qiagen Effectene transfection reagents. 1µg plasmid DNA was mixed with 8µl of Enhancer and 10µl of Effectene Reagent using previously described protocol. After addition of 500µl of complete DMEM medium, the mixture was added to the cells in 4µl of the same medium and incubated at 37°C. After 24-72 hours cells were collected and divided in two fractions for protein and RNA extractions. RT-PCR from total RNA was performed as described above. The whole protein extracts were obtained by cell resuspension in 1X lysis buffer (100mM KCl, 200µM EDTA pH 8.0, 20mM Hepes pH 7.5, 0.5mM DTT, 1% NP-40, 6% glycerol, 1X proteinase inhibitor). Each siRNA experiment was repeated at least three times.

SF3b1 siRNA was introduced into MEC-1 cells with the Nucleofector system (Amaxa, Lonza Cologne). Briefly, $0.6x10^7$ cells were resuspended in 100µl of Cell Line Nucleofector Solution L and mixed with 2.5µg of siRNA. Nucleofection was performed on an Amaxa Nucleofector II device using the C-005 program. 72 hours after nucleofection the cells were collected and processed as described above.

2.21 miRNA mimics reverse transfection and cell proliferation assay

miRNA mimics (miRIDIAN miRNA mimics, GE Dharmacon) are double-stranded RNA molecules designed to mimic the function of endogenous, mature microRNAs.
miRNA mimics used to suppplement miRNA activity were the following:

hsa-miR-936 mimic	5'-ACAGUAGAGGGAGGAAUCGCAG-3'
hsa-miR-4260 mimic	5'-CUUGGGGCAUGGAGUCCCA-3'
hsa-miR-711 mimic	5'-GGGACCCAGGGAGAGACGUAAG-3'

miRNA mimics were transfected into HaCaT cells using a standard reverse transfection protocol, at a final miRNA concentration of 50nM for 72 hours. Briefly, the transfection reagent (Lipofectamine RNAiMAX, Life Technologies) was diluted in Opti-MEM medium and left at RT for 5 minutes. The mix was added to the miRNAs already placed into 96-well plates. 30' after, $1.0x10^3$ cells were seeded per well. The final volume was of 150µl. Transfection was optimized using a toxic siRNA targeting ubiquitin C (UBC), which results in the death of all cells. NT siRNA #1 is used as negative control. The experiment was performed in triplicate. 60h after transfection, the medium was replaced with medium containing 5µM 5-ethynyl-2'-deoxyuridine (EdU, Thermo Fisher Scientific), a nucleoside analog of thymidine, which was incorporated into DNA during active DNA synthesis. 12 hours after EdU addition, cells were washed in PBS, fixed in 4% PFA (Paraformaldehyde) for 20', permeabilized with 0.5% Triton X-100 in PBS solution for 20'. EdU detection is carried out using Click-iT EdU Alexa Fluor 488 Imaging Kit (Thermo Fisher Scientific), according to the manufacturer's instructions. Nuclei were counterstained with Hoechst 33342. Cells were processed for immunofluorescence using the Molecular Devices ImageXpress Micro High-Content Screening Microscope at the ICGEB High-Throughput Screening Facility (http://www.icgeb.org/high-throughput-screening.html).

2.22 Small RNA sequencing and bioinformatic analysis

Small RNA sequencing (small RNA-seq) was performed by IGA Technology Services (Udine, Italy). 2µg of total RNAs from MEC-1 cells, at a minimum concentration of 200ng/µl and with a 260/280 ratio >1.8, whose RNA quality was checked on a 0.8% agarose gel-electrophoresis and on an Agilent 2100 Bioanalyzer, were processed with the llumina HiSeq2000 (run lenght 1x50

bp) platform. One lane in 12-plex was run obtaining 10 millions of single-reads per sample, 50 bp long. Four control sample treated with Luciferase siRNA and four sample treated with the SF3b1 siRNA, derived from four independent experiments were send to sequencing.
Real time image analysis and base calling were performed on the the HiSeq2000 instrument using the HiSeq Sequencing Control Software (HCS). CASAVA software (version 1.8.2) was used for de-multiplexing and production of FASTQ sequence files. FASTQ raw sequence files were subsequently quality checked using FASTQC software (version 0.11.3 http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc). We used the Bowtie 2 tool for aligning sequencing reads to long reference sequences and Bioconductor based Expression analysis. Briefly, reads were trimmed by removing the index and adaptor sequences. Subsequently, sequences including adaptor dimers, mitochondrial or ribosomal sequences were discarded. The resulting set of trimmed reads were then mapped against Human GRCh38/hg38 and to known mature miRNAs (miRBase release 21) using Bowtie 2 software [231], due to the tool accuracy of aligning reads of about 50 up to 100 bases to long genomes. Rounded Gene counts were used as input to perform differential gene expression analysis using Bioconductor package DESeq2 (version 1.4.5) [232], estimating the per-gene Negative binomial distribution dispersion parameter. To detect outlier data after normalization we used the R package arrayQualityMetrics [233] and before testing differential gene expression we dropped all genes with low normalized mean counts to improve testing power while maintaining type I error rates. Estimated false discovery rate (FDR) values for each gene was adjusted using the Benjamini-Hochberg method. Features with adjusted FDR value <0.01 and absolute logarithmic base 2 fold change ($\log_2$FC) >1 were considered having a significant altered expression as previously reported. Raw count data were transformed to log2-counts per million (logCPM) before compute appropriate self-contained gene set test, in the sense defined by Goeman and Bühlmann [234], using the ROAST [235] gene set test. This analysis was useful to test for enrichment of a particular group of miRNAs amongst the differential expression results from a global analysis. A specific group of miRNAs is treated as a unit and a single FDR value is evaluated for the group rather than evaluating individual FDR values for individual miRNAs. Specifically ROAST tests whether any of the genes in the set are differentially expressed. For assessing differential expression, the empirical Bayes moderated t-statistics test is used to moderate the standard errors of the estimated $\log_2$FC.

2.23 Denaturing polyacrylamide gel electrophoresis (SDS-PAGE)

30μg of protein were mixed with the appropriate volume of 2X Protein loading buffer (the max volume that can be loaded is 55μl) and the mixture was incubated at 94°C for 4 minutes before loading on NuPAGE 4-12% Bis-Tris Gels (Invitrogen Pre-Cast Protein gels). The gels were run at 115mA in 1X NuPAGE MES SDS Running Buffer, until the desired bands are well separated. After running and separating on NuPAGE gels, proteins were electroblotted onto PVDF membrane according to standard protocols (Amersham Bioscences): sandwich was prepared with 3 pieces of paper, the gel, the membrane (previously activated with 30" in methanol, two times for 1' in H_2O, 5' in 1X Blotting buffer) and other 3 pieces of paper and everything was wet with buffer. After elimination of air bubbles the sandwich was closed and put in the blotting support filled with 1X Blotting buffer. The sandwich was blot at 200-250mA for 75-150 minutes depending on the proteins that have to be transferred.

2.24 Western blotting

The membrane was blocked by incubation in 5% dried milk in 0.1M PBS, 0.1% Tween20. Membranes were probed with different primary antibodies, then incubated with secondary antibodies conjugated with horseradish peroxidase (HRP) enzyme and the resulting immunoreactive bands were detected with enhanced chemiluminescence reagent (ECL, Amersham Biosciences).
The primary antibodies that were applied in western blotting analysis are: goat polyclonal anti-DGCR8 antibody (Santa Cruz; 1:500), rabbit polyclonal anti-Drosha antibody (Abcam; 1:625), rabbit monoclonal anti-SF3b1 antibody (Invitrogen; 1:1000), mouse monoclonal anti-Tubulin antibody (kindly provided by Dr. Muro; 1:5000). The secondary antibodies are respectively: polyclonal rabbit anti-goat immunoglobulins/HRP (Dako; 1:2000), polyclonal goat anti-rabbit immunoglobulins/HRP (Dako; 1:2000), polyclonal goat anti-mouse immunoglobulins/HRP (Dako; 1:2000).

2.25 Northern blot analysis of small RNAs

Approximately 30μg of RNA were mixed with Formamide Loading Dye (98% formamide, 10mM EDTA pH 8.0, Bromophenol Blue and Xylene Cyanol), denatured by incubation at 70°C, 3'

and loaded onto 13% polyacrylamide denaturing gels (5g Urea, 3ml 5X TBE, 5ml of Accugel 19:1 40% acrylamide:bis-acrylamide in a final volume of 15ml). The gel was stained with Ethidium Bromide to check the quality of RNA. The RNA was electroblotted to a nitrocellulose membrane (Amersham Biosciences) using a semi-dry apparatus (EuroClone) and 2mA per cm^2 of membrane for 1 hour. The RNA was UV cross-linked with 1200$\mu J/cm^2$. The membrane was pre-hybridized with Church buffer (0.25M Na_2HPO_4 pH 7.2, 1mM EDTA pH 8.0, 7% SDS in H_2O) at 37°C for 1 hour. In the meantime, the probe was prepared by kinase labelling of DNA using radioactive γ-32^P ATP. The kinase reaction included 10pmol of DNA probe, 1X T4 Polynuclotide Kinase Buffer, 20U of T4 Polynucleotide Kinase (New England Biolabs) and 15pmol of the radioactive ATP in a final volume of 50µl. The reaction was incubated 1 hour at 37°C and the DNA was precipitated with ethanol. The probe was added to the membrane and hybridization was performed overnight at 37°C (for miR-34b) or 2 hour at 37°C (for the U6). The membrane was washed with 2X SSC (300mM NaCl, 30mM sodium citrate dehydrate) with 0.1% SDS. Membrane was exposed to a cyclone screen overnight (miR-34b) or 30' (U6) and the screen was revealed using the cyclone reader. The probes used for the analysis were the following:

miR-34b probe	5'-ATGGCAGTGGAGTTAGTGATTG-3'
U6 probe	5'-ATATGGAACGCTTCACGAATT-3'

2.26 Bioinformatic analyses

All data are presented as mean ± standard deviation (SD). Statistical analysis was carried out using Prism Software (GraphPad). RNA secondary structure predictions were performed, under default parameters, using the freely available **Mfold server**. This web server uses mfold (version 3.5) by M.Zuker, 2003 (http://mfold.rna.albany.edu/?q=mfold/RNA-Folding-Form). Analysis of splice site strength was performed using the **Neural Network** program, lowering the threshold of splice site detection to 0.01 (http://www.fruitfly.org/seq_tools/splice.html). *In silico* prediction of protein binding sites was performed using the free access **ESEfinder** program (http://exon.cshl.edu/ESE/).

Chapter 3

Results

3.1 Identification of human Splice site Overlapping miRNAs

To identify pri-miRNA hairpins overlapping with the intron-exon junctions of non-coding and protein-coding host transcripts, I used an algorithm (designed by Dennis Prickett at CBM, Trieste, Italy) that compares the coordinates of pri-miRNAs, derived from miRBase database (release 21) along with the coordinates of putative splice sites derived from spliced expressed sequence tags (ESTs, GenBank - hg38 assembly). The algorithm compares the position of the 5' and 3'ss of the ESTs with the position of 5' and 3' ends of the miRNAs. I selected pri-miRNA hairpins with a distance from putative EST-derived splice sites <100bp. These candidates were carefully checked manually in order to pick real SO-miRNAs, excluding miRNAs found on non-annotated genes, miRNAs transcribed in the opposite direction than the host transcripts and annotated mirtrons. Through this manual annotation, I identified 52 SO-miRNAs: 32 pri-miRNA hairpins overlap with 3'ss, 18 with 5'ss. In 2 cases, both the 3' and the 5'ss are embedded in the pri-miRNA hairpin. 5 SO-miRNAs are encoded by putative open reading frames and 2 are localized within annotated non-coding transcripts, but the most abundant group, that comprises 45 SO-miRNAs (28 3' SO-miRNAs, 15 5' SO-miRNAs and 2 3' $\rightarrow$ 5' SO-miRNAs) are embedded in protein-coding genes (table 3.1).

3' SO-miRNAs

miRNA	Chr	Transcript
miR-205	1	MIR205HG (nc)
miR-5187	1	TOMM40L
miR-3605	1	PHC2
miR-3606	2	COL3A1
miR-6513	2	TMBIM1
miR-943	4	WHSC2
miR-5004	6	SYNGAP1
miR-6721	6	AGPAT1
miR-5090	7	LRWD1
miR-939	8	CPSF1
miR-6848	8	DGAT1
miR-6852	9	TLN1
miR-936	10	COL17A1
miR-1287	10	PYROXD2
miR-1307	10	USMG5
miR-34b	11	BC021736 (nc)
miR-1178	12	CIT
miR-6861	12	HECTD4
miR-5006	13	VWA8
miR-7855	14	SPTB
miR-5572	15	ARNT2
miR-636	17	SFRS2
miR-4315-2	17	PLEKHM1P
miR-5010	17	ATP6V0A1
miR-6510	17	KRT15
miR-365b	17	AC003101.1 (nc)
miR-3614	17	TRIM25
miR-4321	19	AMH
miR-5196	19	CD22
miR-133a-2	20	C20orf166
miR-1292	20	NOP56
miR-7109	22	PISD

5' SO-miRNAs

miRNA	Chr	Transcript
miR-4260	1	LAMB3
miR-761	1	NRD1
miR-555	1	ASH1L
miR-4793	3	CELSR3
miR-3655	5	IK
miR-4658	7	C7orf43
miR-202	10	MIR202HG (nc)
miR-611	11	TMEM258
miR-3656	11	TRAPPC4
miR-4721	16	TUFM
miR-638	19	DNM2
miR-1909	19	REX01
miR-3940	19	KSRP
miR-6789	19	PLEKHJ1
miR-6800	19	MED25
miR-4758	20	LAMA5
miR-8069-2	21	CH507-210P18.4 (nc)
miR-4761	22	COMT

3' → 5' SO-miRNAs

miRNA	Chr	Transcript
miR-711	3	COL7A1
miR-937	8	SCRIB

Table 3.1: **Human Splice site Overlapping miRNAs (SO-miRNAs)**
List of 52 SO-miRNAs identified through a bioinformatics analysis, followed by manual annotation. They are classified according to the localization of the hairpin on acceptor, donor or both splice sites in 3' SO-miRNAs, 5' SO-miRNAs and 3' → 5' SOmiRNAs, respectively. The genomic position and the correspondent host gene are indicated. nc: non-coding.

3.2 Relationship between spliceosome and MPC machineries in the processing of SO-miR-34b

3.2.1 pcDNA3pY7 miR-34b minigene

In order to explore the functional relationship between the spliceosome and the MPC, I initially focused on 3' SO-miR-34b, located on the intron 1 - exon 2 junction of the non-coding transcript BC021736, taking advantage of the pcDNA3pY7 miR-34b minigene, already available in laboratory. The minigene is composed of the α-tropomyosin exon 2 and the BC021736 exon 2, separated by a 176nt hybrid intronic sequence. It contains the human pri-miR-34b hairpin, which includes the 3'ss junction between intron 1 and exon 2 (fig. 3.1, upper panel).
HeLa cells were transiently transfected with the wild type (WT) pcDNA3pY7 miR-34b minigene and the spliced mRNA, treated with DNase to remove any possible DNA contamination, was analyzed through RT-PCR using pY7 ex2 dir and 2505 XbaI rev oligonucleotides which are specific for the hybrid minigene sequence. The PCR products detected upon transfection of the WT minigene showed the presence of two bands (fig. 3.1, lower panel): the upper one corresponds to intron retention (416 bp) and the lower one to intron splicing (240 bp). Quantitative analysis of band intensity showed that about 86% of the transcript is spliced. The identity of the bands was verified by direct sequencing after elution of the bands from the gel. This result shows that the sequence of miR-34b transcript inserted in the minigene is sufficient for the correct selection of the 3'ss and therefore this minigene represents a reliable tool to explore the *cis* and *trans*-acting elements involved in splicing regulation.

3.2.2 Acceptor splice site mutation completely abolishes splicing and increases mature miR-34b production

To evaluate the importance of the 3'ss in splicing process, I made a mutant minigene construct to disrupt the acceptor site. It contains two mutations in position -1 (-1G>A) and +1 (+1G>C), that disrupt the consensus 3'ss, plus an additional mutation in position +6 (+6G>C) to disrupt a downstream AG site. Transient transfection of the 3'ss mut minigene in HeLa cells completely impaired splicing and only the intron retention isoform was detected on the gel (fig. 3.2).

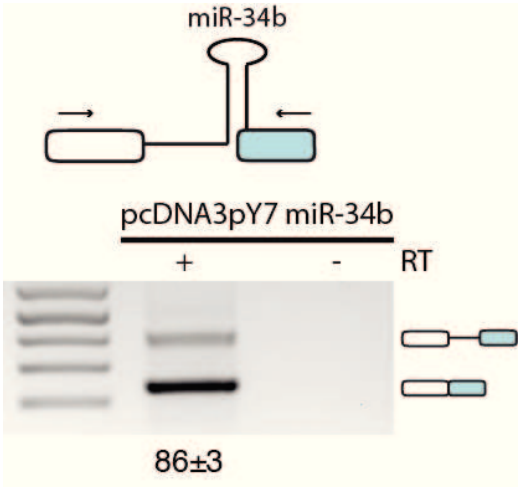


Figure 3.1: **pcDNA3pY7 miR-34b minigene splicing pattern**
[Upper panel] Schematic representation of the pcDNA3pY7 miR-34b minigene system. Boxes indicate exons, in particular light blue box is BC021736 exon 2. Thin line represents intron and pri-miR-34b hairpin. The black superimposed arrows represent primers used for PCR amplification: pY7 ex2 dir and 2505 XbaI rev. [Lower panel] Splicing pattern of pcDNA3pY7 miR-34b minigene after transfection in HeLa cells. The percentage of splicing expressed as means ±SD of three independent experiments is indicated below the gel. RT: reverse transcriptase.

To explore the effect of splicing on miRNA processing, I have evaluated the amount of mature miR-34b in the WT and 3'ss mut minigenes (Chiara Mattioli performed the Northern blot experiments). Northern blot showed that, in comparison to the WT construct, the mutation at the 3'ss, which caused a complete splicing inhibition, significantly increases the amount of miR-34b by 3.5 fold (fig. 3.2). These results suggest a direct competition between the spliceosome and the MPC machineries on the nascent pre-mRNA.

3.2.3 A consensus branch point sequence within pri-miR-34b hairpin is required for 3'ss selection

A canonical 3'ss is composed of three different elements: the branch point, the polypyrimidine tract and the AG acceptor site. In order to understand if this peculiar 3'ss, that lacks polypyrimidine tract, is regulated by a branch point sequence, I created the ΔC2 minigene: it contains a 13 bp deletion of the right part of the stem of the hairpin, that spans from 20 to 33 bp upstream of the AG dinucleotide, in accordance with the optimal distance between the BP and the acceptor site, that ranges from 20 to 40 bp [55].

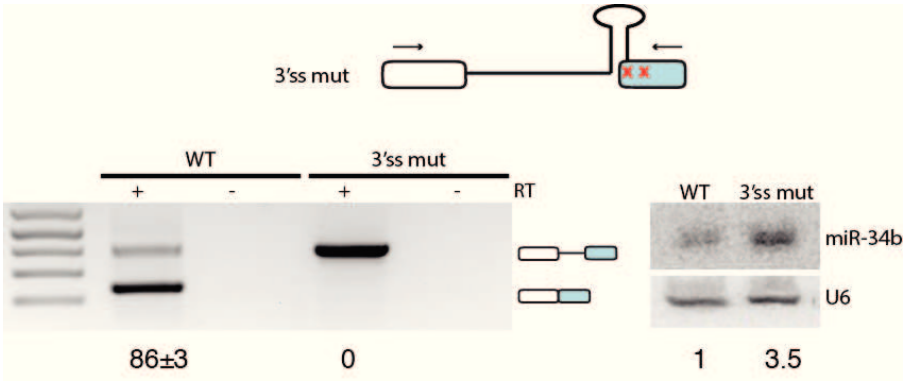


Figure 3.2: **3'ss mutation impairs splicing and increases miR-34b expression level**
[Upper panel] Schematic representation of pcDNA3pY7 miR-34b 3'ss mut minigene. Red x indicates the mutations. [Lower panel - left] Splicing pattern of pcDNA3pY7 miR-34b WT and 3'ss mut minigenes after transfection in HeLa cells. The identity of the bands is depicted on the right. Splicing percentage, expressed as means ±SD, is indicated below the gel. RT: reverse transcriptase. [Lower panel - right] Northern blot analysis of miR-34b and U6 snRNA in WT and 3'ss mut constructs. Numbers below the panel indicate the fold increase of miR-34b normalized to U6. The abundance of miR-34b in WT minigene is set to 1.

Transfection of the ΔC2 minigene in HeLa cells severely affected splicing, dropping the percentage of intron processing to 15% (fig. 3.3) and indicating, therefore, that the C2 sequence is fundamental for the correct selection of the 3'ss. Sequence inspection of the C2 portion showed that it contains the CACUAAC stretch of nucleotides that perfectly matches the highly degenerated branch point consensus sequence YNYURAC. To perform fine mapping of the putative branch point, I made a point mutation of the conserved uridine at the position -2 that, according to Vorechovsky et al. [236], is associated to splicing defects and considered to disrupt the branch point function more severely than mutations to the adenine. Transfection in HeLa cells of the resulting -2U>A minigene induced almost complete intron retention. These results strongly suggest that this sequence is a *bona fide* branch point and it is fundamental for the correct processing of the pri-miR-34b transcript.

3.2.4 miR-34b splicing requires a purine-rich exonic splicing enhancer

The recognition of the 3'ss can be influenced by exonic splicing regulatory elements located downstream of the splice site in exon 2 of the pri-miR-34b transcript. To test this hypothesis, I analyzed the sequence of the exon and I identified a GAA-rich sequence that may act as an exonic splicing enhancer (ESE).

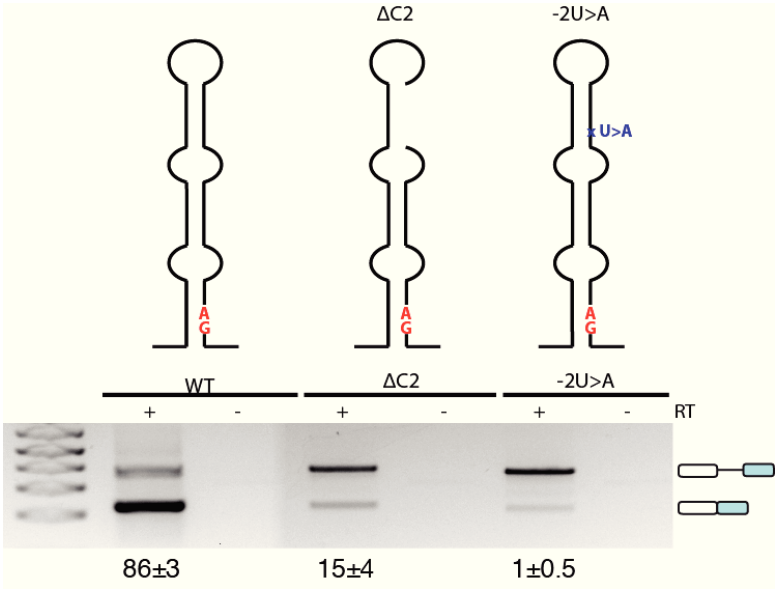


Figure 3.3: **A branch point sequence located in the hairpin promotes the 3'ss recognition** Schematic representation of pri-miR-34b hairpin, with the AG dinucleotide highlighted in red. The deleted part of the stem in ΔC2 construct is represented as empty tract. Blue ics indicates U>A point mutation on the branch point site. The lower panel shows the splicing pattern of the WT and mutant plasmids after transfection in HeLa cells. The identity of the bands is depicted on the right. Splicing percentage, expressed as means ±SD, is indicated below the gel. RT: reverse transcriptase.

To fine map this region, I deleted an internal portion of the exon, spanning from base +50 to +74 (Δ2457-2482) as illustrated in figure 3.4. Transfection of Δ2457-2482 minigene in HeLa cells drastically reduced splicing efficiency to 6%, confirming the presence of a strong enhancer element in this region. To fine map the GAA-rich region, I subjected this regulatory element to site-directed mutagenesis, substituting four purines with pyrimidines. The introduction of these four nucleotide changes (+63A>C, +65A>C, +67*A* >T and +69G>T) across the ESE (ESE mut) has a dramatic effect on the splicing outcome, leading to complete disruption of splicing (fig. 3.4, lower panel). These data support the hypothesis that the exon contains a purine-rich element (GAGAGAAGA) located 61 bp downstream of the 3'ss, that facilitate the AG dinucleotide recognition. Northern blot analysis of Δ2457-2482 and ESEmut mutant minigenes revealed that, in comparison to the WT, miR-34b production was increased by 2.0 fold and 4.0 fold, respectively. These experiments indicate that mutations that inhibit splicing, either at 3'ss or at the ESE, increase the production of mature miR-34b, confirming the direct competition between the two processing machineries.

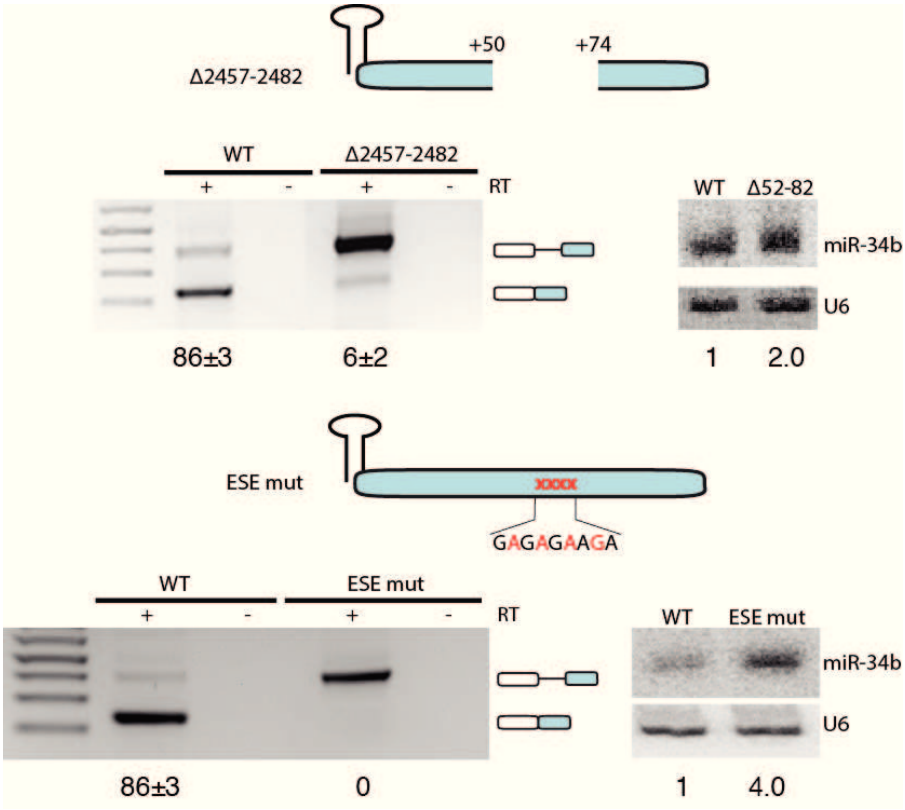


Figure 3.4: **ESE mutations impair splicing and increase miR-34b expression levels**
Schematic representation of pcDNA3pY7 Δ2457-2482 and ESE mut minigenes. Thin lines and light blue boxes correspond to the intronic pri-miR-34b hairpin and exonic sequence, respectively. Red x indicates the ESE mut minigene mutated nucleotides that are highlighted in red in the sequence. Splicing pattern of pcDNA3pY7 miR-34b WT, Δ2457-2482 and ESE mut minigenes after transfection in HeLa cells. The identity of the bands is depicted on the right. Splicing percentage, expressed as means ±SD, is indicated below the gel. RT: reverse transcriptase. Northern blot analysis of miR-34b and U6 snRNA in WT, Δ2457-2482 and ESE mut constructs. Numbers below the panel indicate the fold increase of miR-34b normalized to U6. The abundance of miR-34b in WT minigene is set to 1.

3.2.5 siRNA-mediated silencing of Drosha and DGCR8 increases miR-34b exon 2 splicing

Because splicing influences miR-34b production, I was interested in evaluating if the MPC activity influences splicing efficiency. To this aim, I performed silencing experiments to knockdown Drosha and DGCR8, the two main components of MPC. HeLa cells treated with siRNA against Drosha and DGCR8 alone (siDrosha and siDGCR8) or in combination (siD/D) were transfected with the pcDNA3pY7 2482 minigene (mutant with deletion of last part of exon 2), in which approximately 40% of the transcript is spliced, as shown in figure 3.5. Silencing of Drosha and

DGCR8 resulted in a nearly complete disappearance of the corresponding bands in western blot experiments (fig. 3.5, middle panel). In addition, the expression levels of endogenous mature miR-26b and transfected miR-34b, evaluated by qRT-PCR, were significantly decreased in double siRNA-treated cells (fig. 3.5, lower panel). In this construct, siRNA-mediated depletion of Drosha increased the percentage of splicing of the mutant from 40% to 70%. A smaller but still significant effect is observed for the depletion of DGCR8 and for the combined knockdown of the two enzymes, with a percentage of about 62% of spliced form (fig. 3.5, upper panel). The improvement of splicing efficiency after depletion of Drosha and DGCR8 indicates that MPC activity interferes with the splicing reaction acting on the nascent transcript.

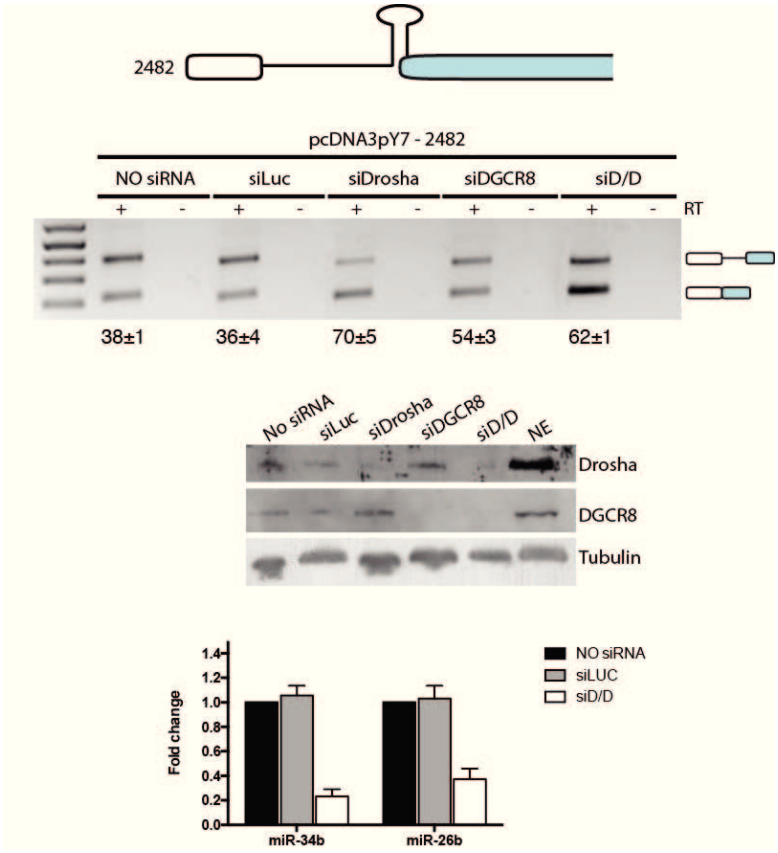


Figure 3.5: **Silencing of Drosha and/or DGCR8 increases exon 2 splicing efficiency**
[Upper panel] Schematic representation of pcDNA3pY7 2482 minigene. Splicing pattern of pcDNA3pY7 2482 minigene in non-treated cells (NO siRNA) and after transfection of a control siRNA (siLUC) or siRNA against Drosha and DGCR8 alone (siDrosha and siDGCR8) or in combination (siD/D). The identity of the bands is depicted on the right. Splicing percentage, expressed as means ±SD, is indicated below the gel. RT: reverse transcriptase. [Middle panel] Western Blot for Drosha, DGCR8 and control Tubulin. NE: nuclear extract. [Lower panel] qRT-PCR of transfected miR-34b and endogenous miR-26b, normalized for GAPDH. The abundance of miRNAs in untreated cells is set to 1.

3.2.6 Drosha and DGCR8 overexpression decreases miR-34b exon 2 splicing

To further clarify the functional role of MPC in pri-miR-34b splicing regulation, I performed overexpression experiments. HeLa cells were transiently co-transfected with pcDNA3pY7 2482 hybrid minigene along with plasmids coding for Drosha and DGCR8. Overexpression of either Drosha or DGCR8 (+Drosha and +DGCR8) showed reduced level of splicing, with an increase in intron retention, in comparison to the untransfected control. Overexpression of both coding plasmids (+D/D) showed a further decrease of splicing efficiency from 40% to 14% (fig. 3.6). Overexpression of MPC proteins had, therefore, a negative effect on splicing, increasing intron retention.

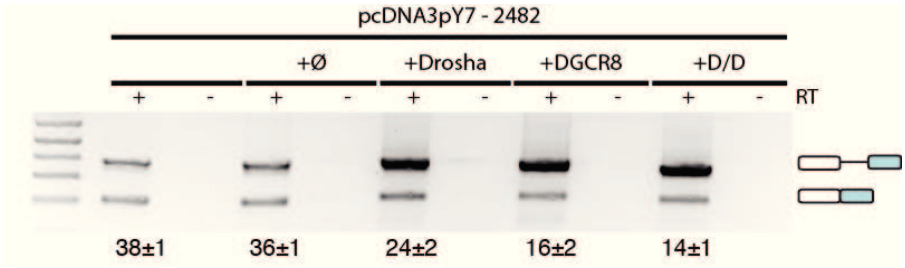


Figure 3.6: **Overexpression of Drosha and/or DGCR8 decreases exon 2 splicing efficiency**
Splicing pattern of pcDNA3pY7 2482 minigene co-transfected with an empty vector (+∅) or vectors expressing Drosha and DGCR8 alone (+Drosha and +DGCR8) or in combination (+D/D). The identity of the bands is depicted on the right. Splicing percentage, expressed as means ±SD, is indicated below the gel. RT: reverse transcriptase.

All together the silencing and overexpression experiments indicate that the MPC-dependent pri-miR-34b processing interferes with spliceosome and affects the splicing of miR-34b transcript.

3.3 Relationship between spliceosome and MPC machineries in the processing of protein-coding SO-miRNAs

3.3.1 SO-miRNA exons are not alternatively spliced, even if mature miRNAs are produced

Since most of the SO-miRNAs are located on intron-exon junction of protein-coding genes, their regulated cropping or splicing could affect final protein expression. For this reason, I decided to analyze if the exons that contain part of the pri-miRNA sequence are included in the final transcript or if some patterns of alternative splicing are detectable. I performed an RT-PCR analysis on five different cell lines available in the laboratory (HeLa, HEK293T, Hep3B, BC-1 and CFPAC-1 cells) to amplify nine endogenous SO-miRNA transcripts using primers located on the exons flanking the one that includes part of the pri-miRNA hairpin (fig. 3.7, left panel). In almost all cases the exon that contains part of the hairpin was included in the final transcript and the exon inclusion represents the unique detectable product (fig. 3.7, panels 1-8). The only exception was the NRD1 transcript: RT-PCR analysis using primers located on exon 2 and exon 6 showed as major product the amplicon without exon 3 and exon 4, which includes part of the hairpin of 5' SO-miR-761 (fig. 3.7, panel 9).

To understand the role of splicing on SO-miRNAs biosynthesis, I have evaluated the effect of SO-miRNA exons inclusion on the production of the correspondent mature miRNAs. For SO-miR-555 and SO-miR-1178, embedded in ASH1L and CIT transcripts, respectively, the available TaqMan miRNAs are not able to detect the mature microRNAs in the different cell lines (Ct values >40). The other mature SO-miRNAs are produced in the cells, with different levels of expression from one cell line to another. SO-miRNAs abundance in HeLa cells is set to 1 and the amount in the other cell lines is expressed as fold increase in comparison to this cell line (fig. 3.8). Interestingly, the Ct values of SO-miR-761, located on the intron-exon junction of the NRD1 transcript are significantly lower (Ct<28, data not shown) compared to Cts of the other miRNAs. Even if a direct comparison between expression levels among different SO-miRNAs is not reliable, this result suggests that the skipping of the exon that contains a SO-miRNA is possibly associated with an higher miRNA biosynthesis.

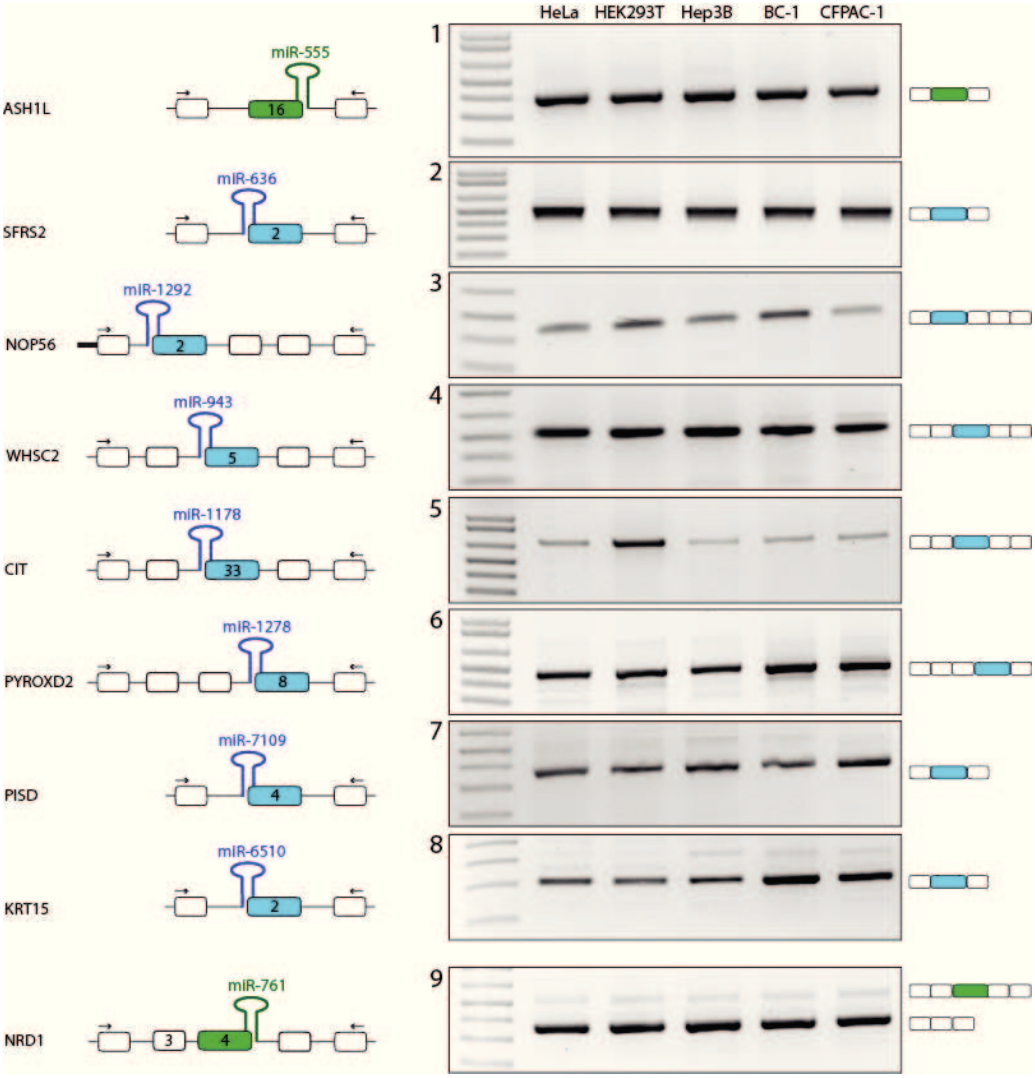


Figure 3.7: **SO-miRNA transcripts amplification in cell lines does not show any pattern of alternative splicing, except for NRD1**

[Left panel] Schematic representation of endogenous SO-miRNA exons, with their flanking exons. Boxes indicate exons. Thin line represents introns. SO-miRNA exons are numbered and highlighted in light blue or green, depending on the position of the SO-miRNAs on the 3' or 5'ss, respectively. 3' SO-miRNAs are thick blue lines and 5' SO-miRNAs thick green lines. The black superimposed arrows represent primers used for PCR amplification, listed in the material and methods. [Right panel] Panels show the pattern of splicing of nine SO-miRNA transcripts in five different cell lines, indicated above the gels. (1-8) The strong bands correspond to exon inclusion. (9) The major product corresponds to exon 3 and 4 skipping. The identity of the bands is depicted on the right.

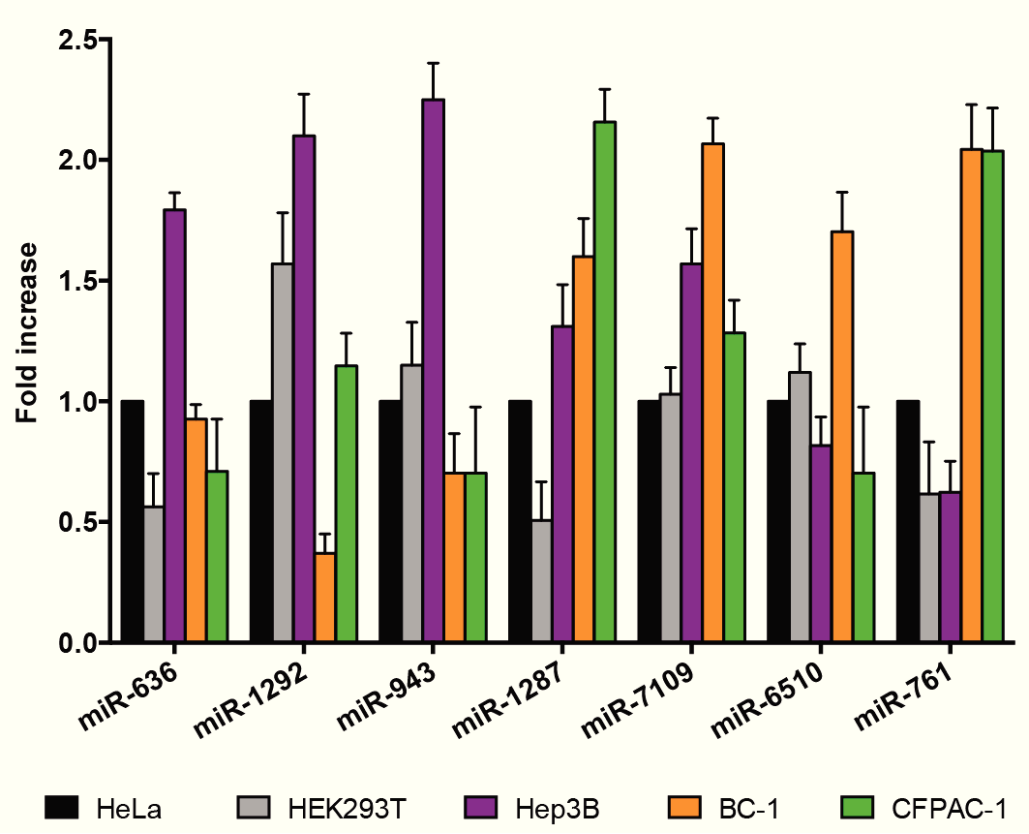


Figure 3.8: **Absence of alternative spliced exons does not affect the production of mature SO-miRNAs in cell lines**

Quantitative analysis of mature SO-miRNAs production in five different cell lines. Black, grey, purple, orange and green column colors correspond to HeLa, HEK293T, Hep3B, BC-1 and CFPAC-1 cells, respectively. TaqMan assays used are listed in the material and methods. Values are normalized to GAPDH and are plotted as fold increase compared to HeLa cells, set to 1.

3.3.2 NRD1 transcript and 5' SO-miR-761 regulation

3.3.2.1 Co-transfection of splicing factors does not change endogenous NRD1 splicing

To analyze the regulation of splicing of the NRD1 transcript, the only SO-miRNA exon that showed evidence of alternative splicing, and to identify trans-acting factors involved in exon 3 and 4 regulation, I performed transfection experiments with plasmids coding for twelve representative splicing factors. This panel of splicing factors includes SR proteins (SRSF1, SRSF3, SRSF4, SRSF7, TRA2B), Py-rich binding proteins (TIA1, PTB1), the general factors U2AF35 and U2AF65 and hnRNP proteins (hnRNP A1, hnRNP A2, hnRNP H). The transient transfection was performed in HeLa cells and the transcripts were analyzed through RT-PCR using NRD1 ex2 dir and NRD1 ex6 rev primers. As shown in figure 3.9, co-transfection with the different expression vectors had no effect on exons 3 and 4, which remain almost completely excluded in the final transcript. The skipping of exon 3 and 4 does not disrupt the reading frame of the transcript, as the missing portion of the transcript is of 204bp.

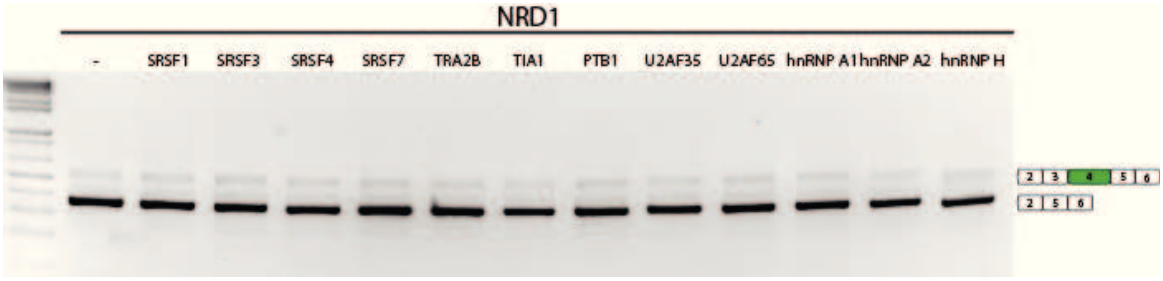


Figure 3.9: **Co-transfection of a panel of splicing factors does not change the percentage of NRD1 exon 3 and 4 inclusion**
RT-PCR analysis of endogenous NRD1 transcript after overexpression of twelve different splicing factors in HeLa cells. The identity of the bands is indicated on the right. Exon 4, containing a portion of 5' SO-miR-761, is highlighted in green.

3.3.2.2 MPC activity does not influence splicing of endogenous NRD1 transcript

In order to understand if the Microprocessor activity modulates NRD1 splicing efficiency, I made silencing experiments to knockdown its two main components Drosha and DGCR8 in HeLa cells. The efficiency of silencing was assessed at mRNA level through RT-PCR analysis using Drosha dir and Drosha rev primers and shows a nearly complete disappearance of the corresponding

band. GAPDH amplification served as a control (fig. 3.10, upper panel). Analysis of the endogenous NRD1 showed that, even if the silencing of Drosha and DGCR8 (siD/D) is effective, the knockdown had no effect on splicing efficiency if compared with the untreated sample (NO siRNA) and the sample treated with a control siRNA (siRNA against luciferase - siLUC). The percentage of exon inclusion was not increased after the treatment and exons 3 and 4 remained almost completely skipped (fig. 3.10, middle panel). Moreover, I performed a qRT-PCR on the endogenous miR-761 and miR-26b, as a control: the levels of mature miRNAs, as expected, were significantly decreased after silencing, in comparison to NO siRNA and siLUC samples (fig. 3.10, lower panel).

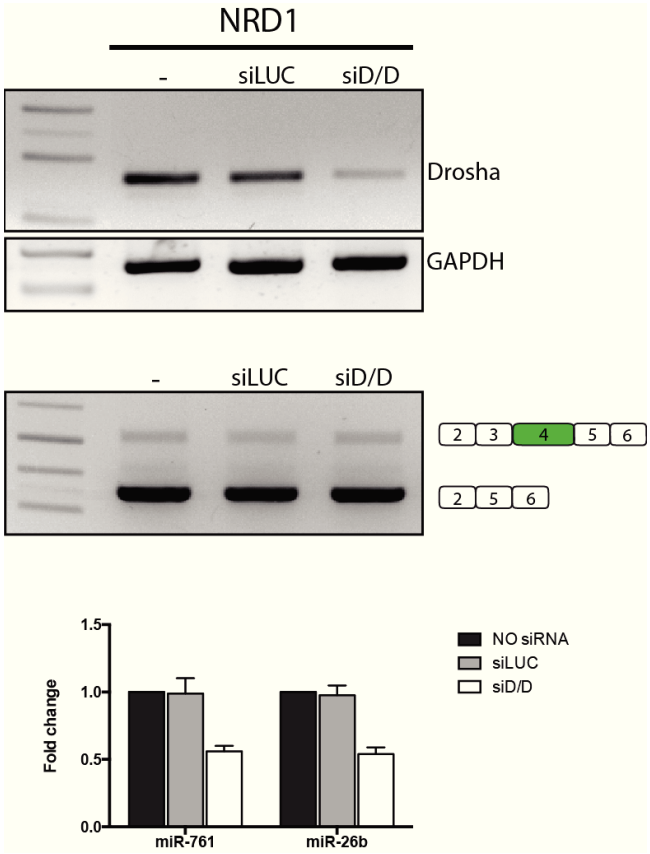


Figure 3.10: **Silencing of Drosha and DGCR8 does not improve NRD1 splicing efficiency**
Splicing pattern of endogenous NRD1 transcript after transfection of a control siRNA (siLUC) or combined silencing of Drosha and DGCR8 (siD/D). Bands identity is indicated on the right. RT-PCR analysis of Drosha and GAPDH is shown in the upper panel. Histograms show the fold change of mature miR-761 and control miR-26b, normalized to GAPDH gene. The abundance of miR-761 and miR-26b in non-treated cells (NO siRNA) is set to 1.

3.4 SF3b1-mediated splicing inhibition and SO-miRNAs profile

3.4.1 SF3b1 positively affects the production of miR-34b in the minigene system

To establish a more direct link between splicing and miRNA changes, I decided to knockdown the SF3b1 factor, that is part of the U2 snRNP and is involved in the recognition of the 3'ss during the first step of splicing. I treated HeLa cells with an siRNA against SF3b1 and analyzed the splicing pattern of transfected pcDNA3pY7 miR-34b minigene. As shown in figure 3.11, treatment with SF3b1 siRNA induced intron retention in the pcDNA3pY7 miR-34b construct, decreasing the percentage of splicing from 86 to 52%, while increased the biosynthesis of the mature miR-34b by 1.5 fold (fig.3.11). These results indicate that SF3b1 positively affects miR-34b biosynthesis through the modulation of splicing of the pcDNA3pY7 miR-34b minigene.

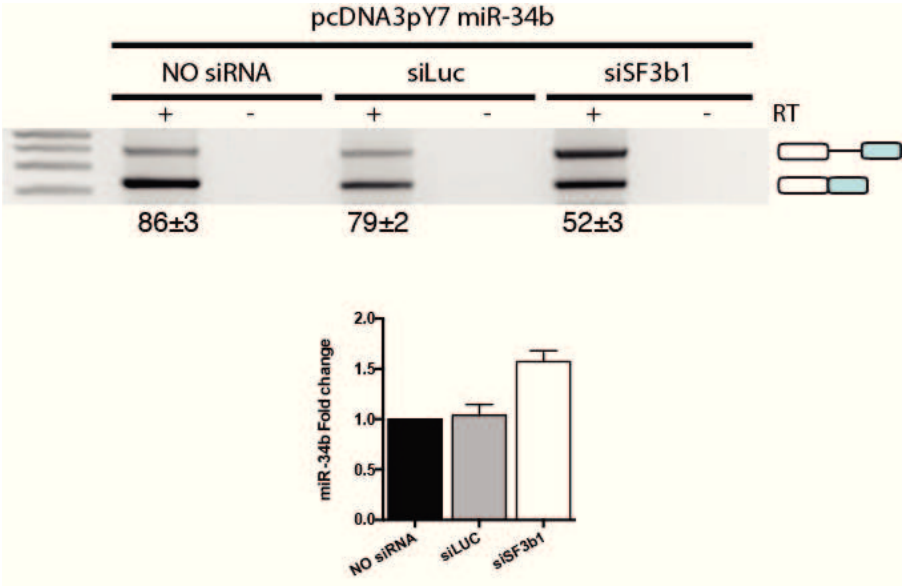


Figure 3.11: **SF3b1 silencing on pcDNA3pY7 miR-34b minigene increases the production of miR-34b**

[Upper panel] Splicing pattern of pcDNA3pY7 miR-34b minigene after treatment with a control siRNA (siLUC) or siRNA against SF3b1 (siSF3b1). Band identity is depicted on the right. Splicing percentage is indicated under the panel. RT: reverse transcriptase. [Lower panel] Quantitative analysis of mature miR-34b derived from pcDNA3pY7 miR-34b minigene, normalized to U6. The abundance of miR-34b in non-treated cells (NO siRNA - black column) is set to 1.

3.4.2 SF3b1 silencing and small RNA sequencing

3.4.2.1 SF3b1 silencing in MEC-1 cells

Based on the observations that changes in pre-mRNA splicing of miRNA-containing transcripts can directly interfere with the miRNA biosynthesis and that SF3b1 is one of the most frequently mutated spliceosomal gene [237] with particular enrichment in hematological malignancies and in solid tumors, I thought to investigate the global effect of SF3b1 silencing on miRNA biosynthesis. As an extremely high frequency of SF3b1 mutations has been reported among chronic lymphocytic leukemia patients and it is predictive of a poor outcome [48, 49], I took advantage of a B-chronic lymphocytic leukemia cell line, the MEC-1, available in the laboratory.

To determine the role of SF3b1 on miRNA biosynthesis, I induced SF3b1 depletion in MEC-1 cells by nucleofection. As detected by Western blot, SF3b1 silencing induces a nearly complete disappearance of the corresponding protein band (fig.3.12, left panel). In addition, I have evaluated through RT-PCR the effect of SF3b1 depletion on the alternative splicing of two previously reported target genes, DUSP11 and RBM5 [243]. As shown in the right panel of figure 3.12, nucleofection with siRNA against SF3b1 resulted in increased levels of DUSP11 exon 6 and RBM5 exon 16 skipping, confirming the efficacy of SF3b1 knockdown.

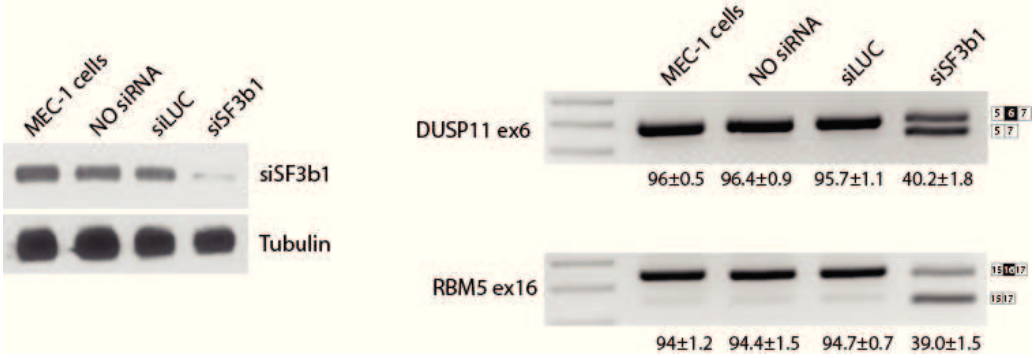


Figure 3.12: **SF3b1 silencing in MEC-1 cells is confirmed by western blot and alternative splicing changes of selected genes**

[Left panel] Western blot analysis of MEC-1 cells, untreated cells (NO siRNA), and cells treated with a control siRNA (siLUC) or siRNA against SF3b1. Tubulin serves as protein loading control. [Right panel] RT-PCR analysis of alternative splicing changes upon SF3b1 depletion in the DUSP11 and RBM5 genes. Band identity is depicted on the right. The percentage of DUSP11 exon 6 and RBM5 exon 16 inclusion, expressed as means ±SD, is indicated below the gel.

3.4.2.2 Global small RNA-seq analysis of MEC-1 cells after SF3b1 silencing

To analyze the effect of SF3b1 depletion on miRNA production in MEC-1 cells, I assessed global miRNA expression changes by deep-sequencing total RNAs extracted from four samples treated with the control siRNA against luciferase (siLUC) and four samples treated with SF3b1 siRNA (siSF3b1), derived from four independent biological replicates. In particular, I performed a small RNA-seq with the llumina HiSeq2000 platform in order to obtain a global analysis of miRNA differential expression between untreated and siSF3b1 treated samples. Reads derived from the sequencing are counted and mapped against human reference genome and to known mature miRNAs using the Bowtie 2 software. For each miRNA, I have considered both the mature product derived from the 3' arm of the hairpin (miRNA-3p) and the one arising from the 5' stem (miRNA-5p). To identify differentially expressed miRNAs, I performed statistical analyses using the Bioconductor package DESeq2. This package firstly normalizes the data, then calculates the variance based on the replicates for each condition and finally computes statistical tests to find differentially expressed miRNAs. I analyzed miRNA expression differences between MEC-1 cells with depletion of SF3b1 and control cells. As shown in figure 3.13, the MEC-1 cells depleted of SF3b1 had a differential miRNA expression profile compared to the control ones. To screen up or downregulated miRNAs I used a fold change $\geq$2.0 and a FDR value $\leq$0.01 as threshold. A total of 266 miRNAs (13.2% of 2020 miRNAs analyzed) were differentially expressed between the siSF3b1-treated and non-treated cells. 212 miRNAs are downregulated in siSF3b1-treated cells, 54 are upregulated, while the expression levels of the remaining 1754 miRNA are not affected by the depletion of SF3b1. Full lists of differentially expressed miRNAs, logarithmic base 2 fold changes and FDR values are reported in appendices 1 and 2. I employed an unsupervised hierarchical clustering based on the variation of expression for each miRNA across the samples examined to explore the correlation between the four biological replicates. As expected, the four control samples were clustered together, as well as the four SF3b1-treated samples (fig. 3.13).

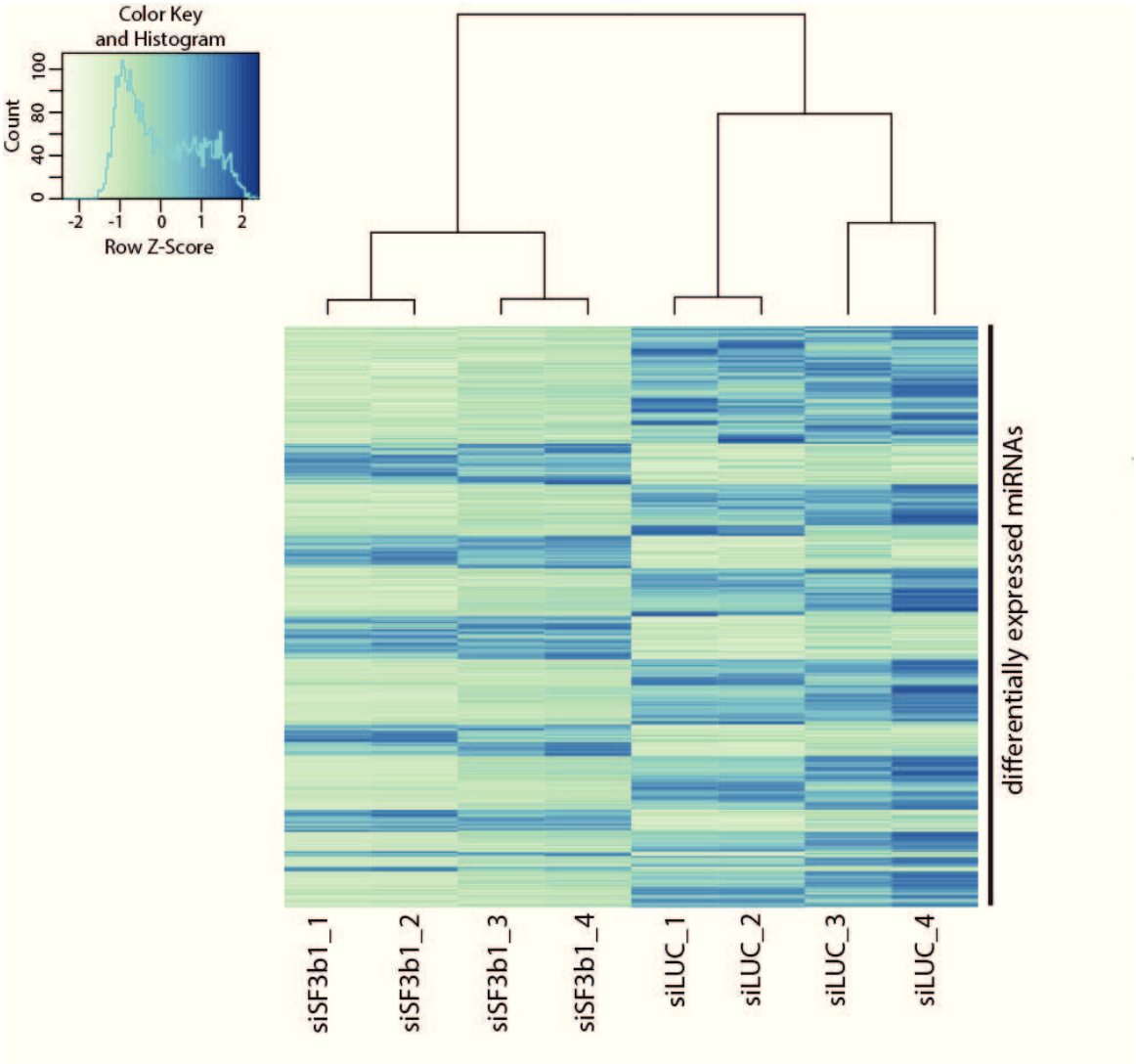


Figure 3.13: **Heat map representation of differential miRNA expression profile of MEC-1 cells with depletion of SF3b1 and control cells**

Heat map of differentially expressed human mature miRNAs in four siSF3b1 treated samples and four control samples (siLUC). Each column represents an individual sample and each row represents one miRNA. Differentially expressed miRNAs are plotted on Y-axis. Cluster of untreated and SF3b1-treated samples is shown by hierarchical tree on the X-axis. Colors represent the expression level fold changes: upregulation - blue, downregulation - light green.

3.4.2.3 Enrichment of upregulated SO-miRNAs after SF3b1 knockdown

To deeply analyze the differential expression profile of miRNAs in MEC-1 cells after knockdown of SF3b1, I divided the miRNAs into three groups:

- the **SO-miRNAs** group, that comprises the 3' SO-miRNAs, the 5' SO-miRNAs and the 3'-5' SO-miRNAs (table 3.1);
- the **mirtrons** group, the peculiar class of intronic Drosha-independent miRNAs, that includes the pure mirtrons, the 3' and the 5' tailed mirtrons;
- the group of **intronic and exonic miRNAs**, composed of the canonical intronic miRNAs, plus the small number of exonic miRNAs.

The differential expression of mature miRNAs within the three categories was evaluated using the ROAST gene set test, which allows the analysis and the assignation of a single FDR value to a group of miRNAs as a unit.
Differential expression analysis of SO-miRNAs group revealed that the depletion of SF3b1 in MEC-1 cells had the tendency to increase the expression levels of the SO-miRNAs. The ROAST gene set test showed that there was a statistically significant (FDR=0.0007) enrichment of upregulated SO-miRNAs in the SF3b1 treated cells, compared to the control cells. As shown in figure 3.14, there was an accumulation of SO-miRNAs in the pink region of the bar that corresponds to upregulated miRNAs. The expression of 39% of mature SO-miRNAs increased after the silencing of SF3b1. For example, among the most upregulated SO-miRNAs, I founded the 3' SO miRNAs miR-205-3p, miR-5187-3p, miR-6510-5p, miR-3614-3p, miR-7109-5p, and the 5'SO miRNAs miR-4260, miR-3655, miR-3656 and miR-8069. These results indicate that splicing inhibition mediated by SF3b1 depletion has a global positive effect on the production of SO-miRNAs.

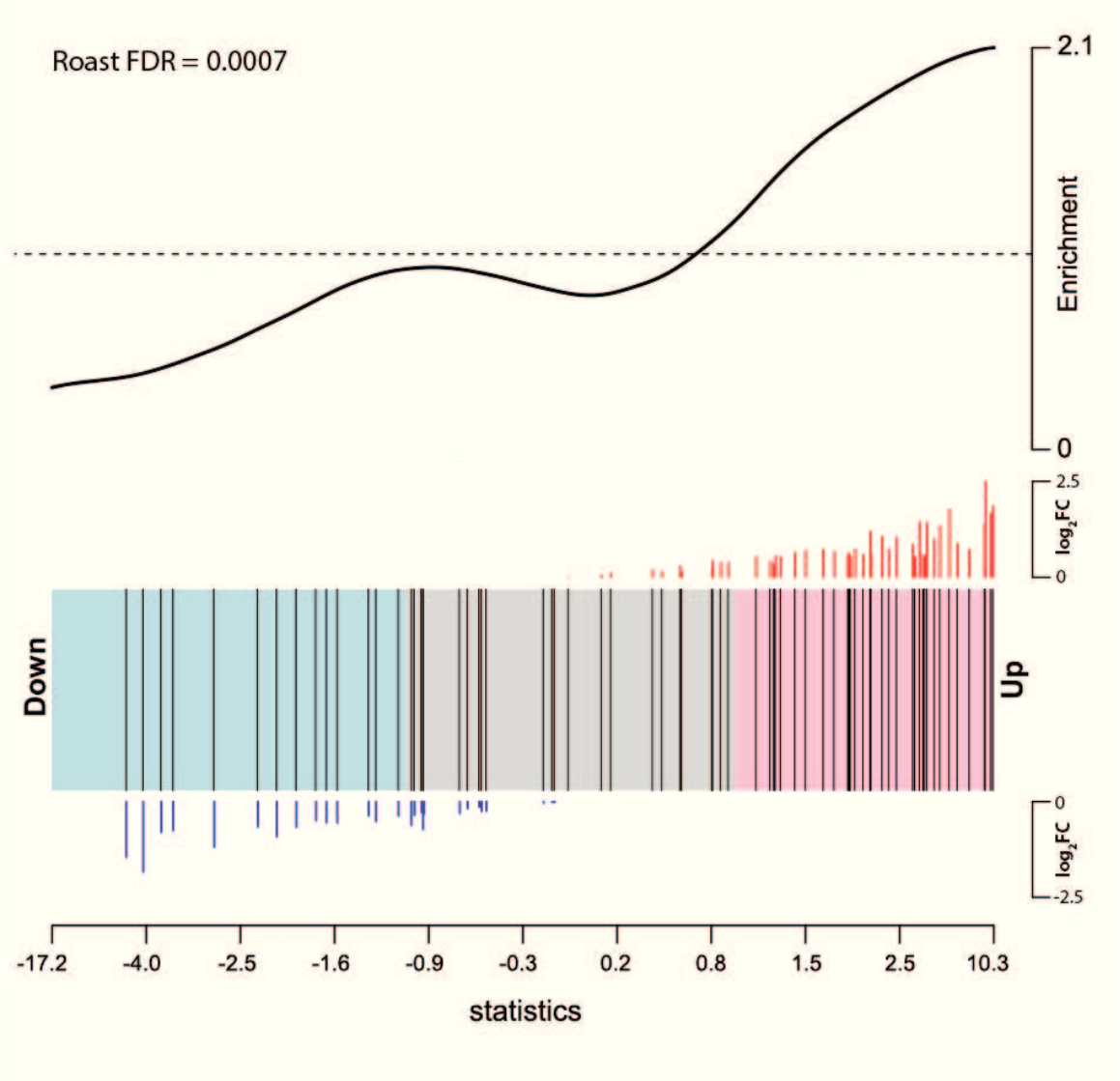


Figure 3.14: **SF3b1 silencing increases the level of SO-miRNAs in MEC-1 cells**
Graphic representation of SO-miRNAs differential expression in MEC-1 cells treated with an siRNA against SF3b1, in comparison to untreated cells. The SO-miRNA enrichment is plotted on the Y-axis in the upper panel. In the bottom panel each black vertical line represents a single SO-miRNA. Pink, grey and light blue regions correspond to upregulated, invariant and downregulated miRNAs. The log_2FC of each single miRNA is indicated above and below the panel as red and blue line, according to their positive or negative value, respectively. The differential expression, expressed by the empirical bayes moderated t-statistics is plotted on the X-axis. The FDR value of ROAST gene set test is indicated.

3.4.2.4 SF3b1 depletion slightly decreases the production of mature intronic and exonic miRNAs

In order to verify that splicing manipulation through SF3b1 knockdown has a specific positive effect on the production of SO-miRNAs, I considered the big group of intronic and exonic miRNAs, whose pri-miRNA hairpins are located in larger introns or in the middle of exons far away from splice sites, avoiding interference between spliceosome and the MPC. Differential expression analysis of this group of miRNAs revealed that depletion of SF3b1 slightly decreases the production of mature intronic and exonic miRNAs in a global manner. Indeed, I observed a small, but statistically significant (FDR=0.001) reduction in the expression levels of intronic and exonic miRNAs (fig. 3.15). 32% of intronic and exonic miRNAs are downregulated in SF3b1 depleted cells (fig. 3.15, light blue region). These data indicate that the effect of splicing manipulation is not general on the production of all mature miRNAs, but is closely related to the position of the pri-miRNA hairpins across the transcripts.

3.4.2.5 Enrichment of upregulated mirtrons after SF3b1 knockdown

In addition, I analyzed the effect of SF3b1 depletion on the peculiar class of mirtrons that arise from spliced-out introns, bypassing Drosha processing. Surprisingly, the differential expression analysis of the mirtrons group revealed that knockdown of SF3b1 had the tendency to upregulate the production of mirtrons. The analysis showed that the production of 41% of mirtrons was significantly (FDR=0.001) increased in the MEC-1 cells treated with the siRNA against SF3b1, compared to the control cells (fig. 3.16, pink region). Together these findings suggest that although this peculiar class of miRNAs is recognized and processed by the spliceosome, via a Drosha-independent pathway, splicing inhibition may enhance recognition of pri-mirtron hairpins by Drosha, inducing the production of mature miRNAs, as previously suggested [179].

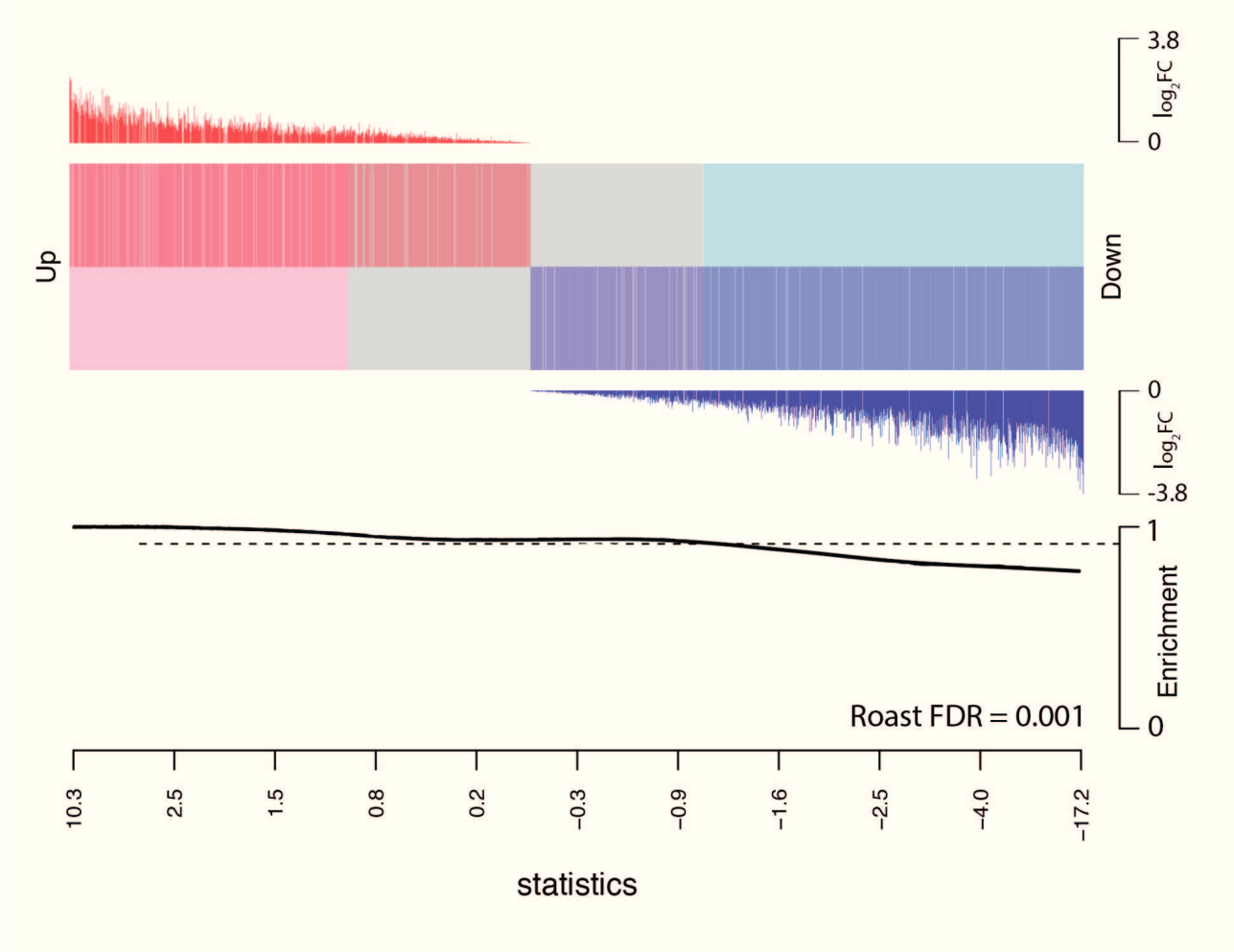


Figure 3.15: **SF3b1 silencing slightly decreases the production of intronic and exonic miRNAs** Graphic representation of intronic and exonic miRNAs differential expression in MEC-1 cells treated with an siRNA against SF3b1, in comparison to untreated cells. The miRNAs enrichment is plotted on the Y-axis in the lower panel. In the upper panel, pink, grey and light blue regions correspond to upregulated, invariant and downregulated miRNAs. The $\log_2$FC of each single miRNA is indicated above and below the panel as red and blue line, according to their positive or negative value, respectively. The differential expression, indicated by the empirical bayes moderated t-statistics is plotted on the X-axis. The FDR value of ROAST gene set test is indicated.

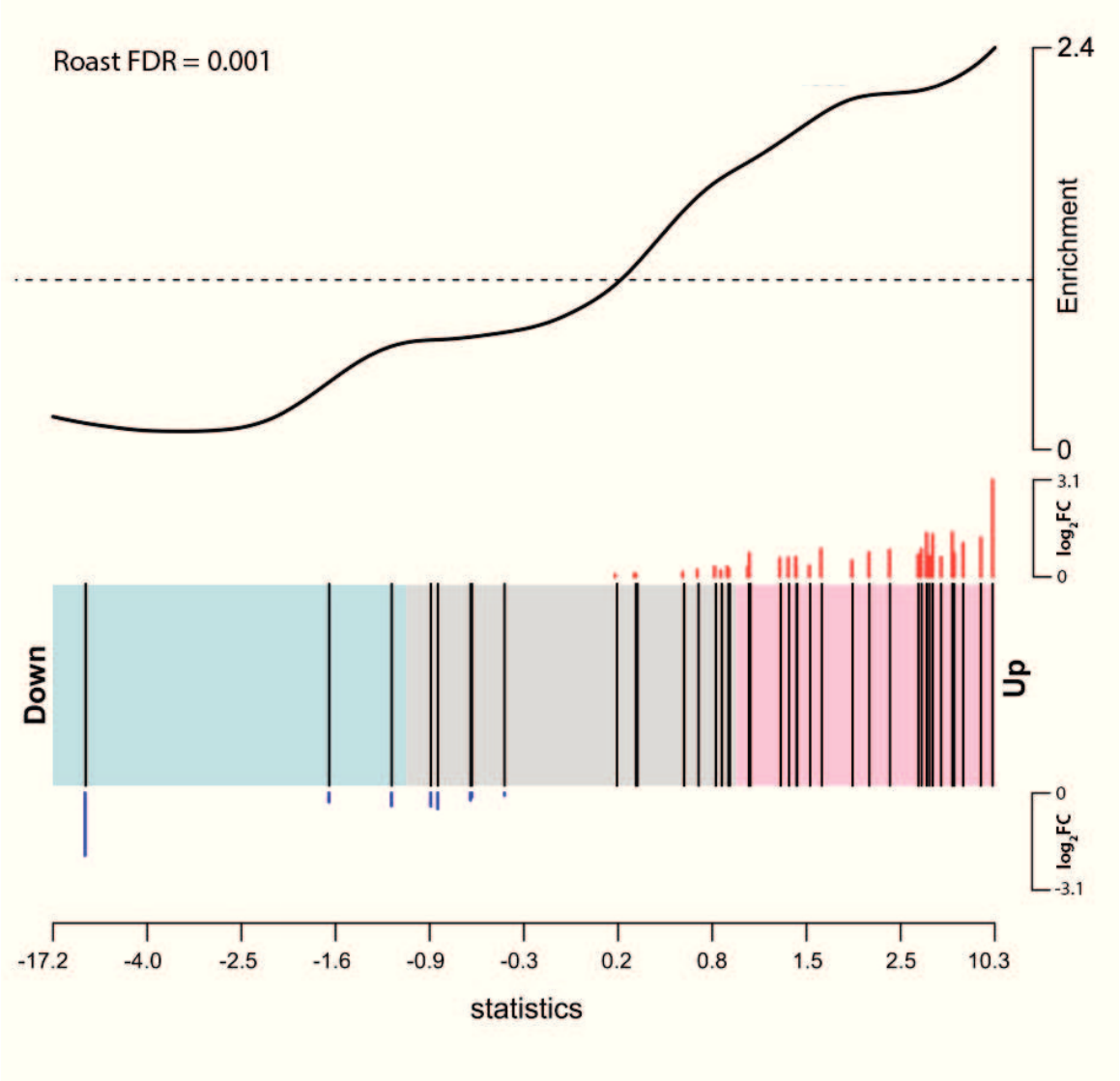


Figure 3.16: **Upregulation of mirtrons after SF3b1 depletion**

Graphic representation of mirtrons differential expression in MEC-1 cells treated with an siRNA against SF3b1, in comparison to untreated cells. The mirtrons enrichment is plotted on the Y-axis in the upper panel. In the bottom panel each black vertical line represents a single mirtron. Pink, grey and light blue regions correspond to upregulated, invariant and downregulated miRNAs. The $\log_2$FC of each single miRNA, is indicated above and below the panel as red and blue line, according to their positive or negative value, respectively. The differential expression, indicated by the empirical bayes moderated t-statistics is plotted on the X-axis. The FDR value of ROAST gene set test is indicated.

3.4.2.6 pCI-neo-PISD and pcDNA3pY7-KRT15 minigene systems for small RNA-seq data validation

To validate the differential expression of SO-miRNAs in the MEC-1 cells depleted of SF3b1, I set up two different minigene systems. The pri-miR-7109 hairpin, located on the intron 3 - exon 4 junction of PISD transcript and the pri-miR-6510, overlapping with the exon 2 of KRT15 transcript, along with their flanking regions, were cloned into the pCI-neo and pcDNA3pY7 constructs, in order to obtain the pCI-neo-PISD and the pcDNA3pY7-KRT15 constructs, respectively (fig. 3.17, upper panel). I chose these two 3' SO-miRNAs because they had a high number of normalized counts both in the control samples and in the samples with SF3b1 deletion, as shown in the normalized counts plot (fig. 3.17, middle panel). SF3b1 knockdown induced an increase of expression of these two SO-miRNAs by about 2.0 fold. I induced SF3b1 depletion in HeLa cells and transiently transfected the pCI-neo-PISD and the pcDNA3pY7-KRT15 constructs. TaqMan miRNA analysis of the transfected mature miR-7109 and miR-6510 showed that both miR-7109 and miR-6510 are upregulated in the siSF3b1-treated cells by about 2.0 fold, as determined by the RNA-seq analyses. In addition, I analyzed the effect of SF3b1 depletion and miRNA biosynthesis upregulation on the splicing pattern of PISD and KRT15 minigenes. Splicing pattern analysis did not show any difference between the untreated samples and samples with silencing of SF3b1. These results indicate that the depletion of SF3b1 increases the production of SO-miRNAs, but does not directly affect the splicing patterns of the corresponding exons.

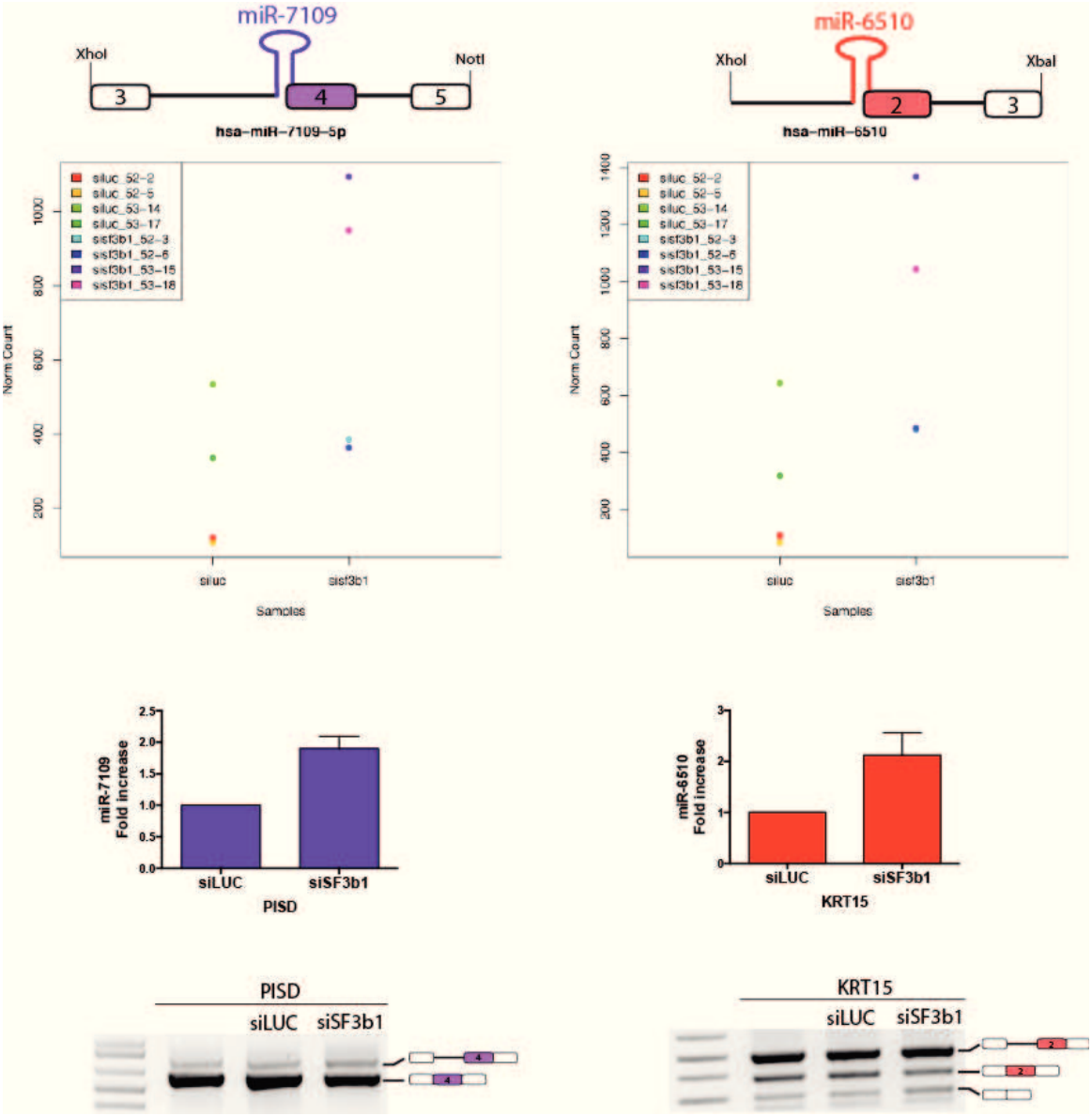


Figure 3.17: **Validation of small RNA-seq data in the pCI-neo-PISD and pcDNA3pY7-KRT15 minigenes**

[Upper panel] Schematic representation of pCI-neo-PISD and pcDNA3pY7-KRT15 minigenes. Light purple and light red boxes are PISD exon 4 and KRT15 exon 2, respectively. pri-miR-7109 and pri-miR-6510 hairpins are highlighted in purple and red. [Middle panel] small RNA-seq normalized counts plot of siLUC and siF3B1 MEC-1 treated cells. Colored dots indicate the counts of each single sample sent to sequence. [Lower panel] Quantitative analysis of mature miR-7109 and miR-6510 expression from minigenes, normalized to U6. The abundance of miR-7109 and miR-6510 in siLUC treated samples is set to 1. Splicing pattern of PISD and KRT15 minigene upon SF3b1 depletion in HeLa cells. Band identity is depicted on the right.

3.5 SO-miRNAs and keratinocytes

Analysis of host transcripts that contain SO-miRNAs revealed that three of them correspond to SO-miRNA coding genes highly expressed in keratinocytes (fig.3.18):

- the 3' SO-miR-936 hairpin overlaps with the intron 28 - exon 29 junction of COL17A1 transcript;
- the 5' SO-miR-4260 hairpin overlaps with the exon 15 - intron 15 junction of LAMB3 transcript;
- the 3'-5' SO-miR-711 hairpin overlaps with both the exon 62 3' and 5'ss of COL7A1 transcript.

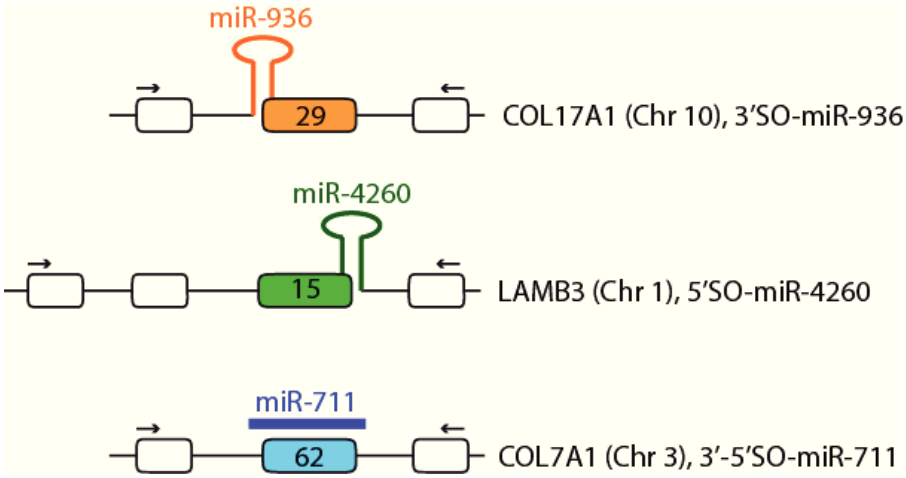


Figure 3.18: **Three SO-pri-miRNA hairpins overlap with the intron-exon junction of genes expressed in keratinocytes**

Schematic representation of keratinocytes SO-miRNAs and their host transcripts: 3' SO-miR-936 (orange) is embedded in COL17A1 transcript (light orange), 5' SO-miR-4260 (green) is embedded in LAMB3 transcript (light green) and 3'-5' SO-miR-711 (blue) is embedded in COL7A1 transcript (light blue). Thin lines, thick colored lines and boxes represent the introns, the pri-miRNA hairpins and the exons, respectively. Exons containing part of the hairpins are colored and the exon number sare indicated.

These three genes are highly expressed at the basal layer of epidermis, where keratinocytes proliferate and mediate the adhesion of the epidermis with the underlying dermis. As basal membrane proteins, their expression levels decrease during differentiation [223, 224].

For this reason, to investigate the crosstalk between the spliceosome and MPC in a physiological context, I took advantage of a normal human primary epidermal keratinocyte (NHEK) culture.

I used the following model of keratinocytes differentiation: basal proliferative NHEK were grown in KGM medium in absence of calcium and then I induced differentiation of NHEK cells by increasing calcium concentration in the medium to a final concentration of 1.2mM. Calcium supplementation of the growth medium activated the differentiation program: the cells stratified after 1-2 days and were terminally differentiated after 3-5 days. I set up three independent NHEK cultures, collecting total RNA samples of time 0 proliferative keratinocytes and after 1 to 5 days of calcium treatment.

In order to verify the successful induction of the differentiation process, I performed a qRT-PCR analysis of the differentiation marker involucrin, which expression increases with differentiation, at each different time of keratinocyte culture (fig.3.19). All the data were normalized with the geometric mean of TBP, RPLP0 and RPL13A that are used as housekeeping genes (HKGs) to study differentiation of keratinocytes by qRT-PCR, as reported in literature [238].

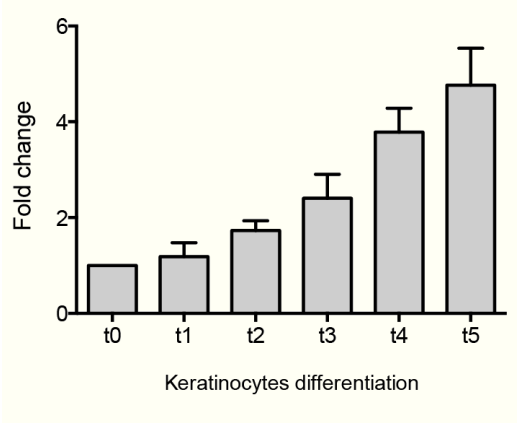


Figure 3.19: **Expression of involucrin differentiation marker in calcium-treated NHEK cells**
qRT-PCR analysis of involucrin marker in undifferentiated keratinocytes (t0) and in calcium-induced differentiated keratinocytes (t1 to t5). The expression of involucrin in undifferentiated cells is set to 1. All data were normalized to the geometric mean of TBP, RPLP0 and RPL13A. Results were presented as mean ±SD of three independent cell cultures.

3.5.1 Keratinocyte differentiation: transcripts abundance and SO-miRNAs inverse correlation

In order to confirm that the differentiation process is associated with a decreased expression of the basal membrane proteins collagen XVII and laminin 5, I analyzed the abundance of the COL17A1 and LAMB3 transcripts at the various time points by qRT-PCR. As expected, the assay showed that the expression levels of COL17A1 and LAMB3 mRNAs were decreased during differentiation (fig.3.20).

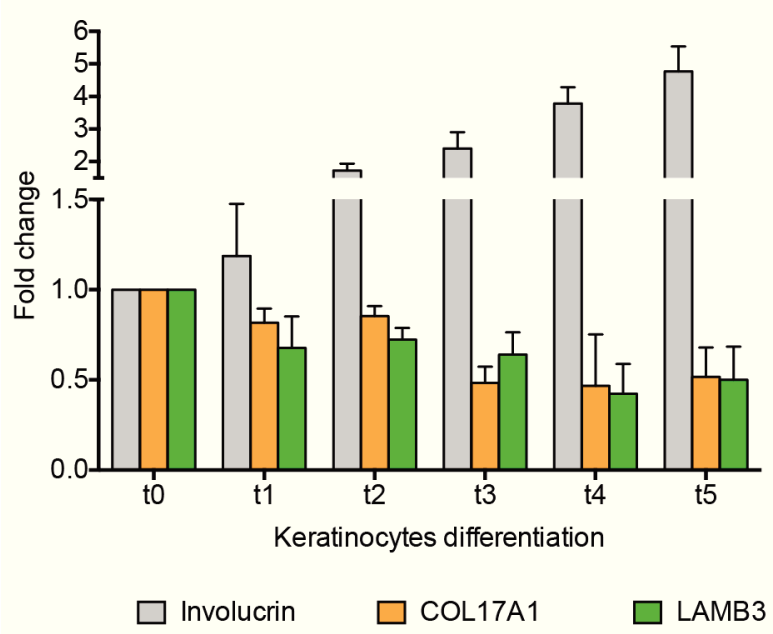


Figure 3.20: **Keratinocytes differentiation is associated with decreased expression levels of COL17A1 and LAMB3 transcripts**

qRT-PCR measurement of involucrin (grey), COL17A1 (light orange) and LAMB3 (light green) mRNAs. The abundance of the transcripts amongst the different stage of differentiation is expressed as fold change relative to undifferentiated cells, set to 1. All data were normalized to the geometric mean of TBP, RPLP0 and RPL13A. Results were presented as mean ±SD of three independent cell cultures.

Subsequently, I wanted to determine if the decrease in the production of COL17A1 and LAMB3 mRNAs during differentiation from the basal layer to the cornified stratum correlates with an increased production of the corresponding miRNAs. TaqMan miRNA assay at various time points during differentiation revealed that the expression of miR-936 and miR-4260, embedded in the COL17A1 and LAMB3 transcripts, respectively, was significantly upregulated as shown in figure 3.21.

Quantification of mature miR-711 was not included in the analysis because the available TaqMan miRNA assay did not detect the mature miRNA, and therefore also the corresponding mRNA quantification was excluded from the analysis.

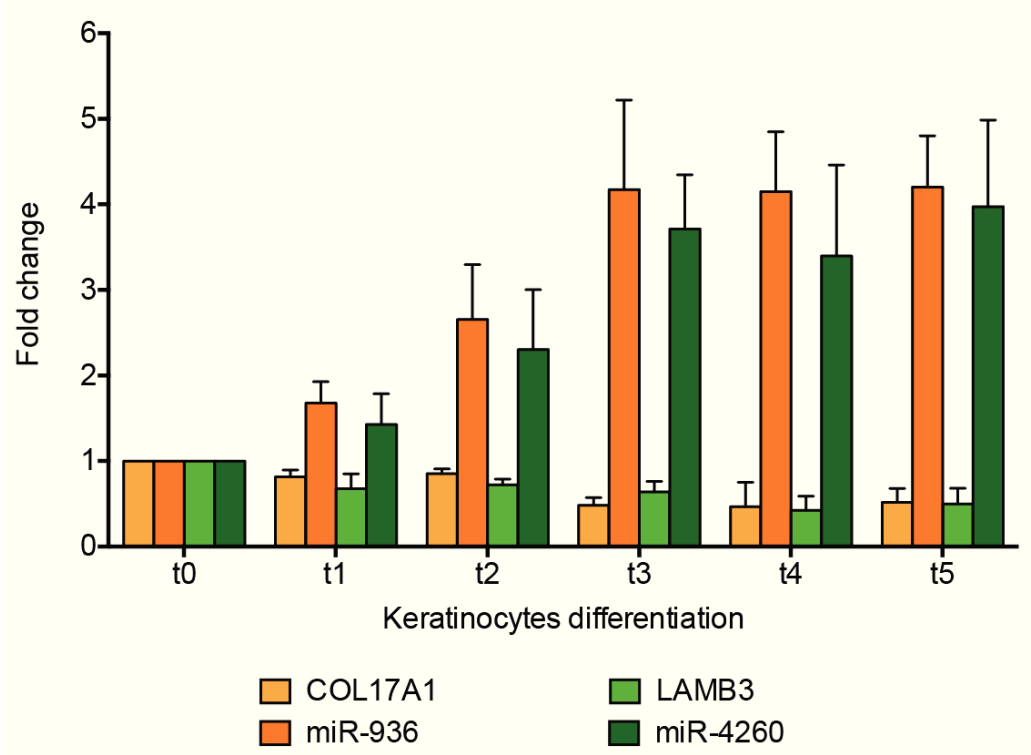


Figure 3.21: **Keratinocytes differentiation is associated with an upregulation of mature miR-936 and miR-4260**

Quantitative analysis of mature miR-936 (orange) and miR-4260 (green) in the undifferentiated cells, set to 1 and after 1 to 5 days of calcium treatment. The graphic summarizes the effect of keratinocytes differentiation on transcript abundance, as shown in previous figure (COL17A1 light orange and LAMB3 light green) and on the mature miRNAs production. All data were normalized to the geometric mean of TBP, RPLP0 and RPL13A. Results were presented as mean ±SD of three independent cell cultures.

To assure that upregulation of mature miR-936 and miR-4260 was strictly related to changes of the corresponding mRNAs expression levels, and was not a general effect of the differentiation process, I decided to quantify the expression level of an intronic miRNA, the miR-330 embedded in the intron of an unrelated gene (EML2). TaqMan miRNA assay showed that, as expected, the expression of the miR-330 did not vary amongst the different steps of differentiation (fig. 3.22).

At the same time, I also quantified the abundance of the major MPC component Drosha, to be sure that changes in mature miRNAs levels are not due to an upregulation of cropping activity.

As shown in figure 3.22 the expression level of Drosha was not significantly affected during differentiation.

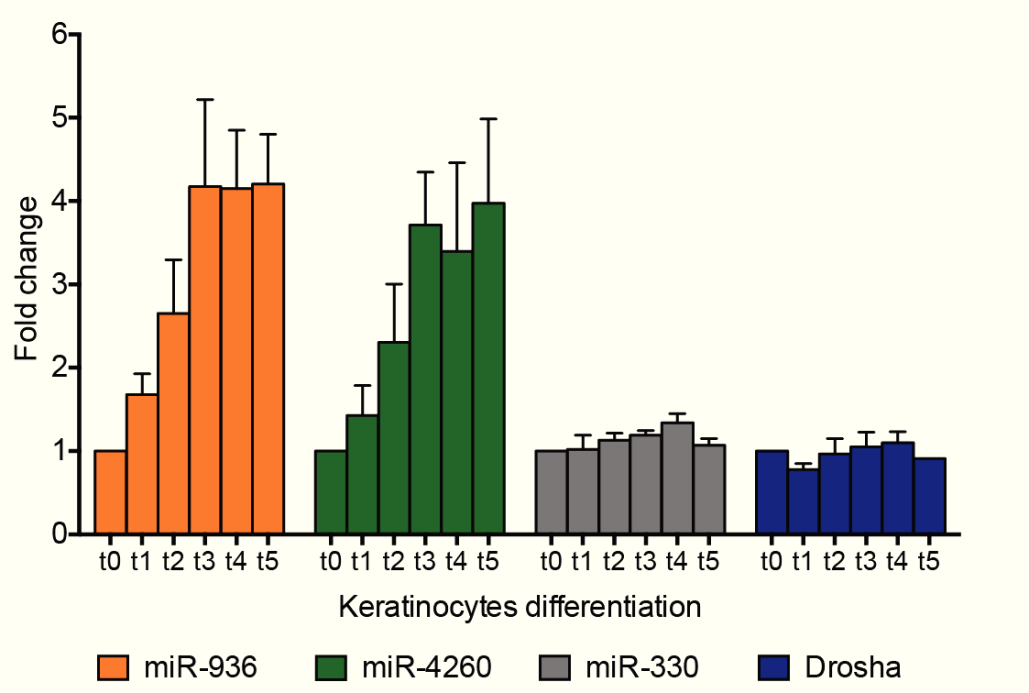


Figure 3.22: **Unchanged expression levels of intronic miR-330 and Drosha during differentiation** Quantitative analysis of intronic miR-330 (grey) and MPC major component Drosha (dark blue) in the undifferentiated cells, set to 1, and after 1 to 5 day of calcium treatment. The graphic summarizes the effect of differentiation on keratinocytes related miRNAs, as shown in previous figure (miR-936 orange and miR-4260 green) and on the mature miR-330 and Drosha production. All data were normalized to the geometric mean of TBP, RPLP0 and RPL13A. Results were presented as mean ±SD of three independent cell cultures.

All together these data indicate that, upon differentiation from basal layer to cornified stratum, the SO-miRNAs expression level and the corresponding transcripts abundance are inversely correlated: decrease in the expression of SO-miRNA transcripts associates to an upregulation of the corresponding SO-miRNA levels, reinforcing the idea of a direct functional relationship between the spliceosome and the MPC.

3.5.2 Splicing pattern of SO-miRNA exons does not change upon keratinocytes differentiation

To further characterize the relationship between the two processing machineries and to verify if increased SO-miRNAs biosynthesis in differentiated keratinocytes was associated to changes

in the splicing pattern of the exons that comprise part of the pri-miRNA hairpins, I performed an RT-PCR analysis on COL17A1 and LAMB3 transcripts, using primers located on the exons flanking the SO-miRNA exons.

The result showed that both COL17A1 exon 29 and LAMB3 exon 15 were always included in the final transcript and no alternative spliced isoforms were detectable at the different time points of differentiation (fig. 3.23).

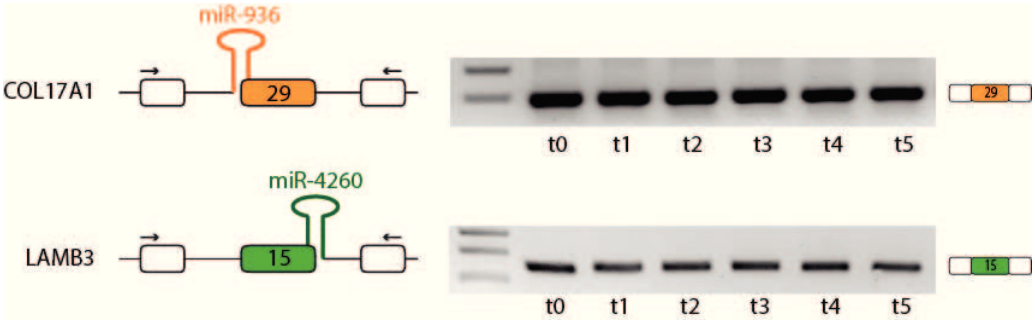


Figure 3.23: **Keratinocytes differentiation is not associated with COL17A1 and LAMB3 splicing pattern changes**

[Left panel] Schematic representation of COL17A1 and LAMB3 transcripts: exon 29 and 15 are highlighted in light orange and green, respectively. The black arrows represent primers used for RT-PCR amplification. [Right panel] Splicing pattern analysis of COL17A1 and LAMB3 transcripts in the undifferentiated cells (t0) and after 1 to 5 day of calcium treatment (t1 to t5). Band identity is depicted on the right.

3.5.3 Anti-proliferative effect of miR-936, miR-4260 and miR-711 on keratinocytes

To better characterize the functional significance of this regulatory switch in the control of keratinocytes differentiation, I decided to test the effect of miR-936, miR-4260 and miR-711 overexpression on keratinocytes proliferation. Since primary keratinocytes have only a finite life span, that does not allow long-term investigations, I decided to perform the experiments in HaCaT cells, a well-known immortalized human keratinocyte cell lines, that can be grown in traditional media and can be maintained in culture for long periods of time. To investigate the effect of exogenous SO-miRNAs administration, HaCaT cells were transfected with the three miRNA mimics separately and EdU incorporation experiments were performed to assess their proliferation. As shown in figure 3.24, overexpression of mir-936, miR-4260 and miR-711 decreased HaCaT cells proliferation.

The strongest effect was obtained with miR-4260, that reduces the percentage of EdU-positive cells from 60% to 22%, while miR-936 and miR-711 reduces the percentage to 36% and 38.8%, respectively.

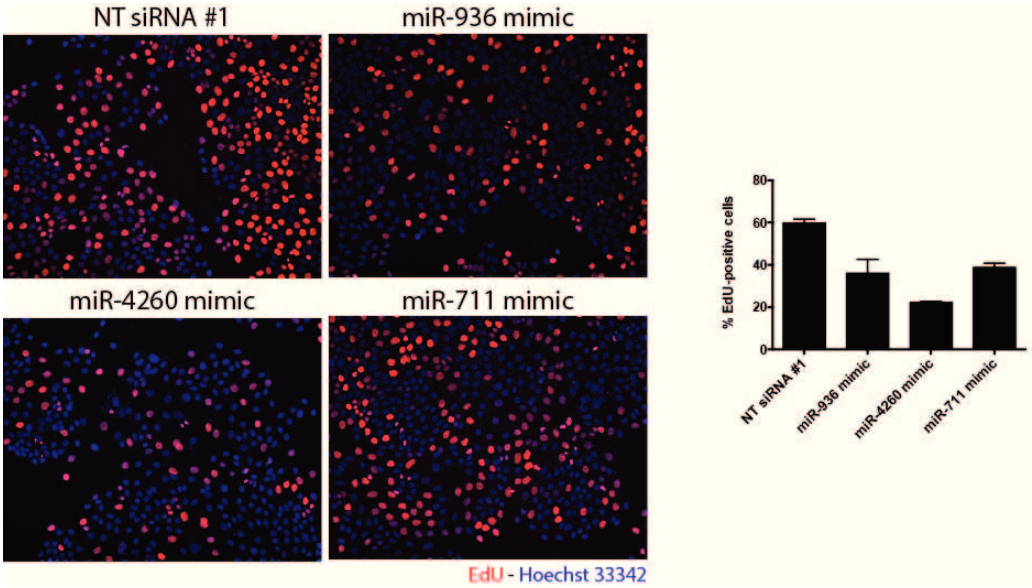


Figure 3.24: **Effect of miRNA mimics transfection on HaCaT cells**
[Left panel] Proliferating HaCaT cells were labeled with EdU. The Click-it reaction revealed EdU staining (red). Cell nuclei were stained with Hoechst 33342 (blue). The merged images are representative of the data obtained. [Right panel] The percentage of EdU-positive HaCaT cells was quantified. Results were presented as mean ±SD.

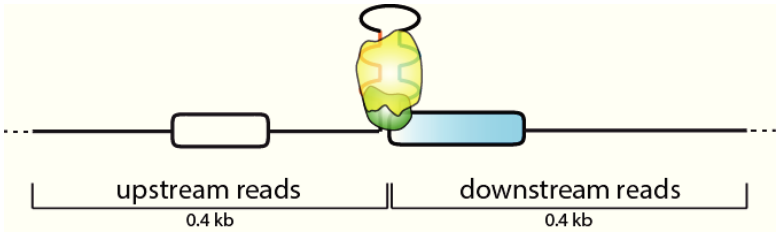
The inhibitory effect of miR-936, miR-4260 and miR-711 on HaCaT cells proliferation emphasizes the importance of the inverse correlation between the expression of COL17A1, LAMB3 and COL7A1 transcripts and their corresponding SO-miRNAs in the transition from basal proliferative to highly differentiated keratinocytes.

3.6 Microprocessor-mediated termination of SO-miRNAs transcripts

Recently, some miRNA hairpins located in the last exons of long non-coding genes were shown to be involved in transcriptional termination [239]. The hypothesis of a Microprocessor-mediated mechanism of transcriptional termination was proposed by Proudfoot and colleagues, using a genome-wide chromatin RNA-seq analysis where the authors found that MPC mediates the transcriptional termination of pri-miR-122 long non-coding RNA (lncRNA) and other lnc-pri-miRNA transcripts. They demonstrated that lnc-pri-miRNAs 3'-ends were generated by the

MPC and that MPC cleavage inhibition by depletion of either Drosha or DGCR8 leads to an extensive transcriptional read-through downstream of lnc-pri-miRNAs.

To test the hypothesis that the MPC could also be involved in transcriptional termination and 3'-end formation of SO-pri-miRNA transcripts, I took advantage of the chromatin RNA-seq data deposited in the Gene Expression Omnibus database (accession number GSE58838). I analyzed the effect of Drosha depletion on SO-pri-miRNA transcripts comparing the chromatin RNA-seq data from HeLa cells treated with a control siRNA (siCTR) and with an siRNA against Drosha (siDrosha). I have considered a region of 0.8kb, which includes 400 bp upstream and downstream of the SO-miRNA hairpin and I have counted the reads upstream and downstream of the SO-pri-miRNA hairpin, normalizing these latter for the first ones. Finally, I made the ratio between the normalized read counts (RPKM) of Drosha siRNA and control siRNA-treated cells. I expressed this ratio as fold increase of siDrosha-treated cells in comparison to siCTR-treated one (3.25, middle table). Fold increase of SO-miRNAs, for which the RNA-seq gives reads and therefore are expressed in HeLa cells, were plotted in a graph. As shown in figure 3.25 the ratio for SO-pri-miRNA transcripts is greater than 1, with the only two exceptions of SO-pri-miR-1287 and pri-miR-636. This means that there is an accumulation of reads downstream of the hairpins after Drosha depletion. I observed variable fold increase values after Drosha knockdown among different SO-pri-miRNA transcripts. For example, cropping inhibition had a strong effect on transcription extension downstream of SO-pri-miRNAs on AHM, NOP56, LAMB3 and DNM2 transcripts (fold increase$>$ 2.0) and a mild effect on NELFA, CIT, PLEKHM1P, NRD1, ASH1L, COL7A1 and SCRIB transcripts (1$<$fold increase$<$2). All together these data suggest that MPC cleavage could mediate transcriptional termination on some endogenous SO-miRNA transcripts.



SO-miRNA	Transcript	Normalized read counts (RPKM) siCTR ratio	siDro ratio	Fold increase siDrosha
hsa-miR-943	NELFA	33,11	40,84	1.23
hsa-miR-1287	PYROXD2	13,14	5,12	0.39
hsa-miR-1178	CIT	8,02	10,81	1.35
hsa-miR-636	SFRS2	0,10	0,08	0.79
hsa-miR-4315-2	PLEKHM1P	0,26	0,41	1.56
hsa-miR-4321	AMH	0,21	0,73	3.46
hsa-miR-1292	NOP56	267,34	1244,31	4.65
hsa-miR-4260	LAMB3	1,15	3,91	3.40
hsa-miR-761	NRD1	0,62	0,75	1.22
hsa-miR-555	ASH1L	14,91	29,08	1.95
hsa-miR-638	DNM2	7,54	20,82	2.76
hsa-miR-711	COL7A1	0,57	0,92	1.59
hsa-miR-937	SCRIB	0,11	0,14	1.32

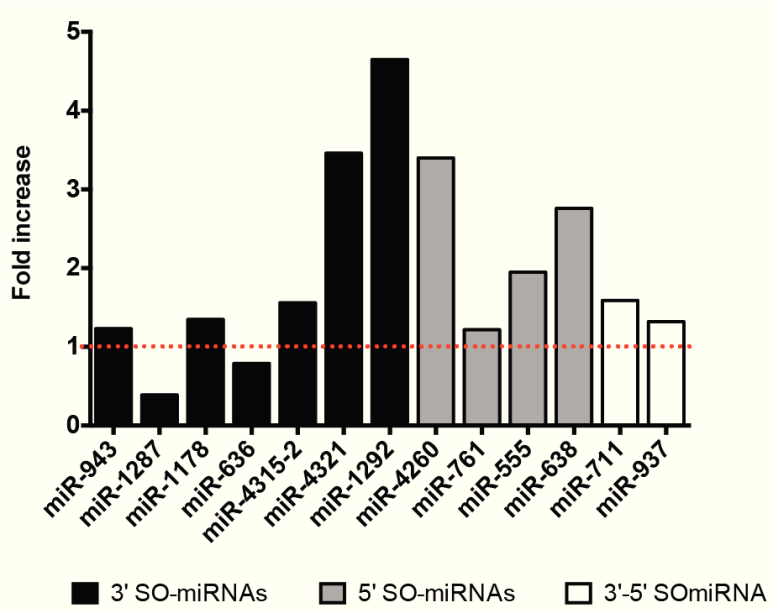


Figure 3.25: **Chromatin RNA-seq analysis of Drosha depletion effect on SO-pri-miRNA transcripts expressed in HeLa cells**

[Upper panel] Schematic representation of SO-pri-miRNA transcript. Chromatin RNA-seq reads were counted 0.4kb upstream and 0.4kb downstream of the SO-miRNA hairpin. [Middle table and lower panel] Chromatin RNA-seq analysis of SO-miRNAs expressed in HeLa cells. Fold increase of Drosha siRNA-treated cells, in comparison to siCTR-treated cells were plotted in the graph. 3', 5' and 3'-5' SO-miRNAs are indicated in black, grey and white, respectively. Red dotted line is set to 1.

Chapter 4

Discussion

4.1 miRNA biogenesis in light of pre-mRNA splicing

Cropping of pri-miRNAs and cleavage of pre-mRNAs mediated by the MPC and the spliceosome, respectively, are two processes that occur co-transcriptionally in the nucleus [4, 169] on nascent transcripts. They generate mature miRNAs and spliced mature transcripts, respectively. Most of pri-miRNA hairpins are hosted in the intronic environment, far away from splice sites, a position that could limit the interference between the two processing complexes. However, also for miRNAs located in large introns, splicing and miRNA processing can be connected and co-regulated. Some studies showed that MPC cropping of the pri-miRNA hairpins can either enhance splicing efficiency [169, 176] or, conversely, diminish its processivity [174]. On the other hand, the levels of pri-miRNA transcribed from introns can be increased in the presence of flanking exons, due to prolonged retention of the nascent transcript at the site of transcription [177], suggesting that changes in splicing can affect miRNA biosynthesis. In addition, several splicing factors can act like a bridge between splicing and miRNA biogenesis. For example, SF2/ASF, hnRNP A1 and KSRP splicing factors can promote cleavage by Drosha on specific pri-miRNA hairpins [147–149]. All together these evidences suggest a functional connection between the spliceosome and the MPC.

The purpose of my research was to study the relationship between pre-mRNA splicing and miRNA biogenesis in a newly identified group of miRNAs, the SO-miRNAs, which have an intriguing genomic location: their pri-miRNA hairpins are juxtaposed to an intron-exon junction

in their hosting transcript. I focused initially on one of these SO-miRNAs, the 3' SO-miR-34b, which derives from a non-coding transcript and its hairpin is located on the 3'ss of the last exon. Using minigene reporter assay I identified the *cis*-acting elements involved in splicing recognition and showed that the Microprocessor complex and the spliceosome compete on the nascent transcript: only the spliced RNA or the mature miRNA, but not both, can derive from the shared transcript. To better clarify the mechanism that regulates primary transcripts processing into both mature miRNA and mRNA in this novel class of SO-miRNAs, I looked more in detail at the fate of the resulting mRNAs. To this aim I performed a global analysis of miRNAs changes in a splicing inhibition context and evaluated the changes in splicing pattern and SO-miRNA expression in a more physiological context of keratinocytes differentiation. Small RNA-seq showed that, in comparison to other miRNAs, SO-miRNAs are significantly upregulated in the SF3b1 depleted cells, indicating that splicing has a direct influence on the biosynthesis of SO-miRNAs. Similarly, quantitative analysis of mRNAs and miRNAs expression levels showed that keratinocyte differentiation is associated to a decreased mRNAs expression and to an increased biosynthesis of the corresponding SO-miRNAs. My results showed also that the competition between the MPC and spliceosome processing machineries is not associated with the generation of alternative spliced transcripts at the SO-miRNA exons. Thus, the competition between splicing and SO-miRNA cleavage seems an important way to regulate the production of two alternative gene products, a mature miRNA and a full-length spliced mRNA, from a single shared transcript. My results suggest that MPC-mediated cropping of the SO-miRNA hairpins causes the premature termination of the transcript.

4.2 Identification of SO-miRNAs

To identify pri-miRNA hairpins located at the intron-exon boundary, I performed a bioinformatics analysis using an algorithm drew to compare the coordinates of annotated miRNAs with putative splice sites derived from ESTs, followed by a manual annotation to discard from the bioinformatics outputs miRNAs located on non-annotated genes, miRNAs that are transcribed in the opposite direction than the host gene and annotated mirtrons. With this approach, I have identified 52 *bona fide* SO-miRNAs (table 3.1). Considering the 1881 precursor miRNAs annotated in human genome (miRBase, release 21 - June 2014), the SO-miRNAs represent a

small, but significant portion of annotated human pri-miRNAs (2.76%). They can be located on acceptor or donor splice sites in either non-coding or protein-coding transcripts (table 3.1). The majority is placed in protein-coding transcripts (45 miRNAs) and only five in non-coding genes. Taking into account the localization of the SO-miRNAs hairpins in the genomic context, in all protein-coding transcripts and in two non-coding transcripts (miR-202 and miR-8069-2), their hairpins are located on internal exons and thus can potentially be associated to alternative splicing. Only three non-coding transcripts (miR-205, miR-34b and miR-365b) are located in the last exon, which cannot be alternatively spliced. Interestingly, there are reports describing their role in physio-pathological conditions mainly for non-coding SO-miRNAs. For example, miR-205 is both a tumor suppressor and an oncogene [186], and miR-133a-2 is fundamental for the maintenance of skeletal muscle structure [190] and in orchestrating cardiac development and function [191]. miR-34b is part of the p53 tumor suppressor network [192, 194], is involved in osteoblast proliferation [211], in pathological cardiac remodeling [210], Parkinson [216] and Huntington's diseases [217]. To clarify the functional relationship between the spliceosome and the MPC I focused on this last SO-miRNA, due to its already established role in several physio-pathological conditions. SO-miR-34b is positioned in the last exon on the junction between intron 1 and exon 2 of the non-coding BC021736 transcript.

4.3 Spliceosome and MPC have an antagonistic effect on the processing of miR-34b transcript

To clarify the functional relationship between the spliceosome and the MPC I focused on SO-miR-34b. I initially identified the key splicing regulatory elements required for the recognition of the 3'ss of pri-miR-34b. To this aim I prepared a minigene carrying the genomic region of the miR-34b hairpin (fig. 3.1). Through the generation of a series of mutant minigenes, I have identified two important *cis*-acting elements, localized within the hairpin structure and in the exon sequence, that are required for the recognition of the 3'ss of pri-miR-34b. The recognition of correct splice sites is the result of a combinatorial regulatory mechanism that involves canonical and auxiliary *cis*-acting elements [23]. In my case two elements contribute to the identification of the 3'ss: the BP and the ESE. The branch point is located in the 3'-arm of pri-miR-34b hairpin, 18 bp upstream of the 3'ss AG dinucleotide (fig. 3.3), and its sequence

perfectly matches the consensus branch point sequence [55]. Deletions or point mutations that functionally inactivate this element completely inhibit splicing (fig. 3.3). The ESE is a purine-rich exonic splicing enhancer and is located 61 bp downstream of the 3'ss (fig. 3.4). In general, purine-rich ESE binds to SR proteins, which are well-established activators of splicing, facilitating the recognition of splice sites [63]. In the future, it would be interesting to identify the *trans*-acting factor(s) that bind to this ESE and promote recognition of the upstream 3'ss. The identification of the critical *cis*-acting elements involved in the recognition of the 3'ss of pri-miR-34b was used to clarify the effect of the spliceosome on the MPC. For this I have tested splicing defective minigenes for miR-34b production by northern blot. My results clearly show that splicing inhibition obtained either affecting the AG dinucleotide of the 3'ss (fig. 3.2) or the ESE (fig. 3.4), significantly increases the production of mature miR-34b by ∼3.5-4 fold. These results provide evidence that splicing modulation affects the biogenesis of the 3' SO-miR-34b. In this experiment I could not use the splicing defect induced by disruption of the BP as mutants that affect this element also change the miRNA sequence. On the other hand, I obtained two complementary evidences that cropping inhibition affects splicing. Silencing of Drosha and DGCR8, the two major components of MPC, increases the splicing efficiency of pri-miR-34b transcript (fig. 3.5) and conversely, overexpression of these two proteins has the opposite effect, reducing splicing efficiency (fig. 3.6).

All together these experiments indicate that the MPC and the spliceosome act on the pri-miR-34b transcript in a mutually exclusive manner. The preferential recognition of the transcript by the spliceosome machinery leads to the production of the mature spliced transcript, while the mature miR-34b is not produced. On the other hand, reduction of spliceosome efficiency gives to the MPC more substrate, favoring the cropping of the pri-miRNA hairpin and the production of the mature miR-34b, but blocking the production of the mature transcript (fig. 4.1). Therefore, either blocking the spliceosome or the MPC gives to the other machinery more substrate for processing, allowing us to hypothesize that the interference between these two complexes is important for the regulation of miRNA expression.

In addition, my results indicate that the miR-34b is not processed as a 5' tailed mirtron. This hypothesis arises from the observation that the usage of the BP that I have identified will produce a lariat that binds exactly in the correspondence of the mature miR-34b sequence. Since mirtron pathway requires the debranching enzyme DBR1 to adapt a pre-miRNA fold [179], it

might be possible that the pri-miR-34b can be processed as a non canonical mirtron and that the debranching enzyme is required for the complete processing of miR-34b. However, in contrast to this hypothesis, I showed that the silencing of Drosha and DGCR8 reduced the levels of miR-34b derived from the minigene. This excludes that miR-34b is processed as a 5' tailed mirtron since this group of miRNAs does not need Drosha for its maturation [179].

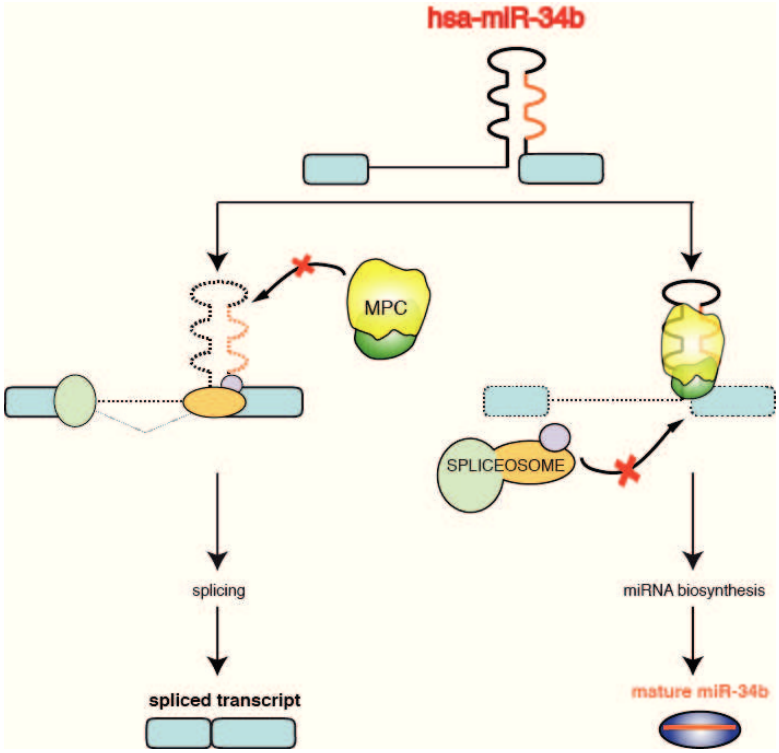


Figure 4.1: **Competition between MPC and spliceosome on non-coding 3' SO-miR-34b**
Schematic representation of the human BC021736 non-coding transcript. pri-miR-34b hairpin overlaps with the 3'ss of the last exon (exon 2) of the gene. Light blue boxes represent exons, thin line the intron and hairpin the miRNA. The mature miR-34b is highlighted in orange. [Left panel] When splicing processing prevails (spliceosome assembly on the 5' and 3'ss is schematic represented as light green, orange and violet circles) the transcript is preferentially spliced and the mature miR-34b is not produced. [Right panel] The preferential recognition of the hairpin by the MPC (Drosha (green) and DGCR8 (yellow) complex) induces the biosynthesis of the mature miR-34b (orange line) that is incorporated in the miRISC complex (blue circle). The continuos and dashed lines indicate the processed and unprocessed part of the transcript, respectively. Red ics indicate the inhibited processes.

The competition between the MPC and spliceosome in the processing of SO-miRNAs was confirmed by the work published in 2013 by the Gil Ast group [171]: they analyzed the genomic location of miRNA precursors within the hosting genes of 18 invertebrate and vertebrate

species, including human. They identified 24 human-specific SO-miRNAs, 15 of which are in common with my analysis. They focused on the mouse 3'SO-miR-412 (mmu-miR-412), which is located on an alternatively spliced exon of the non-coding Mirg gene, providing evidence the biogenesis of miR-412 is regulated by alternative splicing: exon inclusion negatively influenced the production of mature miRNA. In addition, overexpression of Drosha reduced the levels of spliced Mirg transcript. These data confirmed that the MPC and the spliceosome compete for processing of the same RNA region and that only one mature product can be generated from the shared transcript.

4.4 SO-miRNAs embedded in protein-coding transcripts

Whereas the analysis performed on miR-34b using artificial minigenes clearly shows a competition between the spliceosome and the MPC, it presents some limitations. In fact, the SO-miR-34b is positioned in the last exon of a non-coding transcript and accordingly cannot generate *in vivo* exon skipping isoforms while the majority of SO-miRNAs I have identified is embedded in protein-coding transcripts and in theory can generate alternatively spliced transcripts, being located in central exons (table 3.1). In addition, the competition between the MPC and the spliceosome on protein-coding transcripts may have an influence not only on the biosynthesis of miRNAs but also on the mRNAs and the corresponding proteins.

To answer to these questions, I looked *in vivo* at the splicing pattern of nine protein-coding genes that contain a SO-miRNA in an internal exon. I analyzed the pattern of splicing of these transcripts in five different cell lines (fig. 3.7) in order to examine if the presence of a hairpin on the splice sites can modulate splice-site selection and induce alternative splicing. Surprisingly, I did not see alternative splicing of the SO-miRNA exons that remained always included in the final transcript amongst the different cell lines. The only exception that emerged from the analysis was the NRD1 gene (see next paragraph). In parallel, I examined the amount of mature miRNAs produced and qRT-PCR analysis showed that the different cell lines produced different amounts of SO-miRNAs (fig. 3.8). This suggests either that cropping of the SO-miRNAs does not necessarily result in alternatively spliced isoforms that lack the target SO-miRNA exon or that the levels of alternative splicing are minimal and below the level of detection of the assay. In addition, even if I cannot compare the absolute Ct values of different

miRNAs, the lower number of Ct cycles for the miR-761 embedded in the NRD1 transcript suggests, at least in a semi-quantitative way, that there were more copies of miR-761 compared to all the other miRNAs.

4.5 NRD1 transcript and 5' SO-miR-761

As NRD1 was the only SO-miRNA transcript that showed some evidence of alternative splicing, I have evaluated more in detail this gene. Interestingly, in this case skipping affects not only the SO-miRNA exon 4, but also exon 3 suggesting a complex regulation (fig. 3.7, panel 10). To understand the possible functional relationship between alternative splicing pattern and miRNA biosynthesis in this gene, I tested the effect of the MPC, as well as of several splicing factors in regulating the alternative splicing of the NRD1 transcript. Co-transfection of a panel of twelve splicing factors in HeLa cells had no effect on splicing of the endogenous NRD1 exons 3 and 4 (fig. 3.9), suggesting that other *trans*-acting factors might be involved in alternative splicing regulation. On the other hand, silencing of the two MPC components Drosha and DGCR8 reduced the amount of miR-761 but did not change the alternative splicing pattern of NRD1 transcript (fig. 3.10). Thus, inhibition of MPC-dependent cropping probably provides more substrate to the spliceosome but this does not lead to the production of more full-length transcript. This suggests that also in this particular case, miRNA cropping efficiency does not influence alternative splicing of the SO-miRNA exon.

4.6 SF3b1-mediated splicing inhibition and SO-miRNAs profile

To investigate the effect of splicing on processing of the SO-miRNAs on a global basis, I decided to inhibit splicing by silencing the SF3b1 factor. SF3b1 is a well-studied spliceosome component that is involved in 3'ss recognition during the early stages of RNA splicing and its knockdown alters the fidelity of branch site recognition by the U2 snRNP [243]. I induced SF3b1 silencing in MEC-1 cells and afterwards, I performed a comprehensive analysis of miRNAs expression changes. I take advantage of the small RNA sequencing technology to analyze the differences in miRNAs expression between wild type samples and SF3b1-depleted samples. To verify if changes in pre-mRNA splicing of miRNA-containing transcripts can directly interfere with the miRNA biosynthesis, I analyzed the differential expression of miRNAs after SF3b1 knockdown:

in total 266 miRNAs are differentially expressed in the siSF3b1-treated samples, compared to untreated ones (fig. 3.13). To analyze more in detail the effect of splicing manipulation, I took into consideration the location of the miRNA hairpins in the corresponding pre-mRNA transcripts. As expected, and consistently with a competition between the spliceosome and the MPC on SO-miRNAs, miRNA expression profile in SF3b1-depleted samples revealed a significant enrichment of upregulated SO-miRNAs (fig. 3.14). SF3b1 silencing treatment positively affects the level of expression of several SO-miRNAs, suggesting that SF3b1-associated changes in pre-mRNA splicing can affect their MPC-dependent cropping. In addition, I analyzed the effect of changes in splicing efficiency on the large group of intronic and exonic miRNAs: the differential expression analysis showed a slight downregulation of intronic and exonic miRNAs after depletion of SF3b1 (fig. 3.15). These results on a global basis are consistent with previous analysis that suggested that spliceosome assembly might facilitate cropping of intronic pri-miRNA hairpins, increasing the mature miRNAs production [177]. Moreover and unexpectedly, I observed that knockdown of SF3b1 altered the production of mirtrons, with the tendency to upregulate the production of this peculiar class of miRNAs (fig. 3.16). Since the mirtron pathway utilizes splicing rather than microprocessor processing to generate intermediates of the RNAi pathway [179]. I would have expected a decrease in mirtrons processing efficiency after the knockdown of SF3b1. However, it is also possible that the MPC can efficiently process mirtrons when they are not processed by the spliceosome. Although the contribution of Drosha in pri-miRNA processing is less crucial for mirtrons rather than for canonical miRNAs, its role in the first step of miRNAs processing could not be totally excluded. Indeed, there are evidences that indicate that whereas in normal situations mirtrons are processed by the spliceosome, in a splicing-inhibition context Drosha cleavage become effective [178]. It is known that the miRNA pathway is saturable, with Drosha processing appearing to be a rate-limiting step [244]. Therefore the mirtron pathway could be emerged as a mean of generating important miRNAs using the non-saturable splicing pathway. It is highly possible that in our context, the inhibition of splicing through the depletion of SF3b1 provides more substrate to the MPC, which crops more efficiently the pri-mirtron hairpins, inducing an increase in the production of mature miRNAs.

In order to validate the results of the small RNA-seq and to confirm the central role of the pri-miRNA hairpins position across the transcripts, I prepared two hybrid minigenes derived

from PISD and KRT15 genes and containing the 3' SO-miR-7109 and 3' SO-miR-6150 hairpins and their flanking regions, respectively (fig. 3.17). Analysis by TaqMan assay of the mature miRNAs expression levels derived from the transfection of the minigenes in HeLa cells confirmed that both miRNAs are upregulated in the SF3b1 depleted samples (fig. 3.17). Moreover, I analyzed the effect of SF3b1 depletion on the splicing pattern of the two minigenes: silencing of SF3b1 increases the production of SO-miRNAs, but does not reflects in changes of alternative splicing pattern of the PISD and KRT15 transcripts (fig. 3.17, bottom panel). Valcárcel and collaborators showed that 3'ss of genes have a different sensitivity to SF3b1 knockdown and, therefore, SF3b1 depletion does not cause a general inhibition of splicing, but rather is involved only in a minor subset of alternative splicing events [243]. Thus, it is possible that either splicing of PISD and KRT15 transcripts is not SF3b1-dependent or that the effect on alternative splicing is minimal and below the sensitivity of our assay.

In the last years, several publications reported the association of defined hotspot mutations in SF3b1 with hematological tumors [43–49]. According with the well-established role of SF3b1 in splicing, mutations of this factor have been found to cause aberrant splicing of a sub-population of transcripts in sample from CLL patients [48]. At the same time, changes in the expression level of miRNAs have been detected initially in CLL [245] and subsequently in many types of human tumors [246], contributing to oncogenesis either as tumor suppressors or oncogenes. Since our evidence indicates that splicing can directly influence miRNA biosynthesis, it will be interesting to investigate the connections between splicing alteration and miRNA biogenesis in tumors associated to SF3b1 mutations. Changes in miRNAs expression levels induced by SF3b1 mutations might be relevant for the development and the progression of these types of tumors.

4.7 SO-miRNAs and keratinocytes differentiation

Among the 45 coding SO-miRNAs I have identified, miR-936, miR-4260 and miR-711 are hosted in three genes (COL17A1, LAMB3 and COL7A1, respectively) that are highly expressed at the basal layer of epidermis (fig. 3.17). The epidermis is composed of a basal layer of proliferating cells, a spinous post-mitotic cell layer, a granular layer, and a stratum corneum of terminally differentiating keratinocytes. Keratinocyte differentiation is a multistep process that requires

a highly coordinated program of gene expression. To study skin differentiation process *in vitro* I used the calcium-induced NHEK differentiation system and involucrin as a marker for differentiation. In this model involucrin expression starts to increase when cells reach confluence and increases four fold in cells after 5 days of treatment with calcium (fig. 3.19). At the same time, during differentiation I observed a decrease in the expression of COL17A1, LAMB3 and COL7A1, which are basal membrane components that mediate the adhesion of the epidermis with the underlying dermis (fig. 3.20). In order to analyze if there is a correlation between the levels of SO-miRNA hosting transcripts and the production of mature miRNAs, I analyzed through TaqMan miRNA assays the expression of two SO-miRNAs, miR-936 and miR-4260 contained in COL17A1 and LAMB3 gene at the different time points of differentiation. In my experimental condition I could not detect miR-711. As shown in figure 3.21, I observed an inverse correlation between the expression of the SO-miRNA hosting transcripts and the expression of the corresponding mature miRNAs: during differentiation COL17A1 and LAMB3 are downregulated, whereas the corresponding SO-miRNAs are upregulated. As a control, the levels of intronic miR-330 and the amount of Drosha were not significantly affected during keratinocytes differentiation (fig. 3.22).

To go further into the biological meaning of this switch, I overexpressed miR-936, miR-4260 and miR-711 by miRNA mimics. I tested the effect of miRNAs overexpression on the proliferation of the immortal human keratinocyte line HaCaT that is commonly used in scientific research for its high capacity to differentiate and proliferate *in vitro*. Consistent with their role in regulating keratinocyte differentiation, the EdU proliferation assay showed that the three SO-miRNAs have an anti-proliferative effect on keratinocytes (fig. 3.24). The inverse correlation between the expression of COL17A1 and LAMB3 transcripts and their corresponding embedded SO-miRNAs let us hypothesize the existence of a regulatory switch that controls keratinocytes transition from the basal to the cornified layer. In the basal layer, where keratinocytes proliferate, there is a high expression of COL17A1, LAMB3 and COL7A1 proteins and a low expression of corresponding miRNAs, possibly due to a low MPC-dependent cropping of the SO-miRNAs. The activation of the differentiation program shifts the balance toward a preferential cropping of the nascent transcripts that increases the production of mature SO-miRNAs, dropping the expression of the mature transcripts.

A regulatory switch between two alternative gene products has been recently described in the

FSTL1 mRNA and is fundamental to orchestrate keratinocyte migration, re-epithelialization and wound healing [184]. In this case in the healthy epidermis the FSTL1 primary transcript is preferentially processed, in a KSRP-mediated manner, by the MPC that crops the pri-miR-198 hairpin, located in the 3'UTR of the transcript. When epidermis is wounded, the downregulation of KSRP by the TGF-β factor drops off the expression of miR-198, switching on the production of FSTL1 protein and other wound healing factors.

More in general, miRNAs play an important role in regulating keratinocyte proliferation and differentiation. Their global inhibition *in vivo* by epithelium specific depletion of Dicer or DGCR8 genes leads to barrier defects, abnormal hair follicle development, and hyper proliferation of basal interfollicular keratinocytes [247–249]. miR-203 promotes epidermal differentiation by restricting proliferative potential and inducing cell-cycle exit by silencing p63, which is an essential regulator of stem-cell maintenance in stratified epithelial tissue [250]. In addition other miRNAs have been found to be associated with human keratinocyte differentiation *in vitro* and *in vivo* [251]. The three miRNAs I have identified here can be inserted in the list of miRNAs that regulate keratinocyte differentiation.

4.8 Does the MPC induce premature transcriptional termination of SO-miRNA transcripts?

I observed that the vast majority of SO-miRNA exons does not undergo alternative splicing *in vivo*: in different cell types, in MEC-1 cells after SF3b1 splicing inhibition and also during keratinocyte differentiation I could not observe changes in alternative splicing pattern. In keratinocytes, the increased biosynthesis of miR-936 and miR-4260 had no apparent effect on the inclusion of the exons that contain part of the pri-miRNA hairpins. Indeed, although the levels of expression of the two miRNAs significantly increase during differentiation, exon 29 of COL17A1 and exon 15 of LAMB3 are always included in the final transcript (fig. 3.23). There are two possibilities that could explain the absence of detection of alternative spliced RNA isoforms: either the amount of alternative splicing is very low and below the detection limit of our experimental assay or the SO-miRNAs may function as dead-end processing signal inside genes. To explore the hypothesis that SO-miRNA hairpins represent potential site of transcription termination inside genes, I took advantage from the available chromatin RNA sequencing

data from HeLa cells [239]. I found that Drosha depletion leads to extensive transcriptional readthrough downstream of the pri-miRNA hairpins in eleven SO-miRNA transcripts out of the thirteen that are expressed in HeLa cells (fig. 3.25). These results suggest the possibility that MPC cropping of SO-pri-miRNA hairpins may be important for transcriptional termination inside the transcripts. The MPC-mediated cleavage generates the mature SO-miRNAs and shorter, non-polyadenylated transcripts that are unstable and, therefore, rapidly degraded. As suggested for lncRNA transcripts [239], this non canonical transcriptional termination at SO-pri-miRNA sites may represent a way to regulate the production of two alternative gene products from a single shared transcript, allowing the production of high amount of mature SO-miRNAs, without the concomitant generation of high levels of host transcripts.

The data presented in this thesis support the view that when the miRNA hairpins overlap with the splice sites the competition between the MPC and the spliceosome processing machineries regulates the production of mature mRNA and miRNA from a shared precursor RNA transcript. On the nascent transcript the relative activities of the spliceosome and the MPC will determine the execution of one process rather than the other: in the presence of more splicing enhancing regulatory factors, the spliceosome will prevail and the transcript will be preferentially spliced. On the other hand, an increase in the relative activity of the MPC will cause the production of mature miRNA and the generation of a dead-end transcript (fig. 4.2).

In general, transcription termination of mRNA is normally achieved after polyadenylation by an XRN2-dependent 5'-3' exoribonuleolytic degradation of the downstream transcript. Polyadenylation sites are normally located at end of genes but in several cases they can also be present inside genes as alternative poly(A) sites. In the last years a series of alternative transcriptional termination mechanisms have been described. For instance, lncRNA MALAT1 termination is mediated by enzymes normally involved in tRNA processing [252], histone mRNAs possess a hairpin structure and a downstream purine-rich element that recruit the U7 snRNP to terminate transcription [253]. In yeast, RNA transcripts are processed by the RNase III-type enzyme Rnt1 and do not require polyadenylation signals for termination [254]. An additional non-canonical mechanism of transcriptional termination involves the human RNase III-type enzyme Drosha [239]. Similar to the mechanism we are proposing here for SO-miRNA hairpins, the authors showed that the MPC could mediate transcriptional termination. This effect was reported to occur mainly on miRNA hairpins located in the last exon of non-coding genes, just

upstream of the poly(A) site [239]. By sequencing analysis of chromatin-associated RNA from HeLa cells, they identified defects in transcriptional termination after MPC depletion that led to extensive transcriptional readthrough of lnc-pri-miRNA transcripts. In addition, a previous study showed that Drosha processing of pri-miRNA hairpins can attenuate downstream transcription by providing an entry site for XRN2 [255].

In the big group of protein-coding SO-miRNA transcripts, the peculiar position of the pri-miRNA imposes an additional level of transcriptional regulation. The competition between the spliceosome and the MPC apparently does not produce alternatively spliced transcripts, but it is a way to regulate the production of a mature miRNA or a mature mRNA from a shared transcript. When the transcript is preferentially recognized by the spliceosome, the mature mRNA is produced. On the other hand, if the pri-miRNA is a preferential substrate for the MPC, the mature miRNA is produced, but the cropping of the pri-miRNA hairpin serves as a transcriptional termination signal inside the transcript, that does not produce an alternative spliced mRNA, but instead a shorter, non-stable and non-polyadenylated transcript, with the 3'-end defined by Drosha cleavage, that is retained in the nucleus and rapidly degraded (fig. 4.2). Therefore, the MPC-mediated transcriptional termination replaces the canonical cleavage and polyadenylation pathway, inducing the production of mature miRNAs, but preventing the accumulation of potentially undesirable high amounts of mRNAs. In addition, since some SO-miRNAs are found in cluster with other pri-miRNA hairpins, it will be interesting to investigate if the MPC-mediated transcriptional termination at the level of the SO-miRNAs may affect the processing and the relative expression of the clustered miRNAs, providing an additional level of miRNA biogenesis regulation.

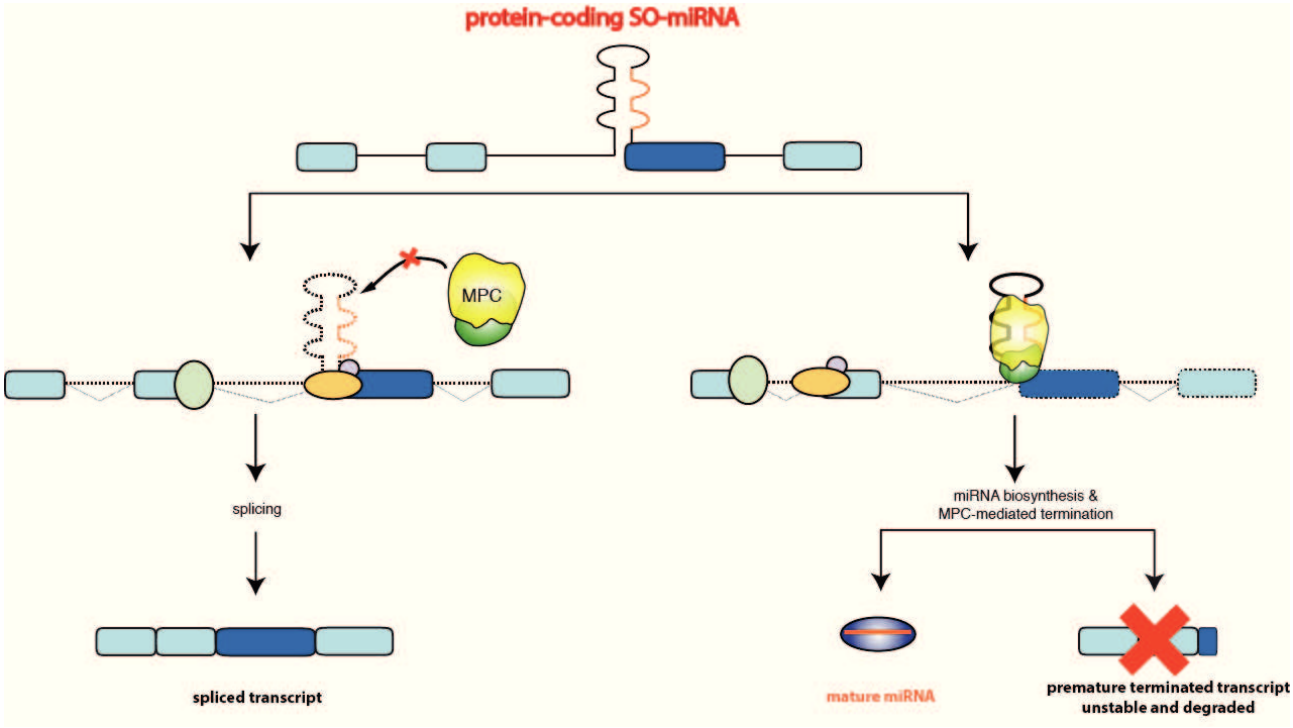


Figure 4.2: **Competition between MPC and spliceosome on protein-coding SO-miRNAs**
Schematic representation of a SO-miRNA overlapping within a protein-coding transcript. Exons flanking the SO-miRNA and the exon containing part of the pri-miRNA hairpin are represented as light blue and blue boxes, respectively. Black thin line and hairpin represent the intron and the miRNA, respectively. The mature miRNA is highlighted in orange. [Left panel] When splicing processing prevails (spliceosome assembly on the 5' and 3'ss is schematic represented as light green, orange and violet circles) the transcript is preferentially spliced and the mature SO-miRNA is not produced. [Right panel] The preferential recognition of the hairpin by the MPC (Drosha (green) and DGCR8 (yellow) complex) induces the biosynthesis of the mature miRNA (orange line) that is incorporated in the miRISC complex (blue circle). In addition, MPC cropping of the SO-pri-miRNA hairpin mediates the premature termination at the site of cropping of the host transcript that is non-polyadenylated, unstable and, therefore, rapidly degraded. The continuous and dashed lines indicate the processed and unprocessed part of the transcript, respectively.

Appendix 1: List of dowregulated miRNAs in SF3b1-depleted MEC-1 cells

hsa miRNA	Deseq2 log_2FC	Deseq2 FDR
hsa-miR-7974	-3,402	1,98E-49
hsa-miR-155-3p	-2,954	5,43E-20
hsa-miR-4454	-2,833	5,67E-49
hsa-miR-29b-1-5p	-2,729	3,37E-19
hsa-miR-573	-2,409	2,24E-18
hsa-miR-4455	-2,358	2,14E-16
hsa-miR-148a-5p	-2,351	3,01E-25
hsa-miR-23a-5p	-2,263	1,74E-19
hsa-miR-27b-5p	-2,237	3,01E-25
hsa-miR-608	-2,192	3,11E-14
hsa-miR-27a-5p	-2,181	3,43E-10
hsa-miR-92a-1-5p	-2,174	4,35E-18
hsa-miR-92b-3p	-2,160	3,37E-19
hsa-miR-32-3p	-2,112	1,80E-09
hsa-miR-3692-3p	-2,109	4,62E-08
hsa-miR-558	-2,100	3,93E-07
hsa-miR-3122	-2,064	1,43E-13
hsa-miR-5194	-2,061	3,16E-07
hsa-miR-20b-5p	-2,026	1,52E-11
hsa-miR-7977	-2,017	2,66E-13
hsa-miR-98-5p	-1,961	4,03E-09
hsa-miR-6773-3p	-1,936	1,13E-06
hsa-miR-4729	-1,919	1,85E-08
hsa-miR-4483	-1,893	1,42E-08
hsa-miR-30b-3p	-1,885	1,06E-06
hsa-miR-941	-1,885	1,04E-07
hsa-miR-21-3p	-1,857	2,38E-08
hsa-miR-671-3p	-1,841	1,79E-21
hsa-miR-374b-3p	-1,835	2,14E-08
hsa-miR-149-5p	-1,817	2,25E-05
hsa-let-7b-3p	-1,813	2,14E-07
hsa-miR-3935	-1,776	2,23E-08
hsa-miR-1247-3p	-1,757	9,40E-12
hsa-miR-6715b-3p	-1,732	2,40E-07
hsa-miR-3617-3p	-1,730	5,53E-07
hsa-miR-744-5p	-1,726	4,71E-12
hsa-miR-3529-3p	-1,722	5,91E-07
hsa-miR-148a-3p	-1,722	1,15E-09
hsa-miR-374a-3p	-1,717	1,08E-07
hsa-miR-30d-3p	-1,699	1,42E-06
hsa-miR-1272	-1,694	3,47E-07
hsa-miR-766-3p	-1,684	5,16E-09
hsa-miR-335-3p	-1,681	2,08E-05
hsa-miR-26a-2-3p	-1,677	4,60E-06
hsa-miR-3136-3p	-1,647	7,86E-06
hsa-miR-20a-5p	-1,642	5,68E-05
hsa-miR-6835-5p	-1,626	4,60E-06
hsa-miR-4518	-1,622	1,98E-07
hsa-let-7e-5p	-1,608	1,30E-06
hsa-let-7f-1-3p	-1,607	2,71E-07
hsa-miR-5589-5p	-1,604	1,52E-04
hsa-miR-645	-1,604	1,31E-04
hsa-miR-17-5p	-1,603	4,48E-05
hsa-miR-6834-3p	-1,599	2,03E-06
hsa-miR-548aa	-1,598	5,04E-05
hsa-miR-98-3p	-1,597	1,12E-07
hsa-miR-212-5p	-1,591	1,58E-05
hsa-miR-151a-5p	-1,589	1,08E-06
hsa-miR-3184-3p	-1,587	1,33E-07
hsa-miR-378a-5p	-1,568	4,84E-08
hsa-miR-143-3p	-1,566	5,39E-07
hsa-miR-16-2-3p	-1,563	2,62E-06
hsa-miR-423-5p	-1,555	9,47E-08
hsa-let-7d-5p	-1,551	4,29E-09
hsa-miR-1268a	-1,539	4,67E-09
hsa-let-7a-5p	-1,529	4,62E-06
hsa-miR-5188	-1,525	4,26E-05
hsa-miR-92a-3p	-1,521	5,99E-06
hsa-miR-4743-5p	-1,512	2,14E-08
hsa-miR-18a-3p	-1,510	9,67E-07
hsa-miR-486-5p	-1,505	5,57E-08

miRNA	Deseq2 log_2FC	Deseq2 FDR
hsa-miR-4801	-1,498	8,51E-04
hsa-let-7f-5p	-1,495	1,86E-06
hsa-miR-5000-3p	-1,485	1,09E-04
hsa-miR-27b-3p	-1,482	1,90E-08
hsa-miR-4443	-1,478	5,09E-04
hsa-miR-106a-5p	-1,474	4,02E-04
hsa-miR-3150a-5p	-1,474	5,41E-05
hsa-miR-1260b	-1,466	6,28E-04
hsa-miR-138-5p	-1,462	1,31E-04
hsa-miR-181b-3p	-1,462	6,18E-06
hsa-miR-548x-5p	-1,455	3,82E-04
hsa-miR-589-3p	-1,452	5,80E-07
hsa-miR-3150b-3p	-1,450	6,93E-04
hsa-miR-562	-1,446	1,08E-03
hsa-miR-454-5p	-1,439	8,07E-05
hsa-miR-744-3p	-1,428	8,67E-04
hsa-miR-3186-5p	-1,427	5,33E-05
hsa-let-7g-5p	-1,424	3,62E-04
hsa-miR-3613-3p	-1,407	1,08E-03
hsa-miR-6782-3p	-1,403	1,58E-04
hsa-let-7f-2-3p	-1,402	8,08E-04
hsa-miR-4284	-1,394	1,52E-04
hsa-miR-130b-5p	-1,392	1,24E-03
hsa-miR-6832-3p	-1,386	6,14E-04
hsa-miR-20b-3p	-1,371	1,19E-03
hsa-miR-421	-1,364	2,23E-05
hsa-miR-942-5p	-1,358	5,69E-04
hsa-miR-371b-5p	-1,356	9,09E-05
hsa-miR-873-3p	-1,354	1,17E-03
hsa-miR-7158-3p	-1,349	2,44E-03
hsa-miR-589-5p	-1,346	5,39E-04
hsa-miR-7-5p	-1,341	2,55E-04
hsa-miR-653-3p	-1,331	1,52E-03
hsa-miR-186-5p	-1,331	1,80E-04
hsa-miR-548b-5p	-1,330	6,42E-04
hsa-miR-3074-5p	-1,328	1,99E-04
hsa-miR-26b-3p	-1,325	4,62E-08
hsa-miR-340-5p	-1,325	1,69E-05
hsa-miR-7706	-1,325	8,51E-04
hsa-miR-500a-3p	-1,322	7,97E-05
hsa-miR-1260a	-1,320	3,04E-03
hsa-miR-5195-3p	-1,317	3,29E-05
hsa-miR-17-3p	-1,315	1,83E-03
hsa-miR-6851-3p	-1,314	3,22E-06
hsa-miR-6871-3p	-1,311	1,24E-03
hsa-miR-548c-5p	-1,310	1,99E-03
hsa-miR-330-5p	-1,301	2,16E-03
hsa-miR-629-5p	-1,297	9,74E-05
hsa-miR-125a-5p	-1,291	4,41E-03
hsa-miR-30e-3p	-1,291	4,14E-05
hsa-miR-15a-3p	-1,285	1,87E-03
hsa-miR-9-5p	-1,285	1,36E-07
hsa-miR-484	-1,278	2,28E-03
hsa-miR-1303	-1,278	4,81E-04
hsa-miR-30c-1-3p	-1,277	6,13E-07
hsa-miR-708-3p	-1,271	1,69E-03
hsa-miR-4448	-1,261	3,88E-03
hsa-miR-588	-1,258	2,77E-03
hsa-miR-8077	-1,256	1,28E-03
hsa-miR-548o-3p	-1,255	3,82E-04
hsa-miR-1269b	-1,254	5,65E-03
hsa-miR-18a-5p	-1,251	1,78E-03
hsa-miR-7-1-3p	-1,244	1,61E-03
hsa-miR-155-5p	-1,237	2,30E-06
hsa-miR-664b-3p	-1,236	2,60E-03
hsa-miR-10a-5p	-1,233	4,00E-04
hsa-miR-339-3p	-1,233	1,36E-03
hsa-miR-561-5p	-1,230	6,69E-03
hsa-miR-548g-5p	-1,229	3,08E-03
hsa-miR-1273h-3p	-1,228	5,76E-04
hsa-miR-133a-3p	-1,228	5,36E-06

miRNA	Deseq2 log_2FC	Deseq2 FDR
hsa-miR-4475	-1,227	5,08E-03
hsa-miR-363-3p	-1,227	6,42E-04
hsa-miR-24-3p	-1,226	1,36E-03
hsa-miR-374a-5p	-1,210	3,81E-03
hsa-let-7b-5p	-1,205	2,52E-06
hsa-miR-30a-5p	-1,197	5,89E-03
hsa-miR-660-5p	-1,194	4,96E-04
hsa-miR-33a-5p	-1,194	5,08E-03
hsa-miR-93-5p	-1,191	1,07E-03
hsa-miR-499b-3p	-1,188	5,35E-03
hsa-miR-93-3p	-1,187	2,14E-03
hsa-miR-6742-3p	-1,182	3,81E-03
hsa-miR-499a-5p	-1,179	4,90E-03
hsa-miR-548k	-1,177	2,77E-03
hsa-let-7c-5p	-1,175	1,77E-06
hsa-miR-25-5p	-1,174	3,09E-05
hsa-miR-642b-5p	-1,173	7,33E-04
hsa-miR-4731-3p	-1,171	1,72E-04
hsa-miR-1268b	-1,164	8,52E-06
hsa-miR-450b-5p	-1,163	1,57E-03
hsa-miR-769-5p	-1,163	1,26E-03
hsa-miR-501-3p	-1,161	9,07E-03
hsa-miR-181a-2-3p	-1,159	3,19E-05
hsa-miR-186-3p	-1,159	7,07E-03
hsa-miR-361-5p	-1,150	1,28E-03
hsa-miR-15b-5p	-1,150	7,78E-03
hsa-miR-183-5p	-1,149	4,41E-03
hsa-miR-3128	-1,139	9,61E-03
hsa-miR-28-3p	-1,138	9,52E-05
hsa-miR-1287-5p	-1,132	1,48E-03
hsa-miR-365a-3p	-1,130	8,55E-03
hsa-miR-331-5p	-1,126	3,63E-04
hsa-miR-378h	-1,126	1,26E-03
hsa-miR-200c-3p	-1,123	1,45E-04
hsa-miR-3182	-1,120	7,10E-03
hsa-miR-378a-3p	-1,117	4,82E-04
hsa-miR-222-5p	-1,115	3,96E-04
hsa-miR-146b-3p	-1,114	6,88E-04
hsa-miR-3168	-1,110	7,07E-03
hsa-miR-128-1-5p	-1,108	1,42E-03
hsa-miR-6759-3p	-1,098	6,93E-04
hsa-miR-2467-5p	-1,090	3,63E-04
hsa-miR-502-3p	-1,086	3,29E-03
hsa-miR-6846-5p	-1,084	6,32E-03
hsa-miR-4668-5p	-1,081	7,04E-03
hsa-miR-210-3p	-1,079	1,04E-03
hsa-miR-6799-3p	-1,072	1,97E-03
hsa-miR-185-5p	-1,065	1,24E-03
hsa-miR-548e-5p	-1,062	7,05E-03
hsa-miR-342-3p	-1,058	1,45E-04
hsa-miR-3937	-1,055	3,40E-04
hsa-miR-5705	-1,053	1,04E-03
hsa-miR-574-3p	-1,052	1,09E-04
hsa-miR-1972	-1,050	3,95E-03
hsa-miR-103a-3p	-1,047	8,73E-04
hsa-miR-4428	-1,038	1,85E-03
hsa-miR-574-5p	-1,037	6,88E-04
hsa-miR-503-5p	-1,036	6,34E-03
hsa-miR-132-3p	-1,035	6,28E-04
hsa-miR-6735-3p	-1,030	2,22E-03
hsa-miR-130b-3p	-1,028	9,10E-04
hsa-let-7i-5p	-1,027	9,92E-04
hsa-miR-1285-3p	-1,024	5,69E-04
hsa-miR-126-5p	-1,021	8,46E-03
hsa-miR-378c	-1,020	1,23E-03
hsa-miR-598-3p	-1,015	2,04E-03
hsa-miR-133b	-1,012	8,88E-04
hsa-miR-150-3p	-1,008	5,80E-04
hsa-miR-148b-3p	-1,006	9,93E-03
hsa-miR-101-3p	-1,000	3,22E-03

Appendix 2: List of upregulated miRNAs in SF3b1-depleted MEC-1 cells

hsa miRNA	Deseq2 log_2FC	Deseq2 FDR
hsa-miR-3196	1,004	6,92E-05
hsa-miR-3622a-5p	1,009	9,94E-03
hsa-miR-1471	1,018	3,29E-04
hsa-miR-550b-3p	1,018	4,41E-03
hsa-miR-631	1,019	1,58E-04
hsa-miR-219a-2-3p	1,024	5,50E-03
hsa-miR-489-3p	1,029	1,10E-03
hsa-miR-4781-5p	1,040	1,11E-03
hsa-miR-4707-3p	1,058	6,28E-04
hsa-miR-1825	1,063	1,12E-03
hsa-miR-6831-5p	1,064	1,36E-03
hsa-miR-8058	1,065	6,15E-03
hsa-miR-1236-3p	1,075	8,66E-03
hsa-miR-4512	1,081	1,71E-05
hsa-miR-3161	1,082	8,90E-03
hsa-miR-7155-5p	1,094	1,35E-03
hsa-miR-1976	1,115	6,71E-03
hsa-miR-1910-3p	1,117	4,85E-03
hsa-miR-7107-5p	1,125	4,53E-03
hsa-miR-6879-5p	1,128	1,28E-03
hsa-miR-6732-3p	1,142	3,11E-03
hsa-miR-320c	1,142	5,18E-03
hsa-miR-6726-3p	1,149	5,76E-04
hsa-miR-4749-3p	1,153	3,08E-03
hsa-miR-6131	1,153	9,78E-05
hsa-miR-2116-3p	1,171	1,15E-03
hsa-miR-3614-3p	1,172	1,05E-03

miRNA	Deseq2 log_2FC	Deseq2 FDR
hsa-miR-1281	1,172	5,32E-03
hsa-miR-4632-3p	1,195	1,43E-06
hsa-miR-6776-3p	1,199	4,93E-03
hsa-miR-4651	1,216	4,60E-04
hsa-miR-337-3p	1,233	1,36E-03
hsa-miR-5047	1,239	2,45E-03
hsa-miR-6508-5p	1,242	1,13E-04
hsa-miR-3141	1,284	6,42E-04
hsa-miR-6769a-3p	1,291	1,21E-04
hsa-miR-3656	1,307	5,33E-07
hsa-miR-6751-3p	1,318	3,67E-05
hsa-miR-5003-3p	1,337	2,87E-03
hsa-miR-4474-5p	1,350	1,36E-04
hsa-miR-6886-5p	1,353	6,04E-08
hsa-miR-550a-3-5p	1,380	6,73E-04
hsa-miR-3655	1,408	5,91E-06
hsa-miR-483-3p	1,411	1,01E-04
hsa-miR-4426	1,417	4,98E-04
hsa-miR-3177-3p	1,420	2,58E-06
hsa-miR-6090	1,427	5,33E-07
hsa-miR-4741	1,457	8,48E-11
hsa-miR-8069	1,671	9,81E-11
hsa-miR-205-3p	1,812	4,71E-06
hsa-miR-4800-3p	1,888	2,16E-09
hsa-miR-4516	2,015	8,53E-15
hsa-miR-5088-3p	2,030	1,52E-11
hsa-miR-1236-5p	2,511	5,04E-15

Bibliography

[1] T Maniatis and R Reed. An extensive network of coupling among gene expression machines. *Nature*, 416(6880):499–506, 2002.

[2] DL Bentley. Rules of engagement: co-transcriptional recruitment of pre-mRNA processing factors. *Current Opinion in Cell Biology*, 17(3):251–256, 2005.

[3] R Reed. Coupling transcription, splicing and mRNA export. *Current Opinion in Cell Biology*, 15(3):326–331, 2003.

[4] AR Kornblihtt. Multiple links between transcription and splicing. *RNA*, 10(10):1489–1498, 2004.

[5] CA Beelman and R Parker. Degradation of mRNA in eukariotes. *Cell*, 81(2):179–183, 1995.

[6] JD Lewis, E Izaurralde, A Jarmolowski, C McGuigan, and IW Mattaj. A nuclear cap-binding complex facilitates association of U1 snRNP with the cap-proximal 5' splice site. *Genes & Development*, 10(13):1683–1698, 1996.

[7] JD Gross, NJ Moerke, T von der Haar, AA Lugovskoy, AB Sachs, JE McCarthy, and G Wagner. Ribosome loading onto the mRNA cap is driven by conformational coupling between eIF4G and eIF4E. *Cell*, 115(6):739–750, 2003.

[8] E Beaudoing, S Freier, JR Wyatt, JM Claverie, and D Gautheret. Patterns of variant polyadenylation signal usage in human genes. *Genome Research*, 10(7):1001–1010, 2000.

[9] S West, N Gromak, and NJ Proudfoot. Human 5' $\rightarrow$ 3' exonuclease Xrn2 promotes transcription termination at co-transcriptional cleavage sites. *Nature*, 432(7016):522–525, 2004.

[10] SM Berget, C Moore, and PA Sharp. Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *PNAS*, 74(8):3171–3175, 1977.

[11] LT Chow, RE Gelinas, TR Broker, and JR Roberts. An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell*, 12(1):1–8, 1977.

[12] AJ Jeffreys and RA Flavell. The rabbit β-globin gene contains a large insert in the coding sequence. *Cell*, 12(4):1097–1108, 1977.

[13] R Breathnach, C Benoist, K O'Hare, F Gannon, and P Chambon. Ovalbumin gene: evidence for a leader sequence in mRNA and DNA sequences at the exon-intron boundaries. *PNAS*, 75(10):4853–4857, 1978.

[14] JL Mandel, R Breathnach, P Gerlinger, M Le Meur, F Gannon, and P Chambon. Organization of coding and intervening sequences in the chicken ovalbumin split gene. *Cell*, 14(3):641–653, 1978.

[15] TW Nilsen and BR Graveley. Expansion of the eukaryotic proteome by alternative splicing. *Nature*, 463(7280):457–463, 2010.

[16] RA Padgett, PJ Grabowski, MM Konarska, S Seiler, and PA Sharp. Splicing of messenger RNA precursors. *Annual Review of Biochemistry*, 55:1119–1150, 1986.

[17] AI Lamond and DL Spector. Nuclear speckles: a model for nuclear organelles. *Nature Reviews Molecular Cell Biology*, 4(8):605–612, 2003.

[18] T Misteli, JF Cáceres, and DL Spector. The dynamics of a pre-mRNA splicing factor in living cells. *Nature*, 387(6632):523–527, 1997.

[19] T Misteli and DL Spector. Protein phosphorylation and the nuclear organization of pre-mRNA splicing. *Trends in Cell Biology*, 7(4):135–138, 1997.

[20] E Meshorer and T Misteli. Splicing misplaced. *Cell*, 122(3):317–318, 2005.

[21] J Glanzer, KY Miyashiro, JY Sul, L Barrett, B Belt, P Haydon, and J Eberwine. RNA splicing capability of live neuronal dendrites. *PNAS*, 102(46):16859–16864, 2005.

[22] MM Denis, ND Tolley, M Bunting, H Schwertz, H Jiang, S Lindemann, CC Yost, FJ Rubner, KH Albertine, KJ Swoboda, CM Fratto, E Tolley, LW Kraiss, TM McIntyre, GA Zimmerman, and AS Weyrich. Escaping the nuclear nonfines: signal-dependent pre-mRNA splicing in anucleate platelets. *Cell*, 122(3):379–391, 2005.

[23] L Cartegni, SL Chew, and AR Krainer. Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nature Reviews Genetics*, 3(4):285–298, 2002.

[24] MJ Moore and PA Sharp. Evidence for two active sites in the spliceosome provided by stereochemistry of pre-mRNA splicing. *Nature*, 365(6444):364–368, 1993.

[25] DS Horowitz. The mechanism of the second step of pre-mRNA splicing. *WIREs RNA*, 3(3):331–350, 2012.

[26] J Rino and M Carmo-Fonseca. The spliceosome: a self-organized macromolecular machine in the nucleus? *Trends in Cell Biology*, 19(8):375–384, 2009.

[27] MC Wahl, CL Will, and R Lührmann. The spliceosome: design principles of a dynamic RNP machine. *Cell*, 136(4):701–718, 2009.

[28] J Rappsilber. Large-scale proteomic analysis of the human spliceosome. *Genome Research*, 12(8):1231–1245, 2002.

[29] MS Jurica and MJ Moore. Pre-mRNA splicing. *Molecular Cell*, 12(1):5–14, 2003.

[30] SB Patel and M Bellini. The assembly of a spliceosomal small nuclear ribonucleoprotein particle. *Nucleic Acids Research*, 36(20):6482–6493, 2008.

[31] CL Will, S Rümpler, JK Gunnewiek, WJ van Venrooij, and R Lührmann. In vitro reconstitution of mammalian U1 snRNPs active in splicing: the U1-C protein enhances the formation of early (E) spliceosomal complexes. *Nucleic Acids Research*, 24(23):4614–4623, 1996.

[32] J Valcárcel, RK Gaur, R Singh, and MR Green. Interaction of U2AF65 RS region with pre-mRNA of branch point and promotion base pairing with U2 snRNA. *Science*, 273(5282):1706–1709, 1996.

[33] S Wu, CM Romfo, TW Nilsen, and MR Green. Functional recognition of the 3' splice site AG by the splicing factor U2AF35. *Nature*, 402(6763):832–835, 1999.

[34] JA Berglund, N Abovich, and M Rosbash. A cooperative interaction between U2AF65 and mBBP/SF1 facilitates branchpoint region recognition. *Genes & Development*, 12(6):858–867, 1998.

[35] O Gozani, R Feld, and R Reed. Evidence that sequence-independent binding of highly conserved U2 snRNP proteins upstream of the branch site is required for assembly of spliceosomal complex A. *Genes & Development*, 10(2):233–243, 1996.

[36] IA Turner, CM Norman, MJ Churcher, and AJ Newman. Roles of the U5 snRNP in spliceosome dynamics and catalysis. *Biochemical Society Transactions*, 32(6):928–931, 2004.

[37] B Schwer and CH Gross. Prp22, a DExH-box RNA helicase, plays two distinct roles in yeast pre-mRNA splicing. *The EMBO Journal*, 17(7):2086–2094, 1998.

[38] JO Ilagan, RJ Chalkley, AL Burlingame, and MS Jurica. Rearrangements within human spliceosomes captured after exon ligation. *RNA*, 19(3):400–412, 2013.

[39] CL Will and R Lührmann. Spliceosome structure and function. In *The RNAworld,* 3rd ed. (ed. R.F. Gesteland et al.) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pages 369–400, 2006.

[40] J Wu and JL Manley. Mammalian pre-mRNA branch site selection by U2 snRNP involves base pairing. *Genes & Development*, 3(10):1553–1561, 1989.

[41] O Gozani, J Potashkin, and R Reed. A potential role for U2AF–SAP 155 interactions in recruiting U2 snRNP to the branch site. *Molecular and Cellular Biology*, 18(8):4752–4760, 1998.

[42] CC Query, SA Strobel, and PA Sharp. Three recognition events at the branch-site adenine. *The EMBO Journal*, 15(6):1392–1402, 1996.

[43] E Papaemmanuil, M Cazzola, J Boultwood, L Malcovati, P Vyas, D Bowen, et al. Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. *WIREs RNA*, 365(15):1384–1395, 2011.

[44] K Yoshida, Y Sanada, M adn Shiraishi, D Nowak, Y Nagata, R Yamamoto, et al. Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature*, 478(7367):64–69, 2011.

[45] JW Harbour, ED Roberson, H Anbunathan, MD Onken, LA Worley, and AM Bowcock. Recurrent mutations at codon 625 of the splicing factor SF3B1 in uveal melanoma. *Nature Genetics*, 45(2):133–135, 2013.

[46] SJ Furney, M Pedersen, D Gentien, AG Dumont, A Rapinat, L Desjardins, et al. SF3B1 mutations are associated with alternative splicing in uveal melanoma. *Cancer discovery*, 3(10):1122–1129, 2013.

[47] K Yoshida and S Ogawa. Splicing factor mutations and cancer. *WIREs RNA*, 5(4):445–459, 2014.

[48] L Wang, MS Lawrence, Y Wan, P Stojanov, C Sougnez, K Stevenson, et al. SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. *The New England Journal of Medicine*, 365(26):2497–2506, 2011.

[49] V Quesada, L Conde, N Villamor, GR Ordonez, P Jares, L Bassaganyas, et al. Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. *The New England Journal of Medicine*, 44(1):47–52, 2011.

[50] Y Wan and CJ Wu. SF3B1 mutations in chronic lymphocytic leukemia. *Blood*, 121(23):4627–4634, 2013.

[51] MB Shapiro and P Senapathy. RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. *Nucleic Acids Research*, 15(17):7155–7174, 1987.

[52] CJ Langford, FJ Klinz, C Donath, and D Gallwitz. Point mutations identify the conserved, intron-contained TACTAAC box as an essential splicing signal sequence in yeast. *Cell*, 36(3):645–653, 1984.

[53] M Aebi, H Hornig, and C Weissmann. 5' cleavage site in eukaryotic pre-mRNA splicing is determined by the overall 5' splice region, not by the conserved 5' GU. *Cell*, 50(2):237–246, 1987.

[54] R Reed. The organization of 3' splice-site sequences in mammalian introns. *Genes & Development*, 3(12B):2113–2123, 1989.

[55] R Reed and T Maniatis. The role of the mammalian branchpoint sequence in pre-mRNA splicing. *Genes & Development*, 2(10):1268–1276, 1988.

[56] RF Roscigno, M Weiner, and MA Garcia-Blanco. A mutational analysis of the polypyrimidine tract of introns. Effects of sequence differences in pyrimidine tracts on splicing. *The Journal of Biological Chemistry*, 268(15):11222–11229, 1993.

[57] H Sun and LA Chasin. Multiple splicing defects in an intronic false exon. *Molecular and Cellular Biology*, 20(17):6414–6425, 2000.

[58] F Pagani, C Stuani, M Tzetis, E Kanavakis, A Efthymiadou, S Doudounakis, T Casals, and FE Baralle. New type of disease causing mutations: the example of the composite exonic regulatory elements of splicing in CFTR exon 12. *Human Molecular Genetics*, 12(10):1111–1120, 2003.

[59] C Dominguez and FH Allain. NMR structure of the three quasi RNA recognition motifs (qRRMs) of human hnRNP F and interaction studies with Bcl-x G-tract RNA: a novel mode of RNA recognition. *Nucleic Acids Research*, 34(13):3634–3645, 2006.

[60] Z Wang, ME Rolish, G Yeo, V Tung, M Mawson, and CB Burge. Systematic identification and analysis of exonic splicing silencers. *Cell*, 119(6):831–845, 2004.

[61] K Han, G Yeo, P An, CB Burge, and PJ Grabowski. A combinatorial code for splicing silencing: UAGG and GGGG motifs. *PLoS Biology*, 3(5):e158, 2005.

[62] Black DL. Mechanisms of alternative pre-messenger RNA splicing. *Annual Review of Biochemistry*, 72:291–336, 2003.

[63] BJ Blencowe. Exonic splicing enhancers: mechanism of action, diversity and role in human genetic diseases. *Trends in Biochemical Sciences*, 25(3):106–110, 2000.

[64] H Tian and R Kole. Selection of novel exon recognition elements from a pool of random sequences. *Molecular and Cellular Biology*, 15(11):6291–6298, 1995.

[65] TD Schaal and T Maniatis. Selection and characterization of pre-mRNA splicing enhancers: identification of novel SR protein-specific enhancer sequences. *Molecular and Cellular Biology*, 19(3):1705–1719, 1999.

[66] LR Coulter, MA Landree, and TA Cooper. Identification of a new class of exonic splicing enhancers by in vivo selection. *Molecular and Cellular Biology*, 17(4):2143–2150, 1997.

[67] L Cartegni, J Wang, Z Zhu, MQ Zhang, and AR Krainer. ESEfinder: a web resource to identify exonic splicing enhancers. *Nucleic Acids Research*, 31(13):3568–3571, 2003.

[68] Z Zheng. Regulation of alternative RNA splicing by exon definition and exon sequences in viral and mammalian gene expression. *Journal of Biomedical Science*, 11(3):278–294, 2004.

[69] EM McCarthy and JA Phillips. Characterization of an intron splice enhancer that regulates alternative splicing of human GH pre-mRNA. *Human Molecular Genetics*, 7(9):1491–1496, 1998.

[70] JP Venables. Downstream intronic splicing enhancers. *FEBS Letters*, 581(22):4127–4131, 2007.

[71] WG Fairbrother and LA Chasin. Human genomic sequences that inhibit splicing. *Molecular and Cellular Biology*, 20(18):6816–6825, 2000.

[72] G Dreyfuss, MJ Matunis, S Piñol-Roma, and CG Burd. hnRNP proteins and the biogenesis of mRNA. *Annual Review of Biochemistry*, 62:289–321, 1993.

[73] AM Krecic and MS Swanson. hnRNP complexes: composition, structure, and function. *Current Opinion in Cell Biology*, 11(3):363–371, 1999.

[74] EJ Wagner and MA Garcia-Blanco. Polypyrimidine tract binding protein antagonizes exon definition. *Molecular and Cellular Biology*, 21(10):3281–3288, 2001.

[75] J Sauliere, A Sureau, A Expert-Bezancon, and J Marie. The polypyrimidine tract binding protein (PTB) represses splicing of exon 6B from the -tropomyosin pre-mRNA by directly interfering with the binding of the U2AF65 subunit. *Molecular and Cellular Biology*, 26(23):8755–8769, 2006.

[76] F Pagani, E Buratti, C Stuani, M Romano, E Zuccato, M Niksic, L Giglio, D Faraguna, and FE Baralle. Splicing factors induce cystic fibrosis transmembrane regulator exon 9 skipping through a nonevolutionary conserved intronic element. *Journal of Biological Chemistry*, 275(28):21041–21047, 2000.

[77] E Buratti and FE Baralle. Influence of RNA secondary structure on the pre-mRNA splicing process. *Molecular and Cellular Biology*, 24(24):10505–10514, 2004.

[78] M Hiller, Z Zhang, R Backofen, and S Stamm. Pre-mRNA secondary structures influence exon recognition. *PLoS Genetics*, 3(11):e204, 2007.

[79] AA Patel and JA Steitz. Splicing double: insights from the second spliceosome. *Nature Reviews Molecular Cell Biology*, 4(12):960–970, 2003.

[80] CL Will and R Lührmann. Spliceosomal UsnRNP biogenesis, structure and function. *Current Opinion in Cell Biology*, 13(3):290–301, 2001.

[81] C Kambach, S Walke, and K Nagai. Structure and assembly of the spliceosomal small nuclear ribonucleoprotein particles. *Current Opinion in Structural Biology*, 9(2):222–230, 1999.

[82] U Fisher, V Sumpter, M Sekine, T Satoh, and R Lührmann. Nucleo-cytoplasmic transport of U snRNPs: definition of a nuclear location signal in the Sm core domain that binds a transport receptor independently of the m3G cap. *The EMBO Journal*, 12(2):573–583, 1993.

[83] H Stark, P Dube, R Lührmann, and B Kastner. Arrangement of RNA and proteins in the spliceosomal U1 small nuclear ribonucleoprotein particle. *Nature*, 409(6819):539–542, 2001.

[84] CJ David and JL Manley. The search for alternative splicing regulators: new approaches offer a path to a splicing code. *Genes & Development*, 22(3):279–285, 2008.

[85] AM Zahler, WS Lane, JA Stolk, and MB Roth. SR proteins: a conserved family of pre-mRNA splicing factors. *Genes & Development*, 6(5):837–847, 1992.

[86] E Birney, S Kumar, and AR Krainer. Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors. *Nucleic Acids Research*, 21(25):5803–5816, 1993.

[87] JL Manley and R Tacke. SR proteins and splicing control. *Genes & Development*, 10(13):1569–1579, 1996.

[88] HX Liu, M Zhang, and AR Krainer. Identification of functional exonic splicing enhancer motifs recognized by individual SR proteins. *Genes & Development*, 12(13):1998–2012, 1998.

[89] D Longman, IL Johnstone, and JF Cáceres. Functional characterization of SR and SR-related genes in *Caenorhabditis elegans*. *The EMBO Journal*, 19(7):1625–1637, 2000.

[90] HZ Ring and JT Lis. The SR protein B52/SRp55 is essential for Drosophila development. *Molecular and Cellular Biology*, 14(11):7499–7506, 1994.

[91] JF Cáceres, GR Screaton, and AR Krainer. A specific subset of SR proteins shuttles continuously between the nucleus and the cytoplasm. *Genes & Development*, 12(1):55–66, 1998.

[92] R Tacke, Y Chen, and JL Manley. Sequence-specific RNA binding by an SR protein requires RS domain phosphorylation: creation of an SRp40-specific splicing enhancer. *PNAS*, 94(4):1148–1153, 1997.

[93] E Izaurralde and IW Mattaj. RNA export. *Cell*, 81(2):153–159, 1995.

[94] S Piñol-Roma and G Dreyfuss. Shuttling of pre-mRNA binding proteins between nucleus and cytoplasm. *Nature*, 355(6362):730–732, 1992.

[95] H Kamma, DS Portman, and G Dreyfuss. Cell type-specific expression of hnRNP proteins. *Experimental Cell Research*, 221(1):187–196, 1995.

[96] M Caputi and AM Zahler. SR proteins and hnRNP H regulate the splicing of the HIV-1 tev-specific exon 6D. *The EMBO Journal*, 21(4):845–855, 2002.

[97] CW Smith and J Valcárcel. Alternative pre-mRNA splicing: the logic of combinatorial control. *Trends in Biochemical Sciences*, 25(8):381–388, 2000.

[98] Q Pan, O Shai, LJ Lee, BJ Frey, and BJ Blencowe. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nature Genetics*, 40(12):1413–1415, 2008.

[99] ET Wang, R Sandberg, S Luo, I Khrebtukova, L Zhang, C Mayr, SF Kingsmore, GP Schroth, and CB Burge. Alternative isoform regulation in human tissue transcriptomes. *Nature*, 456(7221):470–476, 2008.

[100] S Stamm, S Ben-Ari, I Rafalska, Y Tang, Z Zhang, D Toiber, TA Thanaraj, and H Soreq. Function of alternative splicing. *Gene*, 344:1–20, 2005.

[101] M Cesana, D Cacchiarelli, I Legnini, T Santini, O Sthandier, M Chinappi, A Tramontano, and I Bozzoni. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell*, 147(2):358–369, 2011.

[102] MJ Pajares, T Ezponda, R Catena, A Calvo, R Pio, and LM Montuenga. Alternative splicing: an emerging topic in molecular and clinical oncology. *The LANCET Oncology*, 8(4):349–357, 2007.

[103] S Kadener, JP Fededa, M Rosbash, and AR Kornblihtt. Regulation of alternative splicing by a transcriptional enhancer through RNA pol II elongation. *PNAS*, 99(12):8185–8190, 2002.

[104] H Tilgner, C Nikolaou, S Althammer, M Sammeth, M Beato, J Valcárcel, and R Guigó. Nucleosome positioning as a determinant of exon recognition. *Nature Structural & Molecular Biology*, 16(9):996–1001, 2009.

[105] E Batsché, M Yaniv, and C Muchardt. The human SWI/SNF subunit Brm is a regulator of alternative splicing. *Nature Structural & Molecular Biology*, 13(1):22–29, 2006.

[106] V Ambros. A uniform system for microRNA annotation. *RNA*, 9(3):277–279, 2003.

[107] A Kozomara and S Griffiths-Jones. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Research*, 42(D1):D68–D73, 2014.

[108] W Filipowicz, SN Bhattacharyya, and N Sonenberg. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nature Reviews Genetics*, 9(2):102–114, 2008.

[109] DH Kim, P Saetrom, O Snøve, and JJ Rossi. MicroRNA-directed transcriptional gene silencing in mammalian cells. *PNAS*, 105(42):16230–16235, 2008.

[110] DP Bartel. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, 116(2):281–297, 2004.

[111] RC Lee, RL Feinbaum, and V Ambros. The *C. elegans* heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell*, 75(5):843–854, 1993.

[112] B Wightman, I Ha, and G Ruvkun. Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. *Cell*, 75(5):855–862, 1993.

[113] BJ Reinhart, FJ Slack, M Basson, AE Pasquinelli, JC Bettinger, AE Rougvie, HR Horvitz, and G Ruvkun. The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature*, 403(6772):901–906, 2000.

[114] AE Pasquinelli, B J Reinhart, F Slack, MQ Martindale, MI Kuroda, B Maller, et al. Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature*, 408(6808):86–89, 2000.

[115] M Lagos-Quintana, R Rauhut, A Yalcin, J Meyer, W Lendeckel, and T Tuschl. Identification of tissue-specific microRNAs from mouse. *Current Biology*, 12(9):735–739, 2002.

[116] C Ibáñez-Ventoso, M Vora, and M Driscoll. Sequence relationships among *C. elegans*, *D. melanogaster* and Human microRNAs highlight the extensive conservation of microRNAs in biology. *PLoS ONE*, 3(7):e2818, 2008.

[117] YS Lee and A Dutta. MicroRNAs in Cancer. *Annual Review of Pathology: Mechanisms of Disease*, 4(1):199–227, 2009.

[118] X Cai. Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA*, 10(12):1957–1966, 2004.

[119] GM Borchert, W Lanier, and BL Davidson. RNA polymerase III transcribes human microRNAs. *Nature Structural & Molecular Biology*, 13(12):1097–1101, 2006.

[120] TA Nguyen, MH Jo, YG Choi, J Park, SC Kwon, S Hohng, VN Kim, and JS Woo. Functional Anatomy of the Human Microprocessor. *Cell*, 161(6):1374–1387, 2015.

[121] J Han, Y Lee, KH Yeom, JW Nam, I Heo, JK Rhee, SY Sohn, Y Cho, BT Zhang, and VN Kim. Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. *Cell*, 125(5):887–901, 2006.

[122] E Lund. Nuclear export of microRNA precursors. *Science*, 303(5654):95–98, 2004.

[123] MT Bohnsack. Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA*, 10(2):185–191, 2004.

[124] A Calado, N Treichel, EC Müller, A Otto, and U Kutay. Exportin-5-mediated nuclear export of eukaryotic elongation factor 1A and tRNA. *The EMBO Journal*, 21(22):6216–6224, 2002.

[125] G Hutvágner, J McLachlan, AE Pasquinelli, E Bálint, T Tuschl, and PD Zamore. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science*, 293(5531):834–838, 2001.

[126] PW Lau, KZ Guiley, N De, CS Potter, B Carragher, and IJ MacRae. The molecular architecture of human Dicer. *Nature Structural & Molecular Biology*, 19(4):436–440, 2012.

[127] TP Chendrimada, RI Gregory, E Kumaraswamy, J Norman, N Cooch, K Nishikura, and R Shiekhattar. TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature Cell Biology*, 436(7051):740–744, 2005.

[128] A Khvorova, A Reynolds, and SD Jayasena. Functional siRNAs and miRNAs exhibit strand bias. *Cell*, 115(2):209–216, 2003.

[129] DS Schwarz, G Hutvágner, T Du, Z Xu, N Aronin, and PD Zamore. Asymmetry in the assembly of the RNAi enzyme complex. *Cell*, 115(2):199–208, 2003.

[130] Y Lee, K Jeon, JT Lee, S Kim, and VN Kim. MicroRNA maturation: stepwise processing and subcellular localization. *The EMBO Journal*, 21(17):4663–4670, 2002.

[131] RI Gregory, K Yan, G Amuthan, T Chendrimada, B Doratotaj, N Cooch, and R Shiekhattar. The Microprocessor complex mediates the genesis of microRNAs. *Nature Cell Biology*, 432(7014):235–240, 2004.

[132] Y Tomari. A protein sensor for siRNA asymmetry. *Science*, 306(5700):1377–1380, 2004.

[133] B Czech and GJ Hannon. Small RNA sorting: matchmaking for Argonautes. *Nature Reviews Genetics*, 12(1):19–31, 2011.

[134] FV Rivas, NH Tolia, JJ Song, JP Aragon, J Liu, GJ Hannon, and L Joshua-Tor. Purified Argonaute2 and an siRNA form recombinant human RISC. *Nature Structural & Molecular Biology*, 12(4):340–349, 2005.

[135] JF Chen, EM Mandel, JM Thomson, Q Wu, TE Callis, SM Hammond, FL Conlon, and DZ Wang. The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. *Nature Genetics*, 38(2):228–233, 2005.

[136] PK Rao, RM Kumar, M Farkhondeh, S Baskerville, and HF Lodish. Myogenic factors that regulate expression of muscle-specific microRNAs. *PNAS*, 103(23):8721–8726, 2006.

[137] P Laneve, U Gioia, A Andriotto, F Moretti, I Bozzoni, and E Caffarelli. A minicircuitry involving REST and CREB controls miR-9-2 expression during human neuronal differentiation. *Nucleic Acids Research*, 38(20):6895–6905, 2010.

[138] GT Bommer, I Gerin, Y Feng, AJ Kaczorowski, R Kuick, RE Love, Y Zhai, et al. p53-mediated activation of miRNA34 candidate tumor-suppressor genes. *Current Biology*, 17(15):1298–1307, 2007.

[139] L He and GJ Hannon. MicroRNAs: small RNAs with a big role in gene regulation. *Nature Reviews Genetics*, 5(7):522–531, 2004.

[140] V Tarasov, P Jung, B Verdoodt, D Lodygin, A Epanchintsev, A Menssen, et al. Differential regulation of microRNAs by p53 revealed by massively parallel sequencing: miR-34a is a p53 target that induces apoptosis and G1-arrest. *Cell Cycle*, 6(13):1586–1593, 2007.

[141] M Dews, A Homayouni, D Yu, D Murphy, C Sevignani, E Wentzel, EE Furth, et al. Augmentation of tumor angiogenesis by a Myc-activated microRNA cluster. *Nature Genetics*, 38(9):1060–1065, 2006.

[142] N Tanaka, S Toyooka, J Soh, T Kubo, H Yamamoto, Y Maki, T Muraoka, et al. Frequent methylation and oncogenic role of microRNA-34b/c in small-cell lung cancer. *Lung Cancer*, 76(1):32–38, 2012.

[143] J Krol, I Loedige, and W Filipowicz. The widespread regulation of microRNA biogenesis, function and decay. *Nature Reviews Genetics*, 11(9):597–610, 2010.

[144] HI Suzuki, K Yamagata, K Sugimoto, T Iwamoto, S Kato, and K Miyazono. Modulation of microRNA processing by p53. *Nature*, 460(7254):529–533, 2009.

[145] I Heo, C Joo, J Cho, M Ha, J Han, and VN Kim. Lin28 mediates the terminal uridylation of let-7 precursor microRNA. *Molecular Cell*, 32(2):276–284, 2008.

[146] BN Davis, AC Hilyard, G Lagna, and A Hata. SMAD proteins control DROSHA-mediated microRNA maturation. *Nature*, 454(7200):56–61, 2008.

[147] H Wu, S Sun, K Tu, Y Gao, X Bin, AR Krainer, and J Zhu. A splicing-independent function of SF2/ASF in microRNA processing. *Molecular Cell*, 38(1):67–77, 2010.

[148] G Michlewski, S Guil, CA Semple, and JF Cáceres. Posttranscriptional regulation of miRNAs harboring conserved terminal loops. *Molecular Cell*, 32(3):383–393, 2008.

[149] M Trabucchi, P Briata, M Garcia-Mayoral, AD Haase, W Filipowicz, A Ramos, R Gherzi, and MG Rosenfeld. The RNA-binding protein KSRP promotes the biogenesis of a subset of microRNAs. *Nature*, 459(7249):1010–1014, 2009.

[150] BS Heale, LP Keegan, L McGurk, G Michlewski, J Brindle, CM Stanton, JF Cáceres, and MA O'Connell. Editing independent effects of ADARs on the miRNA/siRNA pathways. *The EMBO Journal*, 28(20):3145–3156, 2009.

[151] Y Kawahara, B Zinshteyn, P Sethupathy, H Iizasa, AG Hatzigeorgiou, and K Nishikura. Redirection of silencing targets by adenosine-to-inosine editing of miRNAs. *Science*, 315(5815):1137–1140, 2007.

[152] HW Hwang, EA Wentzel, and JT Mendell. A hexanucleotide element directs microRNA nuclear import. *Science*, 315(5808):97–100, 2007.

[153] J Han, JS Pedersen, SC Kwon, CD Belair, YK Kim, KH Yeom, WY Yang, D Haussler, R Blelloch, and VN Kim. Posttranscriptional crossregulation between Drosha and DGCR8. *Cell*, 136(1):75–84, 2009.

[154] MC Vella. The C. elegans microRNA let-7 binds to imperfect let-7 complementary sites from the lin-41 3'UTR. *Genes & Development*, 18(2):132–137, 2004.

[155] C Shin, JW Nam, KK Farh, HR Chiang, A Shkumatava, and DP Bartel. Expanding the microRNA targeting code: functional sites with centered pairing. *Molecular Cell*, 38(6):789–802, 2010.

[156] JR Lytle, TA Yario, and JA Steitz. Target mRNAs are repressed as efficiently by microRNA-binding sites in the 5'UTR as in the 3'UTR. *PNAS*, 104(23):9667–9672, 2007.

[157] A Grimson, KK Farh, WK Johnston, P Garrett-Engele, LP Lim, and DP Bartel. MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Molecular Cell*, 27(1):91–105, July 2007.

[158] DT Humphreys, BJ Westman, DI Martin, and T Preiss. MicroRNAs control translation initiation by inhibiting eukaryotic initiation factor 4E/cap and poly(A) tail function. *PNAS*, 102(47):16961–16966, 2005.

[159] M Kiriakidou, GS Tan, S Lamprinaki, M De Planell-Saguer, PT Nelson, and Z Mourelatos. An mRNA m7G cap binding-like motif within human Ago2 represses translation. *Cell*, 129(6):1141–1151, 2007.

[160] LN Kinch and NV Grishin. The human Ago2 MC region does not contain an eIF4E-like mRNA cap binding motif. *Biology Direct*, 4(1):2, 2009.

[161] G Mathonnet, MR Fabian, YV Svitkin, A Parsyan, L Huck, T Murata, S Biffo, et al. MicroRNA inhibition of translation initiation *in vitro* by targeting the cap-binding complex eIF4F. *Science*, 317(5845):1764–1767, 2007.

[162] R Thermann and MW Hentze. Drosophila miR2 induces pseudo-polysomes and inhibits translation initiation. *Nature*, 447(7146):875–878, 2007.

[163] K Seggerson, L Tang, and EG Moss. Two genetic circuits repress the *Caenorhabditis elegans* heterochronic gene lin-28 after translation initiation. *Developmental Biology*, 243(2):215–225, 2002.

[164] CP Petersen, ME Bordeleau, J Pelletier, and PA Sharp. Short RNAs repress translation after initiation in mammalian cells. *Molecular Cell*, 21(4):533–542, 2006.

[165] S Nottrott, MJ Simard, and JD Richter. Human let-7a miRNA blocks protein production on actively translating polyribosomes. *Nature Structural & Molecular Biology*, 13(12):1108–1114, 2006.

[166] I Behm-Ansmant. mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. *Genes & Development*, 20(14):1885–1898, 2006.

[167] N Liu, M Landreh, K Cao, M Abe, GJ Hendriks, JR Kennerdell, Y Zhu, LS Wang, and NM Bonini. The microRNA miR-34 modulates ageing and neurodegeneration in Drosophila. *Nature*, 482(7386):519–523, 2012.

[168] S Vasudevan, Y Tong, and JA Steitz. Switching from repression to activation: microRNAs can up-regulate translation. *Science*, 318(5858):1931–1934, 2007.

[169] M Morlando, M Ballarino, N Gromak, F Pagano, I Bozzoni, and NJ Proudfoot. Primary microRNA transcripts are processed co-transcriptionally. *Nature Structural & Molecular Biology*, 15(9):902–909, 2008.

[170] C Mattioli, G Pianigiani, and F Pagani. A competitive regulatory mechanism discriminates between juxtaposed splice sites and pri-miRNA structures. *Nucleic Acids Research*, 41(18):8680–8691, 2013.

[171] Z Melamed, A Levy, R Ashwal-Fluss, G Lev-Maor, K Mekahel, N Atias, S Gilad, et al. Alternative splicing regulates biogenesis of miRNAs located across exon-intron junctions. *Molecular Cell*, 50(6):869–881, 2013.

[172] S Macias, M Plass, A Stajuda, G Michlewski, E Eyras, and JF Cáceres. DGCR8 HITS-CLIP reveals novel functions for the microprocessor. *Nature Structural & Molecular Biology*, 19(8):760–766, 2012.

[173] N Smalheiser and V Torvik. Mammalian microRNAs derived from genomic repeats. *Trends in Genetics*, 21(6):322–326, 2005.

[174] YK Kim and VN Kim. Processing of intronic microRNAs. *The EMBO Journal*, 26(3):775–783, 2007.

[175] MJ Dye, N Gromak, and NJ Proudfoot. Exon tethering in transcription by RNA polymerase II. *Molecular Cell*, 21(6):849–859, 2006.

[176] MM Janas, M Khaled, S Schubert, JG Bernstein, D Golan, RA Veguilla, et al. Feed-forward microprocessing and splicing activities at a microRNA-containing intron. *PLoS Genetics*, 7(10):e1002330, 2011.

[177] JM Pawlicki and JA Steitz. Primary microRNA transcript retention at sites of transcription leads to enhanced microRNA production. *The Journal of Cell Biology*, 182(1):61–76, 2008.

[178] K Okamura, JW Hagen, H Duan, DM Tyler, and EC Lai. The mirtron pathway generates microRNA-class regulatory RNAs in Drosophila. *Cell*, 130(1):89–100, 2007.

[179] JG Ruby, CH Jan, and DP Bartel. Intronic microRNA precursors that bypass Drosha processing. *Nature*, 448(7149):83–86, 2007.

[180] E Ladewig, K Okamura, AS Flynt, JO Westholm, and EC Lai. Discovery of hundreds of mirtrons in mouse and human small RNA data. *Genome Research*, 22(9):1634–1645, 2012.

[181] AS Flynt, JC Greimann, WJ Chung, CD Lima, and EC Lai. MicroRNA biogenesis via splicing and exosome-mediated trimming in Drosophila. *Molecular Cell*, 38(6):900–907, 2010.

[182] MA Havens, AA Reich, DM Duelli, and ML Hastings. Biogenesis of mammalian microRNAs by a non-canonical processing pathway. *Nucleic Acids Research*, 40(10):4626–4640, 2012.

[183] R Triboulet, RJ Chang, HM adn Lapierre, and RI Gregory. Post-transcriptional control of DGCR8 expression by the Microprocessor. *RNA*, 15(6):1005–1011, 2009.

[184] GM Sundaram, FE Common, JE adn Gopal, S Srikanta, K Lakshman, DP Lunny, et al. 'See-saw' expression of microRNA-198 and FSTL1 from a single transcript in wound healing. *Nature*, 495(7439):103–106, 2013.

[185] I Legnini, M Morlando, A Mangiavacchi, A Fatica, and I Bozzoni. A feedforward regulatory loop between HuR and the long noncoding RNA linc-MD1 controls early phases of myogenesis. *Molecular Cell*, 53(3):506–514, 2014.

[186] AY Qin, XW Zhang, L Liu, JP Yu, H Li, SZ Wang, XB Ren, and S Cao. MiR-205 in cancer: an angel or a devil? *European Journal of Cell Biology*, 92(2):54–60, 2013.

[187] S Muratsu-Ikeda, M Nangaku, Y Ikeda, T Tanaka, T Wada, and R Inagi. Downregulation of miR-205 modulates cell susceptibility to oxidative and endoplasmic reticulum stresses in renal tubular cells. *PLoS ONE*, 7(7):e41462, 2012.

[188] WS Chen, CM Leung, HW Pan, LY Hu, SC Li, MR Ho, and KW Tsai. Silencing of miR-1-1 and miR-133a-2 cluster expression by DNA hypermethylation in colorectal cancer. *Oncology Reports*, 28(3):1069–1076, 2012.

[189] T Yamasaki, H Yoshino, H Enokida, H Hidaka, T Chiyomaru, N Nohata, et al. Novel molecular targets regulated by tumor suppressors microRNA-1 and microRNA-133a in bladder cancer. *International Journal of Oncology*, 40(6):1821–1830, 2012.

[190] N Liu, S Bezprozvannaya, JM Shelton, MI Frisard, MW Hulver, RP McMillan, et al. Mice lacking microRNA 133a develop dynamin 2–dependent centronuclear myopathy. *The Journal of Clinical Investigation*, 121(8):3258–3268, 2011.

[191] N Liu, S Bezprozvannaya, AH Williams, X Qi, JA Richardson, R Bassel-Duby, and EN Olson. microRNA-133a regulates cardiomyocyte proliferation and suppresses smooth muscle gene expression in the heart. *Genes & Development*, 22(23):3242–3254, 2008.

[192] L He, X He, LP Lim, E de Stanchina, Z Xuan, Y Liang, W Xue, L Zender, J Magnus, et al. A microRNA component of the p53 tumour suppressor network. *Nature*, 447(7148):1130–1134, 2007.

[193] A Lujambio, GA Calin, A Villanueva, S Ropero, M Sánchez-Céspedes, D Blanco, et al. A microRNA DNA methylation signature for human cancer metastasis. *PNAS*, 105(36):13556–13561, 2008.

[194] DC Corney, A Flesken-Nikitin, AK Godwin, W Wang, and AY Nikitin. MicroRNA-34b and microRNA-34c are targets of p53 and cooperate in control of cell proliferation and adhesion-independent growth. *Cancer Research*, 67(18):8433–8438, 2007.

[195] Q Ji, X Hao, M Zhang, W Tang, M Yang, L Li, D Xiang, JT DeSano, GT Bommer, et al. MicroRNA miR-34 inhibits human pancreatic cancer tumor-initiating cells. *PLoS ONE*, 4(8):e6816, 2009.

[196] OW Rokhlin, VS Scheinker, AF Taghiyev, D Bumcrot, RA Glover, and MB Cohen. MicroRNA-34 mediates AR-dependent p53-induced apoptosis in prostate cancer. *Cancer Biology & Therapy*, 7(8):1288–1296, 2008.

[197] Y Xu, L Liu, J Liu, Y Zhang, J Zhu, J Chen, S Liu, Z Liu, H Shi, H Shen, and Z Hu. A potentially functional polymorphism in the promoter region of miR-34b/c is associated with an increased risk for primary hepatocellular carcinoma. *International Journal of Cancer*, 128(2):412–417, 2010.

[198] M Kato, T Paranjape, R Ullrich, S Nallur, E Gillespie, K Keane, A Esquela-Kerscher, JB Weidhaas, and FJ Slack. The mir-34 microRNA is required for the DNA damage response *in vivo* in *C. elegans* and in vitro in human breast cancer cells. *Oncogene*, 28(25):2419–2424, 2009.

[199] C Welch, Y Chen, and RL Stallings. MicroRNA-34a functions as a potential tumor suppressor by inducing apoptosis in neuroblastoma cells. *Oncogene*, 26(34):5017–5022, 2007.

[200] C He, J Xiong, X Xu, W Lu, L Liu, D Xiao, and D Wang. Functional elucidation of MiR-34 in osteosarcoma cells and primary tumor samples. *Biochemical and Biophysical Research Communications*, 388(1):35–40, 2009.

[201] M Pigazzi, E Manara, E Baron, and G Basso. miR-34b targets cyclic AMP-responsive element binding protein in acute myeloid leukemia. *Cancer Research*, 69(6):2471–2478, 2009.

[202] M Toyota, H Suzuki, Y Sasaki, R Maruyama, K Imai, Y Shinomura, and T Tokino. Epigenetic silencing of microRNA-34b/c and B-Cell translocation gene 4 is associated with CpG island methylation in colorectal cancer. *Cancer Research*, 68(11):4123–4132, 2008.

[203] DC Corney, CI Hwang, A Matoso, M Vogt, A Flesken-Nikitin, AK Godwin, AA Kamat, et al. Frequent downregulation of miR-34 family in human ovarian cancers. *Clinical Cancer Research*, 16(4):1119–1128, 2010.

[204] T Kubo, S Toyooka, K Tsukuda, M Sakaguchi, T Fukazawa, J Soh, H Asano, T Ueno, et al. Epigenetic silencing of MicroRNA-34b/c plays an important role in the pathogenesis of malignant pleural mesothelioma. *Clinical Cancer Research*, 17(15):4965–4974, 2011.

[205] KW Tsai, CW Wu, LY Hu, SC Li, YL Liao, CH Lai, HW Kao, WL Fang, KH Huang, WC Chan, and WC Lin. Epigenetic regulation of miR-34b and miR-129 expression in gastric cancer. *International Journal of Cancer*, 129(11):2600–2610, 2011.

[206] K Kozaki, I Imoto, S Mogi, K Omura, and J Inazawa. Exploration of tumor-suppressive microRNAs silenced by DNA hypermethylation in oral cancer. *Cancer Research*, 68(7):2094–2105, 2008.

[207] KY Wong, RLH Yim, CC So, DY Jin, R Liang, and CS Chim. Epigenetic inactivation of the MIR34B/C in multiple myeloma. *Blood*, 118(22):5901–5904, 2011.

[208] J Mazar, D Khaitan, D DeBlasio, C Zhong, SS Govindarajan, S Kopanathi, S Zhang, A Ray, and RJ Perera. Epigenetic regulation of microRNA genes and the role of miR-34b in cell Invasion and motility in human melanoma. *PLoS ONE*, 6(9):e24922, 2011.

[209] F Bouhallier, N Allioli, F Lavial, F Chalmel, MH Perrard, P Durand, J Samarut, B Pain, and JP Rouault. Role of miR-34c microRNA in the late steps of spermatogenesis. *RNA*, 16(4):720–731, 2010.

[210] BC Bernardo, XM Gao, CE Winbanks, EJ Boey, YK Tham, H Kiriazis, P Gregorevic, et al. microTherapeutic inhibition of the miR-34 family attenuates pathological cardiac remodeling and improves heart function. *PNAS*, 109(43):17615–17620, 2012.

[211] J Wei, Y Shi, L Zheng, B Zhou, H Inose, J Wang, XE Guo, R Grosschedl, and G Karsenty. miR-34s inhibit osteoblast proliferation and differentiation in the mouse by targeting SATB2. *The Journal of Cell Biology*, 197(4):509–521, 2012.

[212] Y Bae, T Yang, HC Zeng, PM Campeau, Y Chen, T Bertin, BC Dawson, E Munivez, J Tao, and BH Lee. miRNA-34c regulates Notch signaling during bone development. *Human Molecular Genetics*, 21(13):2991–3000, 2012.

[213] S Haramati, I Navon, O Issler, G Ezra-Nevo, S Gil, R Zwang, E Hornstein, and A Chen. microRNA as repressors of stress-induced anxiety: the Case of amygdalar miR-34. *Journal of Neuroscience*, 31(40):14191–14203, 2011.

[214] MM Aranha, DM Santos, S Solá, CJ Steer, and CM Rodrigues. miR-34a regulates mouse neural stem cell differentiation. *PLoS ONE*, 6(8):e21396, 2011.

[215] A Zovoilis, HY Agbemenyah, RC Agis-Balboa, RM Stilling, D Edbauer, P Rao, et al. MicroRNA-34c is a novel target to treat dementias. *The EMBO Journal*, 30(20):4299–4308, 2011.

[216] E Minones-Moyano, S Porta, G Escaramis, R Rabionet, S Iraola, B Kagerbauer, et al. MicroRNA profiling of Parkinson's disease brains identifies early downregulation of miR-34b/c which modulate mitochondrial function. *Human Molecular Genetics*, 20(15):3067–3078, 2011.

[217] PM Gaughwin, M Ciesla, N Lahiri, SJ Tabrizi, P Brundin, and M Bjorkqvist. Hsa-miR-34b is a plasma-stable microRNA that is elevated in pre-manifest Huntington's disease. *Human Molecular Genetics*, 20(11):2225–2237, 2011.

[218] E Proksch, JM Brandner, and JM Jensen. The skin: an indispensable barrier. *Experimental Dermatology*, 17(12):1063–1072, 2008.

[219] E Houben, K De Paepe, and V Rogiers. A keratinocyte's course of life. *Skin Pharmacology and Physiology*, 20(3):122–132, 2007.

[220] E Fuchs. Skin stem cells: rising to the surface. *The Journal of Cell Biology*, 180(2):273–284, 2008.

[221] Z Nemes and PM Steinert. Bricks and mortar of the epidermal barrier. *Experimental & Molecular Medicine*, 31(1):5–19, 1999.

[222] FM Watt. Involucrin and other markers of keratinocyte terminal differentiation. *Journal of Investigative Dermatology*, 81(1 Suppl):100–103, 1983.

[223] VS LeBleu, B Macdonald, and R Kalluri. Structure and function of basement membranes. *Experimental Biology and Medicine*, 232(9):1121–1129, 2007.

[224] CL Simpson, DM Patel, , and KJ Green. Deconstructing the skin: cytoarchitectural determinants of epidermal morphogenesis. *Nature Reviews Molecular Cell Biology*, 12(9):565–580, 2011.

[225] D Sawamura, H Nakano, and Y Matsuzaki. Overview of epidermolysis bullosa. *The Journal of Dermatology*, 37(3):214–219, 2010.

[226] CT Chung, SL Niemela, and RH Miller. One-step preparation of competent Escherichia coli: transformation and storage of bacterial cells in the same solution. *PNAS*, 86(7):2172–2175, 1989.

[227] TA Cooper. Use of minigene systems to dissect alternative splicing elements. *Methods*, 37(4):331–340, 2005.

[228] A Deirdre, J Scadden, and CW Smith. Interactions between the terminal bases of mammalian introns are retained in inosine-containing pre-mRNAs. *The EMBO Journal*, 14(13):3236–3246, 1995.

[229] C Mattioli, G Pianigiani, D De Rocco, AM Bianco, E Cappelli, A Savoia, and F Pagani. Unusual splice site mutations disrupt FANCA exon 8 definition. *BBA Molecular Basis of Disease*, 1842(7):1052–1058, 2014.

[230] G Zambruno, PC Marchisio, A Marconi, C Vaschieri, A Melchiori, A Giannetti, and M De Luca. Transforming growth factor-beta 1 modulates beta 1 and beta 5 integrin receptors and induces the de novo expression of the alpha v beta 6 heterodimer in normal human keratinocytes: implications for wound healing. *The Journal of Cell Biology*, 129(3):853–865, 1995.

[231] B Langmead and S Salzberg. Fast gapped-read alignment with Bowtie 2. *Nature Methods*, 9(4):357–359, 2012.

[232] MI Love, W Huber, and S Anders. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15(12):550, 2014.

[233] A Kauffmann, R Gentleman, and W Huber. arrayQualityMetrics–a bioconductor package for quality assessment of microarray data. *Bioinformatics*, 25(3):415–416, 2009.

[234] JJ Goeman and P Bühlmann. Analyzing gene expression data in terms of gene sets: methodological issues. *Bioinformatics*, 23(8):980–987, 2007.

[235] D Wu, E Lim, F Vaillant, ML Asselin-Labat, JE Visvader, and GK Smyth. ROAST: rotation gene set tests for complex microarray experiments. *Bioinformatics*, 26(17):2176–2182, 2010.

[236] J Královicová, S Houngninou-Molango, A Krämer, and I Vorechovsky. Branch site haplotypes that control alternative splicing. *Human Molecular Genetics*, 13(24):3189–3202, 2004.

[237] LM Scott and VI Rebel. Acquired mutations that affect pre-mRNA splicing in hematologic malignancies and solid tumors. *JNCI*, 105(20):1540–1549, 2013.

[238] F Minner and Y Poumay. Candidate housekeeping genes require evaluation before their selection for studies of human epidermal keratinocytes. *Journal of Investigative Dermatology*, 129:770–773, 2009.

[239] A Dhir, S Dhir, NJ Proudfoot, and CL Jopling. Microprocessor mediates transcriptional termination of long noncoding RNA transcripts hosting microRNAs. *Nature Structural & Molecular Biology*, 22(4):319–327, 2015.

[240] N Kataoka, M Fujita, and M Ohno. Functional Association of the Microprocessor Complex with the Spliceosome. *Molecular and Cellular Biology*, 29(12):3243–3254, 2009.

[241] R Scholl, J Marquis, K Meyer, and D Schümperli. Spinal muscular atrophy: position and functional importance of the branch site preceding SMN exon 7. *RNA Biology*, 4(1):34–37, 2007.

[242] PJ Shepard and KJ Hertel. Conserved RNA secondary structures promote alternative splicing. *RNA*, 14(8):1463–1469, 2008.

[243] A Corrionero, B Miñana, and J Valcárcel. Reduced fidelity of branch point recognition and alternative splicing induced by the anti-tumor drug spliceostatin A. *Genes & Development*, 25(5):445–459, 2011.

[244] D Grimm, KL Streetz, CL Jopling, TA Storm, K Pandey, CR Davis, P Marion, F Salazar, and MA Kay. Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. *Genes & Development*, 441(7092):537–541, 2006.

[245] GA Calin, CD Dumitru, M Shimizu, R Bichi, S Zupo, E Noch, H Aldler, et al. Frequent deletions and down-regulation of micro- RNA genes *miR15* and *miR16* at 13q14 in chronic lymphocytic leukemia. *PNAS*, 99(24):15524–15529, 2002.

[246] GA Calin and CM Croce. MicroRNA signatures in human cancers. *PNAS*, 6(11):857–866, 2006.

[247] T Andl, EP Murchison, F Liu, Y Zhang, M Yunta-Gonzalez, JW Tobias, CD Andl, et al. The miRNA-processing enzyme dicer is essential for the morphogenesis and maintenance of hair follicles. *Curr Biol*, 16(10):1041–1049, 2006.

[248] R Yi, D O'Carroll, HA Pasolli, Z Zhang, FS Dietrich, A Tarakhovsky, and E Fuchs. Morphogenesis in skin is governed by discrete sets of differentially expressed microRNAs. *Nature Genetics*, 38(3):356–362, 2006.

[249] R Yi, HA Pasolli, M Landthaler, M Hafner, T Ojo, R Sheridan, C Sander, D O'Carroll, M Stoffel, T Tuschl, and E Fuchs. Dgcr8-dependent microRNA biogenesis is essential for skin development. *PNAS*, 106(2):498–502, 2009.

[250] R Yi, MN Poy, M Stoffel, and E Fuchs. A skin microRNA promotes differentiation by repressing 'stemness'. *Nature*, 452(7184):225–229, 2008.

[251] J Hildebrand, M Rutze, N Walz, S Gallinat, H Wenck, W Deppert, A Grundhoff, and A Knott. A comprehensive analysis of microRNA expression during human keratinocyte differentiation in vitro and in vivo. *Journal of Investigative Dermatology*, 131(1):20–29, 2011.

[252] JE Wilusz, SM Freier, and DL Spector. 3' end processing of a long nuclear-retained noncoding RNA yields a tRNA-like cytoplasmic RNA. *Cell*, 135(5):919–932, 2008.

[253] Z Dominski and WF Marzluff. Formation of the 3' end of histone mRNA: getting closer to the end. *Gene*, 396(2):373–390, 2007.

[254] G Ghazal, J Gagnon, PE Jacques, JR Landry, F Robert, and SA Elela. Yeast RNase III triggers polyadenylation-independent transcription termination. *Molecular Cell*, 36(1):99–109, 2009.

[255] M Ballarino, F Pagano, E Girardi, M Morlando, D Cacchiarelli, M Marchioni, NJ Proudfoot, and I Bozzoni. Coupled RNA processing and transcription of intergenic primary microRNAs. *Molecular and Cellular Biology*, 29(20):5632–5638, 2009.

Appendix 3

Publication related to this project of thesis:

Mattioli C*, **Pianigiani G*** and Pagani F.
A competitive regulatory mechanism discriminates between juxtaposed splice sites and pri-miRNA structures.
Nucleic Acids Research 41(18):8680–8691, 2013
* Joint First Authors

Mattioli C, **Pianigiani G** and Pagani F.
Cross talk between spliceosome and microprocessor defines the fate of pre-mRNA.
Wiley Interdisciplinary Reviews: RNA 5(5):647–658, 2014

Other publications, not related to this project of thesis:

Mattioli C, **Pianigiani G**, De Rocco D, Bianco AM, Cappelli E, Savoia A, Pagani F.
Unusual splice site mutations disrupt FANCA exon 8 definition.
Biochimica et Biophysica Acta 1842(7):1052-1058, 2014

Scudieri P, Caci E, Venturini A, Sondo E, **Pianigiani G**, Marchetti C, Ravazzolo R, Pagani F, Galietta LJ.
Ion channel and lipid scramblase activity associated with expression of TMEM16F/ANO6 isoforms.
The Journal of Physiology 593(17):3829–3848, 2015

Nucleic Acids Research Advance Access published July 17, 2013

Nucleic Acids Research, 2013, 1–12
doi:10.1093/nar/gkt614

A competitive regulatory mechanism discriminates between juxtaposed splice sites and pri-miRNA structures

Chiara Mattioli, Giulia Pianigiani and Franco Pagani*

Human Molecular Genetics, International Centre for Genetic Engineering and Biotechnology, Padriciano 99, 34149, Trieste, Italy

Received April 10, 2013; Revised and Accepted June 21, 2013

ABSTRACT

We have explored the functional relationships between spliceosome and Microprocessor complex activities in a novel class of microRNAs (miRNAs), named Splice site Overlapping (SO) miRNAs, whose pri-miRNA hairpins overlap splice sites. We focused on the evolutionarily conserved SO miR-34b, and we identified two indispensable elements for recognition of its 3′ splice site: a branch point located in the hairpin and a downstream purine-rich exonic splicing enhancer. In minigene systems, splicing inhibition owing to exonic splicing enhancer deletion or AG 3′ss mutation increases miR-34b levels. Moreover, small interfering-mediated silencing of Drosha and/or DGCR8 improves splicing efficiency and abolishes miR-34b production. Thus, the processing of this 3′ SO miRNA is regulated in an antagonistic manner by the Microprocessor and the spliceosome owing to competition between these two machineries for the nascent transcript. We propose that this novel mechanism is commonly used to regulate the relative amount of SO miRNA and messenger RNA produced from primary transcripts.

INTRODUCTION

MicroRNAs (miRNAs) are 21–23-nt long non-coding RNAs that regulate gene expression by affecting translation and/or stability of messenger RNA (mRNAs) (1). Embedded in coding or non-coding genes, the hairpin secondary structure of primary (pri)-miRNAs is initially cropped in the nucleus by Drosha, an RNase III-like enzyme that is part of the Microprocessor Complex (MPC), along with its cofactor DGCR8 (2). The resulting precursor (pre)-miRNA is ~70 nt long and is exported to the cytoplasm where it is cleaved by Dicer to obtain the final mature form. On the nascent transcript, the MPC-dependent processing of the pri-miRNA hairpin is an important and early regulatory event involved in miRNA biogenesis. Indeed, several proteins interfere with the activity of the MPC, including RNA-binding proteins that either affect components of the MPC (3,4) or directly interact with the pri-miRNA hairpins (5–8).

The splicing reaction allows the maturation of a precursor (pre)-mRNA through the joining of the exonic sequences and the excision of the introns; to correctly identify exons, the splicing machinery recognizes the core *cis*-acting elements (9,10) that consist of the 5′ and 3′ splice sites (ss) and include the polypyrimidine tract and the branch point (BP) near the 3′ss. Recognition of the exon requires also splicing regulatory elements that are classified, depending on their location and effect on splicing, as exonic/intronic splicing enhancer and exonic/intronic splicing silencers (9,10). These elements are crucial for alternative splicing regulation, a mechanism present in the majority of human genes that enormously increase the transcript diversity through the selection of alternative splice sites (9). The exonic elements are composed of largely degenerated poorly conserved sequences and interact with splicing factors that may have a positive (serine/arginine-rich (SR) proteins) or a negative heterogeneous nuclear ribonucleoproteins (hnRNPs) effect on exon recognition (11).

Several polymerase II (PolII) precursor transcripts are processed co-transcriptionally by the spliceosome and the MPC into spliced mRNAs and miRNAs, respectively. In the case of intronic miRNA hairpins, which represent almost half of miRNAs (12), Drosha cleavage occurs before splicing and does not significantly affect the amount of mRNA (12,13). On the other hand, intronic pri-miRNA hairpins, both in coding or non-coding transcripts, are preferentially located at a distance from splice sites to avoid possible interference between the two processing machineries (12,13). Experiments using minigenes and *in vivo* analysis indicate that Drosha cleavage

*To whom correspondence should be addressed. Tel: +39 040 375 7342; Fax: +39 040 226 555; Email: pagani@icgeb.org

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

at intronic pri-miRNAs can both increase (12,14) and decrease (13,15) the splicing efficiency. pri-miRNA processing is more efficient if hairpins are retained at the sites of transcription (16), and, in some constructs, splicing disruption of both the 5′ss and 3′ss was found to affect miRNA biosynthesis (13). However, for the intronic miR-211, only mutations at the 5′ss were reported to reduce the biogenesis of the miRNA.

miR-34b, along with the related miR-34a and miR-34c, is involved in several physiological and pathological conditions. Originally identified as a tumour suppressor miRNA (17–19), miR-34b is involved in osteoblast proliferation (20,21), pathological cardiac remodelling (22) and Huntington and Parkinson diseases (23,24). miR-34b and miR-34c are part of the same non-coding transcriptional unit on chr11, possibly regulated by a p53-responsive promoter (25). The transcriptional unit is composed of two exons separated by a ~2 kb long intron. miR-34c is part of the last exon, whereas miR-34b is unexpectedly located on the boundary between intron 1 and exon 2.

In this study, we have identified a peculiar class of miRNAs, including miR-34b, whose hairpins overlap with splice sites and whose biogenesis is regulated by splicing. We have named these Splice site Overlapping (SO)-miRNAs. SO miR-34b overlaps with a non-canonical 3′ss, whose recognition depends on a strong BP and a purine-rich exonic splicing enhancer (ESE). Splicing inhibition by mutation of the 3′ss or the ESE, but not the 5′ss, increases miR-34b biosynthesis, whereas reduction of the Drosha/DGCR8 levels by RNAi knock-down increases splicing efficiency.

MATERIALS AND METHODS

Cell culture, transfections and reverse transcription-PCR analysis

HeLa cell culture and transfection, RNA extraction, reverse transcription (RT)-PCR and quantification of the percentage of splicing were performed as previously described (26). For the analysis of spliced isoforms, pBRA 34b minigenes were amplified with BRC90BsteII for (ctggtgaccaagtttgccagaaaacaccacatcactttaactaatc) and glo800 rev (gctcacagaagccaggaacttgtccagg); pcDNA3pY7 miR-34b constructs were amplified with pY7 ex2 dir (tacaaggcttgtcgaggaggacatc) and miR34b_2505XbaI rev (tatctagaccacgccgacgccgcgct). For co-transfection experiments, HeLa cells were transfected with 500 ng of the minigene construct together with 500 ng of an empty vector or vectors containing the proteins of interest.

Detection of spliced and unspliced miR-34b transcripts in mouse and human tissues

The human total RNA of 20 tissues was purchased from Amsbio, whereas the mouse one was extracted from tissues using TriReagent (Ambion) according to manufacturer's instructions.

The primers used for the RT-PCRs performed to detect the spliced and unspliced isoforms of human miR-34b transcripts were as follows: 34b_131 for (agtaggcaatgcatcttcatgac) and 34b_521 rev (ccttcgagagaagatgcctg) for the splicing form and 34b_233 for (cttttcaaggcatctgaccc) and 34b_435 rev (aatagtcttcattccattaaca) for the unspliced one. For the mouse spliced isoform, we used mmu_34b_4299 for and mmu_34b_6598 rev (CAATGA TAGCTTTGGATGGAAGC), and for the unspliced variant, we used mmu_34b_6397 for (agtagtagaaatagcctccatcc) and mmu_34b_6609 rev (GACAGTTTAT GCAATGATAGCT). For the control GAPDH gene, we used GAPDH for (gacagtcagccgcatcttct) and GAPDH rev (ttaaaagcagccctggtgac).

Quantitative RT-PCR

Quantitative (q)RT-PCR to detect the abundance of mouse and human mature miRNAs was performed using the TaqMan® Assay (Applied Biosystems). Both RT and PCR were performed according to manufacturer instructions.

Minigene design

To clone the hairpin structure of miR-34b and its flanking portions (18 bp upstream and 97 bp downstream), we amplified the region of interest by PCR from human genomic DNA. To clone in the pBRA plasmid (27), we used the following set of primers: miR-34b_2310 Xba for (tatctagacagccgcgggtgcccggtgc) and miR34b_2505 Pst rev (tactgcagccacgccgacgccgcgct). To clone the same region in the pcDNA3pY7 construct (28), we used the following oligonucleotides: miR-34b_2310 Xho for (tatactcgagccgcgggtgcccggtgct) and miR34b_2505Xba rev (tatctagaccacgccgacgccgcgct). To generate the mutated constructs, we amplified the sequence of interest with mutated primers built either for direct or overlapping PCR according to the position of the mutation.

Small interfering transfection and western blot

Small interfering (siRNA) transfections were performed in HeLa cells using Oligofectamine Reagent (Invitrogen) as previously described (26). The sense strand of RNAi oligos (Dharmacon) that were used to silence the target genes are as follows: cgaguaggcuucgugacuu (siDrosha) and caucggacaagagugugau (siDGCR8). The siLuc was the siCONTROL Non-Targeting siRNA #2 from Dharmacon. To check for protein silencing, western blot was performed with the following antibodies: antiDGS8 (N-19) from Santa Cruz Biotechnologies, antiDrosha (ab12286) from Abcam and antiTubulin antibody kindly provided by Dr Muro.

Northern blot for small RNAs

To perform northern blot for small RNA, we followed an already described procedure (29). Briefly, we loaded 30 µg of total RNA on a 13.5% (19:1) acrylamide/bisacrylamide, 5 M urea and a denaturing gel. After the run, we performed semi-dry transfer for 1 h at 2 mA/cm^2 and ultraviolet-cross-linking of the membrane with 1200 µJ. We used probes against miR-34b (atggcagtggagttagtgattg) and U6 snRNA (atatggaacgcttcacgaatt). We radiolabelled 10 pmol of the probe with γ-^{32}P ATP in the presence of T4 Kinase for 1 h at 37°C. Probe

hybridization was performed at 37°C for 2 h (U6snRNA) or O/N (miR-34b).

Bioinformatics analysis and manual annotation of SO pri-miRNAs

The algorithm was designed by Dennis Prickett at CBM, Trieste, Italy. We downloaded the coordinates of human miRNAs (www.mirbase.org) and the coordinates of the spliced expressed sequence tags (ESTs, http://genome.uscs.edu/). The algorithm was designed to compare the position of the 5′ and 3′ ss of the ESTs with the position of 5′ and 3′ ends of the miRNAs. When the distance was <100 bp, the miRNA of interest was printed. Then, through manual annotation, we selected the pri-miRNA hairpins overlapping with splice sites. Donor and acceptor splice site scores were calculated using the Neural Network method available at http://fruitfly.org/seq_tools/splice.html.

RESULTS

Identification of SO pri-miRNAs

To identify pri-miRNAs that overlap with splice sites, we compared the human annotated pri-miRNA hairpin coordinates (from www.mirbase.org) with the position of the splice sites derived from the analysis of the spliced expressed sequence tags (ESTs; extracted from http://genome.uscs.edu/). This analysis, followed by a careful manual annotation to exclude inappropriate intron–exon junctions, identified a group of pri-miRNAs whose predicted hairpins contain a possible splice site. We found 17 pri-miRNA hairpins overlapping with splice sites that accordingly were named SO pri-miRNAs. Eleven SO pri-miRNAs contain a 3′ss, six a 5′ss and eight are evolutionarily conserved among vertebrates (Table 1). Most are located within protein-coding genes, three belong to non-coding transcripts, and two derive from putative open reading frames (ORFs). The splice sites in the hairpins can be located at the base of the hairpin, near the junction between the single-stranded (ss)RNA and the double-stranded (ds)RNA, or in the stems, but none is in the terminal loop. The analysis of the splice site strength, evaluated with a neural network program, showed that most of the splice sites have a good score, but three 3′ss (miR-34b, miR-205 and miR-133a-2) were completely ignored (Table 1). These three non-canonical acceptor sites lack an obvious polypyrimidine tract. Interestingly, these miRNAs are implicated in several physio-pathological conditions, and downstream targets mRNAs have been identified (30,35,39).

The non-canonical 3′ss of SO pri-miR-34b is correctly used *in vivo* and in minigene systems

To investigate the dependence of splicing on these peculiar 3′ss that lack a polypyrimidine tract and are embedded in a pri-miRNA hairpin, we focused on pri-miR-34b. pri-miR-34b is part of a non-coding transcript composed of two exons and is located on the distal junction between intron 1 and exon 2 (Supplementary Figure S1). Interestingly, evolutionarily conserved sequences are present at the promoter, at the splice sites, at the pri-miRNAs and at the polyadenylation site (Supplementary Figure S1). The predicted miR-34b hairpin has a typical conserved pri-miRNA secondary structure with three stems (stems A, B and C) and a terminal loop (Figure 1). The AG dinucleotide of the 3′ss is located at the end of the pri-miRNA structure in stem A, four nucleotides above the ssRNA–dsRNA junction (Figure 1). To understand whether the non-canonical acceptor site is effectively used *in vivo*, we amplified the transcript in a panel of normal human and mouse RNAs with primers located on the exons. The results shown in Figure 2a and Supplementary Figure S2 indicate the presence of the spliced isoform in several human and mouse tissues, and direct sequencing of the bands revealed correct usage of the acceptor site. To understand the role of splicing on miR-34b biosynthesis, we evaluated spliced and unspliced transcripts along with miR-34b abundance in human tissues. In the commercial panel of human total RNAs (Amsbio), miR-34b is significantly expressed in cervix, ovary, trachea, testes and lung (Figure 2b). Interestingly, the relative amount of spliced and unspliced transcripts showed some tissue-specific differences. For example, the spliced and unspliced transcripts are evident in trachea and lung, but the unspliced form is the only product expressed in testes (Figure 2a). Quantitative RT-PCR analysis of the human tissues confirmed this tissue-specific distribution (Supplementary Figure S3). In addition, we observed high levels of the unspliced form also in mouse testes, suggesting that in this tissue production of miR-34b is evolutionarily linked to inefficient intron splicing (Supplementary Figure S2). All together, these data suggest that in some tissues, changes in splicing efficiency might contribute to miR-34b biosynthesis.

To evaluate in more detail how this non-canonical 3′ss is recognized and to identify the minimal sequences required for splicing, we inserted the human pri-miR-34b hairpin structure along with the flanking sequences (18 nt upstream and 97 nt downstream of the ssRNA–dsRNA junctions) in two minigene contexts (Figure 2c): the pcDNA3pY7 minigene (28) in which the 3′ss is before the terminal exon and the pBRA (27) where the 3′ss is part of an alternatively spliced exon. In both minigenes, transfection and RT-PCR analysis in HeLa cells showed correct selection of the acceptor site. In particular, pcDNA3pY7 miR-34b showed two transcripts: the shorter corresponds to splicing of the intron, whereas the longer corresponds to intron retention. In pBRA miR-34b minigene, the upper and the lower bands correspond to exon inclusion and skipping, respectively (Figure 2c). In pcDNA3pY7, miR-34b splicing efficiency was ~86%, whereas in pBRA, it was ~30%. Thus, these two minigenes are suitable systems to explore the splicing regulation of this SO pri-miRNA.

Splicing of SO miR-34b requires a BP sequence and a downstream purine-rich ESE

To identify the splicing regulatory elements involved in recognition of the non-canonical SO pri-miR-34b, we

Table 1. Splice site Overlapping miRNAs

miRNA	chr	miRNA conservation	Transcript	ss strength (neural network)	References
3′ SO miRNAs					
miR-205	1	Y	non-coding	0	(30,31,36)
miR-943	4		WHSC2	0.62	
miR-936	10		COL17A1	0.96	
miR-1287	10	Y	PYROXD2	0.85	
miR-34b	11	Y	non-coding	0	(17,18,23,32–34
miR-1178	12		CIT	0.90	
miR-636	17		SFRS2	0.64	
miR-4315–2	17	Y	PLEKHM1F	0.67	
miR-4321	19		AMH	0.58	
miR-133a-2	20	Y	C20orf166	0	(35,38)
miR-1292	20		NOP56	0.58	
5′ SO miRNAs					
miR-4260	1		LAMB3	0.54	
miR-761	1	Y	NRD1	0.95	
miR-555	1	Y	ASH1L	0.99	
miR-1204	8		non-coding PVT1	0.97	
miR-611	11		C11orf10	0.97	
miR-638	19	Y	DNM2	0.83	(37,53)

performed mutagenesis experiments in the context of the pcDNA3pY7 miR-34b minigene. To evaluate the importance of the AG dinucleotide of the 3′ss, we mutated it to CG, and this activated a 4 bp downstream cryptic acceptor site, located outside the hairpin (3′ss mut 1 in Figure 3a). Mutation of both natural and cryptic sites (3′ss mut 2 in Figure 3a) completely abolished splicing. The disruption of the 5′ss also resulted in complete splicing inhibition, as expected (Figure 3a). To identify the splicing regulatory sequences in the SO pri-miR-34b and explore the possible contribution of the hairpin secondary structure, we performed selective consecutive 5–14-nt long deletions (Figure 3b). Deletions were designed to eliminate the complementary sequences that form the three A, B and C stems of the pri-miRNA secondary structure. Splicing analysis showed that the mutants of the left part of the hairpin, either alone (ΔA1, ΔB1 and ΔC1) or combined (ΔA1+B1+C1), did not affect splicing (Figure 3b). The ΔB2 mutant had only a minor effect. On the contrary, deletion of the 14 bp long C2 (dC2, Figure 3b) or exchange between C2 and C1 (C1fC2, Figure 3c) significantly reduced the splicing efficiency, with only 15% of the intron excised. Fine mapping of the C2 element (Figure 3c) showed that the effect on splicing depends on the C2b portion that affects the 'CACTAAC' sequence. It perfectly matches the consensus for BP (YNYURAC, BP underlined) and is also located 18 bp upstream of the 3′ss, in accordance with the optimal conserved distance of BP in acceptor sites (40). Substitution of the three A with G (3A>G) or mutation of the critical T (41) (T>G and T>A) located upstream the A of the BP inhibited splicing (Figure 3d), strongly suggesting that this BP is critical for the processing of the pri-miR-34b transcript.

To identify additional splicing regulatory elements, we focused on downstream exonic sequences. In several cases, exonic regulatory sequences act on alternative splicing regulating the 3′ ss recognition (42,43). We prepared minigenes with progressive deletions of the exonic sequences. Minigenes 2482, 2469, 2457 and 2436 (Figure 4) contain progressive deletions of 23, 36, 48 and 69 bp of the exon, respectively. Deletion of the last 23 nt of the exon (mutant 2482) reduced the splicing efficiency to 38%, whereas further deletions completely abolished splicing (mutants 2469, 2457 and 2436; Figure 4). To map the regulatory element involved, we made internal 25 bp deletions (Figure 4), and the results showed that the sequence deleted in mutant Δ2457–2482 was sufficient to abolish splicing. This region contains a 9 bp purine-rich GAGA GAAGA sequence, and substitution of the four adenines into pyrimidines induced nearly complete intron retention (ESEmut, Figure 4).

All together, the experiments with the minigenes indicate that splicing of the non-canonical SO pri-miR-34b acceptor site requires an intronic BP located in the hairpin and a downstream purine-rich ESE. In addition, the absence of any effect on splicing of the combined ΔA1+B1+C1 deletion (Figure 3b) and the results obtained with the flip mutants (Figure 3c and d) indicate that the normal secondary structure of the SO pri-miR-34b is not required for the selection of the acceptor site.

Splicing of SO pri-miR-34b influences miRNA biosynthesis

To explore the effect of splicing on miRNA biosynthesis, we analysed the miR-34b derived from processing of the pcDNA3pY7 miR-34b minigenes. We evaluated the wild-type construct, 3′ss mut 1 and 3′ss mut 2 that directly affect the AG dinucleotide (Figure 3a), the 5′ss mutant (Figure 3a) and the ESEmut minigene (Figure 4). In comparison with the wild-type, the amount of mature miR-34b was significantly increased by the mutants that affect 3′ss splicing efficiency (Figure 5). In particular, 3′ss mut 1,

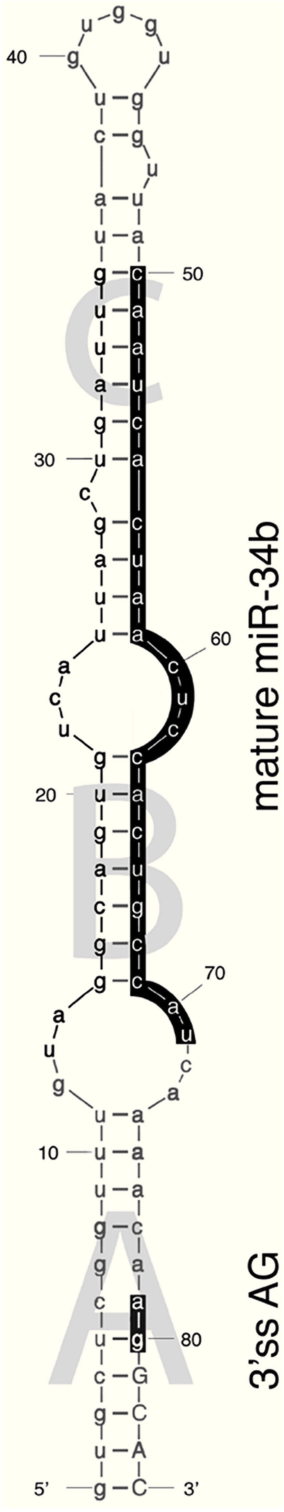


Figure 1. Secondary structure of SO miR-34b hairpin. The secondary structure of SO miR-34b hairpin has been calculated through the mfold web server (http://mfold.rna.albany.edu/?q=mfold/RNA-Folding-Form) with standard parameters. The three stems, named A, B and C, the mature form of miR-34b and the AG dinucleotide of the 3′ss are indicated. Lower and upper cases indicate the intronic and exonic sequences, respectively.

which reduced the splicing efficiency to ~50%, was associated with a ~1.6-fold increase in miR-34b (Figure 5), whereas the 3′ss mut 2 and the ESEmut, which completely abolished splicing, produced ~4-fold more miR-34b (Figure 5). On the other hand, splicing inhibition caused by mutation of the 5′ss only slightly reduced the amount of miR-34b (Figure 5). This decrease can be due to the facilitating effect of U1snRNP on Drosha processing, as recently suggested (14). In northern blot analysis, we did not observe any band corresponding to the pre-miRNA intermediate derived from transfection of normal or mutant minigenes. This may indicate that processing of the pre-miRNA by Dicer is efficient and is not a rate-limiting step for its maturation. In addition, we can exclude that the mutants affect miRNA abundance through Dicer-dependent pre-miRNA processing.

Silencing of the MPC proteins Drosha and DGCR8 improves SO miR-34b splicing efficiency

As splicing inhibition increases miR-34b biosynthesis, we decided to evaluate whether the MPC-dependent pri-miRNA processing affects the splicing efficiency. To this aim, we silenced Drosha and DGCR8 in HeLa cells that do not express miR-34b, followed by the evaluation of the splicing pattern of transfected minigenes (Figure 6). Western blotting showed almost complete silencing of Drosha and DGCR8 in HeLa cells (Figure 6a and b, lower panels) and a significant reduction of the control endogenous miR-26b (Figure 6c).

As the wild-type minigene is nearly completely spliced (pcDNA3pY7 miR-34b) and thus not suitable to appreciate further splicing improvement, we tested Drosha and DGCR8 silencing in pcDNA3pY7 miR-34b 2482 and pBRA miR-34b minigenes. In these two contexts, ~30–40% of transcripts are spliced in normal conditions (see Figures 4 and 2c, respectively). In pcDNA3pY7 miR-34b 2482, silencing of Drosha and DGCR8 either alone or in combination increased the percentage of splicing from 38 to ~70% (Figure 6a). Similarly, in the pBRA miR-34b, splicing efficiency increased, raising exon inclusion from 30 to ~50% (upper band of Figure 6b). On the control pBRAT6 minigene, which does not have the hairpin, we did not observe changes in the splicing pattern (Figure 6d). In addition, we found that Drosha and DGCR8 silencing reduced miR-34b biosynthesis derived from co-transfection of pcDNA3pY7 miR-34b (Figure 6c). As mirtrons (44,45) or 3′-tailed mirtrons (46) do not require the MPC, our results exclude the involvement of these non-canonical pathways in miR-34b biosynthesis.

On the other hand, overexpression of Drosha and DGCR8 in co-transfection experiments reduced the splicing efficiency both in pcDNA3pY7 miR-34b (Supplementary Figure S4a) and pBRA miR-34b (Supplementary Figure S4b) and had no effect on the control pBRA minigene (not shown). Thus, the MPC-dependent pri-miRNA processing interferes with spliceosome and affects the splicing efficiency of SO pri-miR-34b.

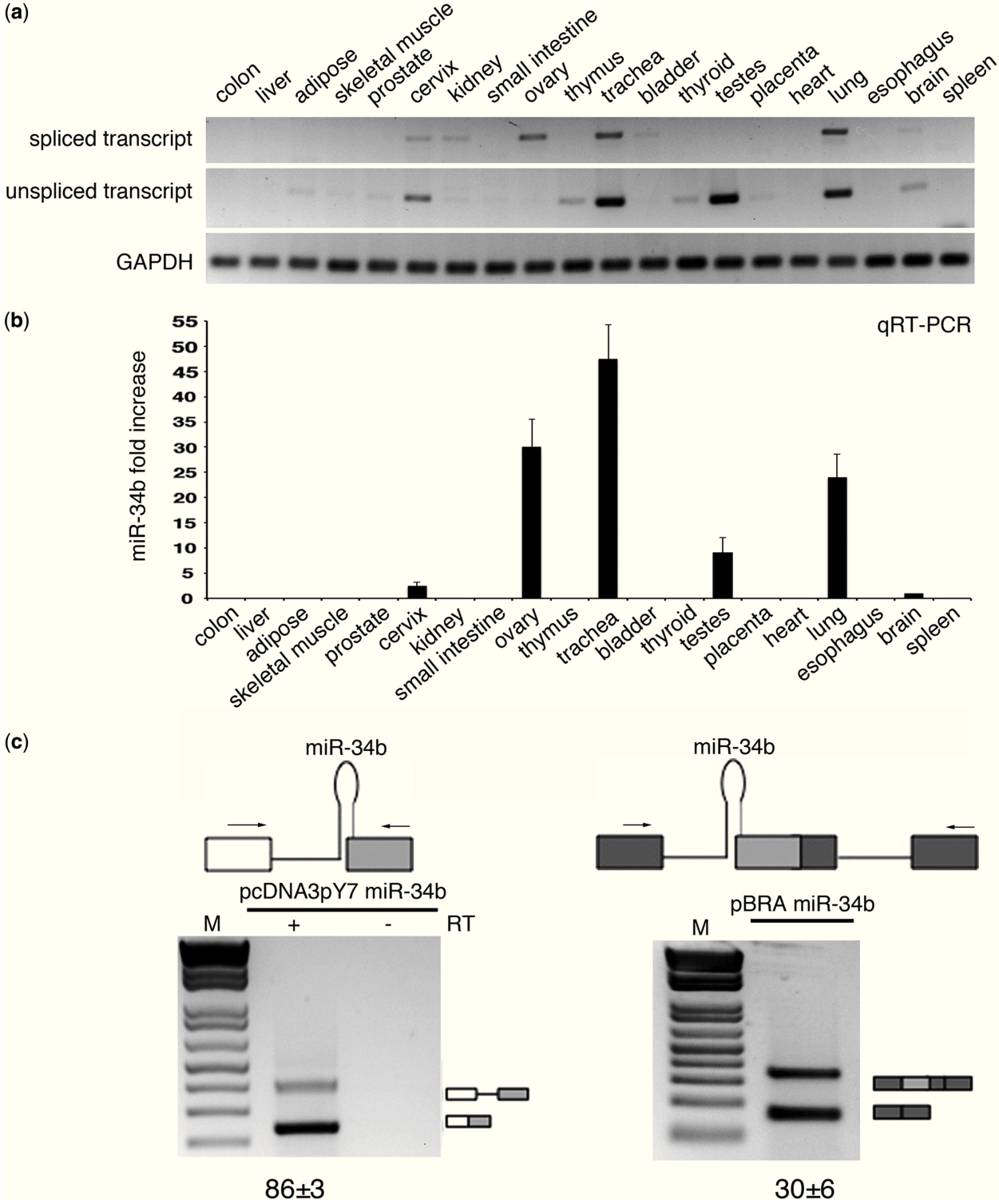


Figure 2. SO pri-miR-34b transcript is spliced *in vivo* and in minigene systems. (**a**) RT-PCR profile of miR-34b transcript in 20 human tissues. The upper panels show amplification of the spliced and unspliced transcripts, the lower one amplification of the control gene GAPDH. The identity of the bands was verified through direct sequencing, and quantitative analysis by qRT-PCR is provided in Supplementary Figure S3. (**b**) qRT-PCR analysis of the mature miR-34b in the different human tissues. Values are normalized for the GAPDH gene. miR-34b abundance in brain is set to 1. (**c**) Schematic representation of the minigene systems: on the left, pcDNA3pY7 miR-34b and on the right pBRA miR-34b. Thin lines represent introns and miR-34b hairpin. Boxes indicate exons, and, in particular, the light grey box is exon2 of miR-34b. Arrows indicate primers used for PCR analysis. On the bottom, the splicing profile of the two systems after transfection in HeLa cells. The identity of the bands is depicted on the right and was verified through direct sequencing. The numbers under the panel indicate the percentage of splicing (for pcDNA3pY7 miR-34b) or exon inclusion (pBRA miR-34b) ± SD.

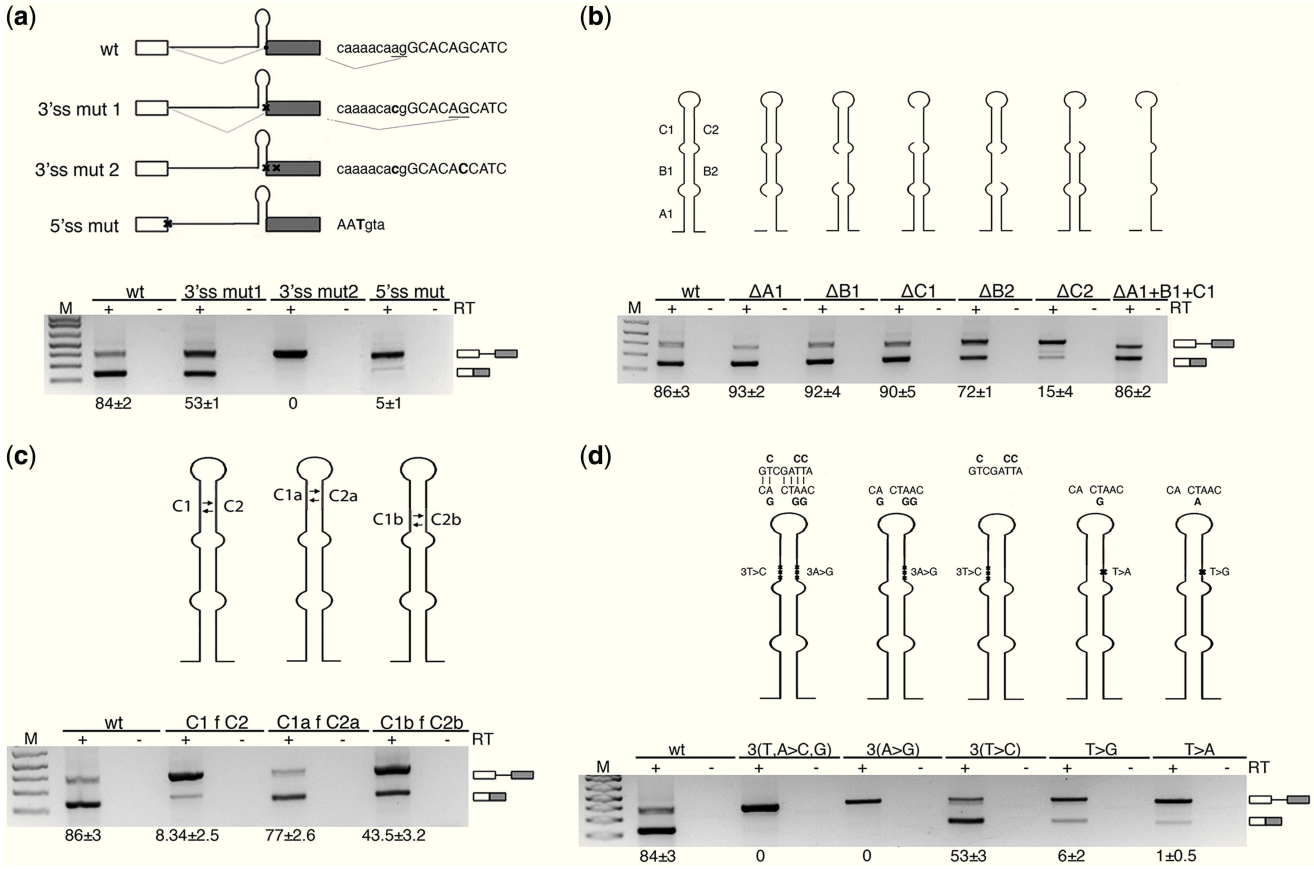


Figure 3. Splicing of SO miR-34b requires the AG dinucleotide and a BP sequence. (**a**) Splicing pattern of pcDNA3pY7 miR-34b mutants that affect the 3′ss and 5′ss after transfection in HeLa cells. Thin lines and boxes represent introns and exons, respectively. The position of miR-34b hairpin is indicated. Splice site mutants are crossed. Lower and upper cases identify intronic and exonic sequences, respectively, and mutated nucleotides are in bold. The results of splice site selection are indicated on the right of each minigene (underlined AG). (**b**) The deletion mutants of pri-miR-34b hairpin are schematically depicted above each lane. The lower panel shows the splicing pattern of the mutant plasmids after transfection in HeLa cells. (**c**) Splicing profile of mutants of the C stem of pri-miR-34b. The left and right arms of the C stem are twisted, as indicated in the scheme. (**d**) Scheme of the BP mutants. Mutants of the left and right arms of miR-34b hairpin are indicated above the pictures. The lower panel indicates the splicing pattern after transfection in HeLa cells. The numbers under each panel indicate the percentage of splicing expressed as mean ± SD of at least three independent experiments.

DISCUSSION

SO miRNAs are a novel class of miRNAs, characterized by the presence of overlapping pri-miRNA hairpins and splice sites on the nascent transcripts. In SO miR-34b, whose hairpin includes an acceptor site, we have shown that mutations that reduce the 3′ss splicing efficiency increase the mature miRNA form. Conversely, reduction of the Drosha/DGCR8 levels by RNAi knock-down improves splicing efficiency, whereas their overexpression reduces splicing. Thus, in the non-coding SO miR-34b and possibly in the other SO miRNAs, miRNA biosynthesis and production of mature mRNAs are regulated by the competition between the MPC and the spliceosome, owing to the overlap between pri-miRNA hairpins and splice sites. This competition represents a novel mechanism to regulate the level of miRNA biosynthesis through control of pre-mRNA processing.

The biosynthesis of miRNAs can be affected at the level of pri-miRNA processing through changes in the efficiency of MPC-dependent cropping of the nascent transcript (47). Some of the factors that affect cropping, such as Ars2 (3) or FUS (4), lack specificity, as they interact directly with component of the MPC to affect the biosynthesis of multiple miRNAs (4). In some cases, RNA-binding proteins that directly interact with the terminal loop sequence of the pri-miRNA hairpin regulate cropping of miRNAs. For example, LIN-28, SF2/ASF and hnRNPA1 bind to the terminal loop of pri-let-7, pri-miR-18a and pri-miR7, respectively, and affect their processing either acting on the MPC or through changes in the RNA secondary structure (5–7). In SO pri-miRNAs, where the splice sites overlap with the hairpins, changes in the splicing efficiency represent a novel system to directly regulate miRNA biosynthesis. In SO miR-34b, two elements are indispensable for splicing: a consensus BP located in the hairpin and a purine-rich ESE. The interaction of the ESE with RNA-binding protein(s) is the most likely mechanism that regulates splicing of SO miR-34b and consequently its biosynthesis. This ESE contains putative binding sites

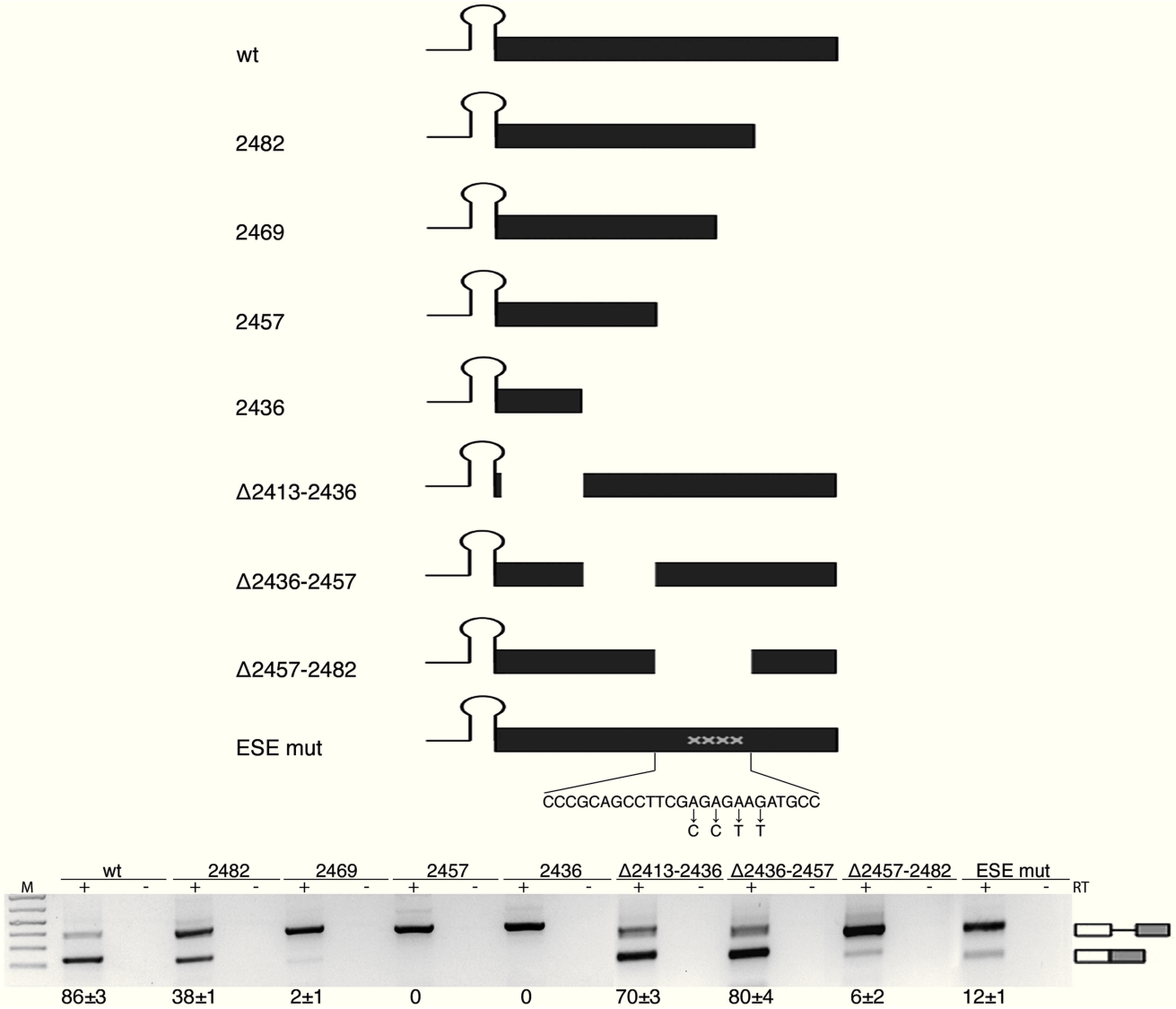


Figure 4. Splicing of SO miR-34b requires a purine-rich ESE. Schematic representation of the pcDNA3pY7 miR-34b mutants of the exon. The lines and black box correspond to the intronic miR-34b hairpin and exonic sequences, respectively. The mutated nucleotides of ESEmut are indicated. Lower panel shows the splicing pattern of pcDNA3pY7 miR-34b exonic mutants after transfection in HeLa cells. The identity of the bands is depicted on the right. Numbers below the panel indicate the percentage of splicing expressed as mean ± SD of at least three independent experiments.

for SF2/ASF (SF2) and Tra2beta (SFRS10), but neither their overexpression nor silencing affected SO miR-34b splicing (data not shown), suggesting that other splicing factors are involved. Additional experiments are required to identify the mechanism of ESE function and characterize the regulatory splicing factors that specifically affect miR-34b and the other SO miRNAs.

Most of the pre-mRNA processing events occur co-transcriptionally and are functionally coupled through the carboxyl-terminal domain of PolII (48,49), but the precise relationship between splicing and pri-miRNA cropping is not completely understood. *In vivo*, MPC-dependent cropping of pri-miRNA hairpins located inside introns has been reported to facilitate (12,14) or inhibit (13) splicing. On the other hand, spliceosome assembly was shown to promote cropping (14,16). However, splicing inhibition due to mutations that specifically affect the 5′ss results in a negative effect on production of miRNAs (12,14). The effect of the 5′ss mutation on the production of miR-34b that we observed here is not totally unexpected and consistent with a positive effect of the donor site and U1 snRNP on the MPC, as recently reported (12,14). In this case, a functional donor site, but not the 3′ss, has been shown to be critical for the biosynthesis of the intronic mir-211. Mutations of the 5′ss reduced the production of this miRNA and Drosha recruitment to the miRNA locus (14). Furthermore, intronic plant miRNAs also

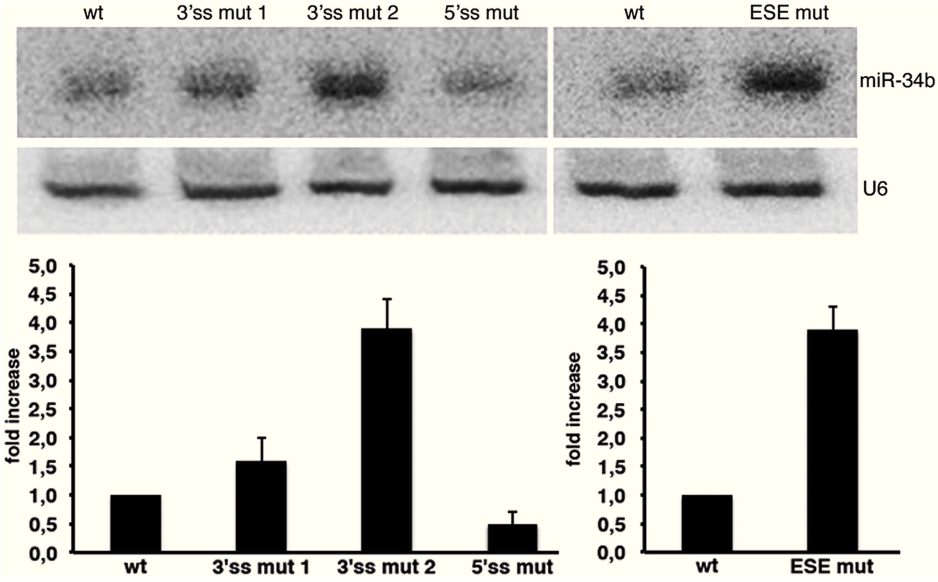


Figure 5. Splicing of SO pri-miR-34b affects miRNA biosynthesis. Northern blot analysis of miR-34b and U6 snRNA. The analysis was performed on the AG dinucleotide mutants, on the 5′ss mutant described in Figure 3 (3′ss mut 1, 3′ss mut 2 and 5′ss mut) and on the mutant of the ESE shown in Figure 4 (ESE mut). Histograms show the fold increase of mature miR-34b normalized to U6. The abundance of miR-34b in the wild-type construct is set to 1.

require a conserved 5′ss for their optimal expression (50). Consistent with the proposed enhancing effect of U1snRNP on the MPC activity (12,14), we also observed that the disruption of the 5′ss reduces the amount of mature miR-34b (Figure 5), in contrast to the mutations that affect splicing of the 3′ss. However, as U1snRNP binding to the first exon can facilitate PolII initiation (51), we cannot exclude a direct effect of the 5′ss mutation on transcription or stability of the nascent transcript. According to Morlando *et al.* (12), the positive effect on splicing of the MPC is mediated by the carboxyl-terminal domain of the PolII through a tethering mechanism that maintains exonic sequences in place during transcription, whereas cropped intronic fragments are degraded by exonucleases (52). However, *in vitro*, the two machineries can compete on preformed transcripts (15). In the case of SO miR-34b, and possibly for the other SO miRNAs, when the pri-miRNA hairpins overlap with a splice site, the MPC and the spliceosome act in a mutually exclusive manner to cleave the precursor transcript. The relative proportion of nascent transcripts that are either spliced or cropped *in vivo* is difficult to assess and might depend on individual SO miRNA features. The analysis of miR-34b in normal human tissues suggests that the spliceosome processes most of the nascent transcripts, as we did not observe a strict reciprocal relationship between the miRNA levels and the amount of spliced RNA. The regulated competition between splicing and the MPC probably affects only a fraction of transcripts in normal conditions, and this explains why in some tissues the levels of mature miR-34b do not correlate with the corresponding amount of spliced transcript (Figure 2a and b). In this case, tissue-specific changes in the splicing efficiency might regulate the competition between the spliceosome and the MPC. Interestingly, several pri-miRNA hairpins do not overlap but are in the proximity of splice sites (within 100 bp) raising the possibility that, in some cases, the position of the miRNA hairpin relative to the intron–exon architecture, along with the presence of intronic splicing regulatory elements, might determine the extent and type of interaction between the MPC and the spliceosome.

The antagonistic effect between the MPC and the spliceosome might have important consequences for SO miRNAs embedded in coding transcripts, which represent a significant proportion of SO miRNAs identified (Table 1). In these cases, the balance between cropping and splicing activities on the nascent transcript might determine how much miRNA is produced at the expense of the mRNA and therefore of protein synthesis. Thus, the coding or non-coding nature of the SO miRNA transcripts targeted by the MPC and the spliceosome might represent an additional level of regulation of gene expression.

Interestingly, several SO miRNAs are associated to physio-pathological conditions. Patients with Parkinson disease or Huntington disease show altered levels of miR-34b (23,24); miR-638 has been correlated with lupus nephritis severity (53); impairment of miR-34b, miR-205 and miR-133a-2 in several types of cancer is a well-established phenomenon (30,35,39). In these cases, the analysis of the splicing pattern of the corresponding transcripts would unveil the unexpected contribution of splicing abnormalities that are at the base of the altered synthesis of miRNAs.

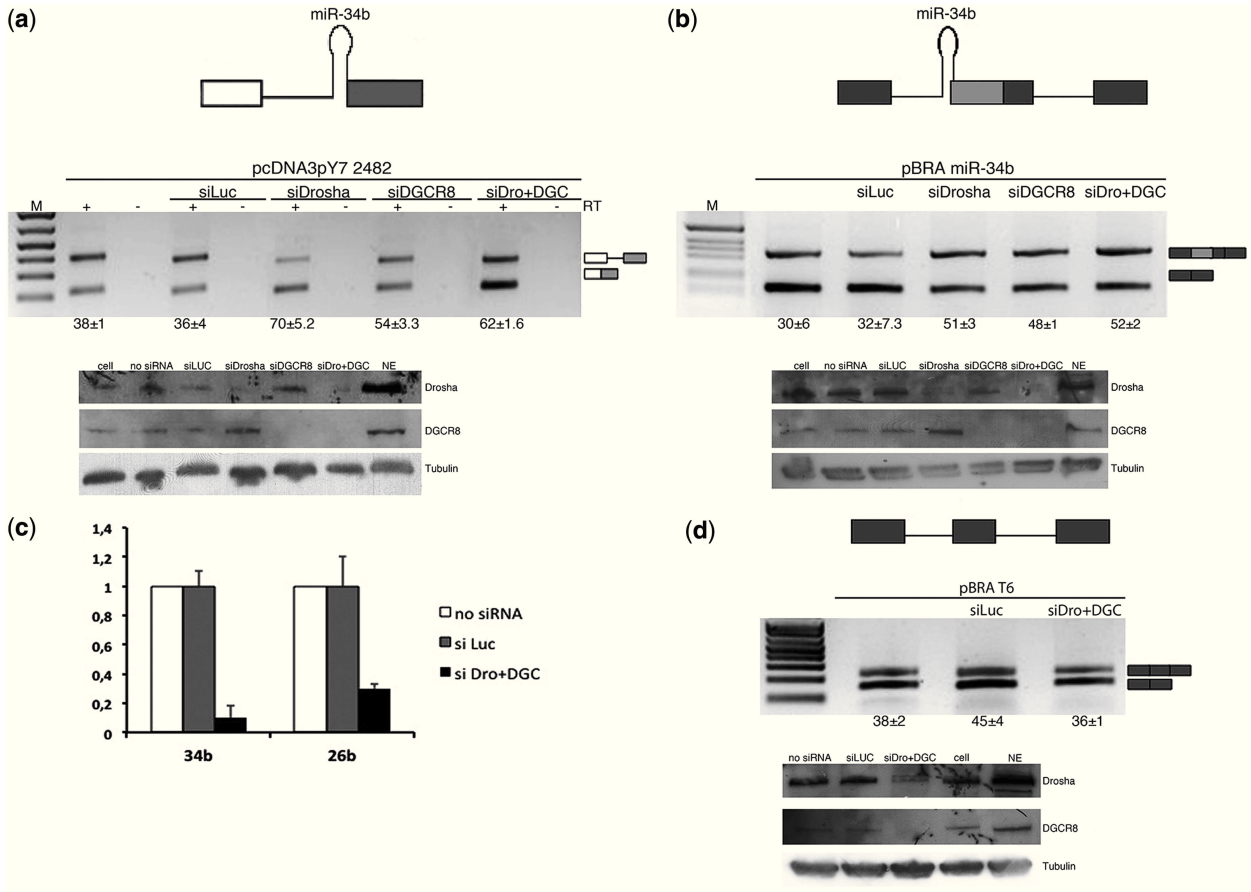


Figure 6. Silencing of Drosha and DGCR8 improves splicing efficiency. (**a** and **b**) Splicing pattern of pcDNA3pY7 2482 plasmid (see Figure 4) and pBRA miR-34b (see Figure 2c) after transfection of a control siRNA (siLuc) or silencing of Drosha and DGCR8, alone (siDrosha and siDGCR8) or together (siDro+DGC). Bands identity is indicated on the right. Numbers below the panel indicate the percentage of splicing (a) or exon inclusion (b) ± SD; western blot for Drosha, DGCR8 and tubulin is shown in the bottom panel (**c**) Quantitative RT-PCR of miR-34b derived from pcDNA3pY7 miR-34b and control endogenous miR-26b, normalized for the GAPDH gene. The abundance of miR-34b and miR-26b in non-treated cells (no siRNA) is set to 1. (**d**) RT-PCR analysis of pBRA T6 after transfection in HeLa cells and silencing of a control gene (siLuc) or silencing of Drosha and DGCR8 together (siDro+DGC). The percentage of exon inclusion ± SD is indicated below the panel. Western blot for Drosha, DGCR8 and tubulin is shown in the bottom panel.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors thank Dennis Prickett, CBM, Trieste for the help with the bioinformatics analysis. They also thank Shona Murphy for her support, discussion and critical reading of the manuscript.

FUNDING

Associazione Italiana per la Ricerca sul Cancro (AIRC) [10387]. Funding for open access charge: AIRC [10387].

Conflict of interest statement. None declared.

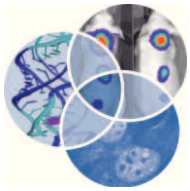
REFERENCES

1. Pillai,R.S. (2005) MicroRNA function: multiple mechanisms for a tiny RNA? *RNA*, **11**, 1753–1761.
2. Gregory,R.I., Yan,K.P., Amuthan,G., Chendrimada,T., Doratotaj,B., Cooch,N. and Shiekhattar,R. (2004) The Microprocessor complex mediates the genesis of microRNAs. *Nature*, **432**, 235–240.
3. Sabin,L.R., Zhou,R., Gruber,J.J., Lukinova,N., Bambina,S., Berman,A., Lau,C.K., Thompson,C.B. and Cherry,S. (2009) Ars2 regulates both miRNA- and siRNA- dependent silencing and suppresses RNA virus infection in. *Drosophila. Cell*, **138**, 340–351.
4. Morlando,M., Dini Modigliani,S., Torrelli,G., Rosa,A., Di Carlo,V., Caffarelli,E. and Bozzoni,I. (2012) FUS stimulates microRNA biogenesis by facilitating co-transcriptional Drosha recruitment. *EMBO J.*, **31**, 4502–4510.
5. Guil,S. and Caceres,J.F. (2007) The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a. *Nat. Struct. Mol. Biol.*, **14**, 591–596.
6. Newman,M.A., Thomson,J.M. and Hammond,S.M. (2008) Lin-28 interaction with the Let-7 precursor loop mediates regulated microRNA processing. *RNA*, **14**, 1539–1549.
7. Wu,H., Sun,S., Tu,K., Gao,Y., Xie,B., Krainer,A.R. and Zhu,J. (2010) A splicing-independent function of SF2/ASF in microRNA processing. *Mol. Cell*, **38**, 67–77.
8. Davis,B.N., Hilyard,A.C., Nguyen,P.H., Lagna,G. and Hata,A. (2010) Smad proteins bind a conserved RNA sequence to promote microRNA maturation by Drosha. *Mol. Cell*, **39**, 373–384.

9. Chen,M. and Manley,J.L. (2009) Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nat. Rev. Mol. Cell. Biol.*, **10**, 741–754.
10. Nilsen,T.W. and Graveley,B.R. (2010) Expansion of the eukaryotic proteome by alternative splicing. *Nature*, **463**, 457–463.
11. Caceres,J.F. and Kornblihtt,A.R. (2002) Alternative splicing: multiple control mechanisms and involvement in human disease. *Trends Genet.*, **18**, 186–193.
12. Morlando,M., Ballarino,M., Gromak,N., Pagano,F., Bozzoni,I. and Proudfoot,N.J. (2008) Primary microRNA transcripts are processed co-transcriptionally. *Nat. Struct. Mol. Biol.*, **15**, 902–909.
13. Kim,Y.K. and Kim,V.N. (2007) Processing of intronic microRNAs. *EMBO J*, **26**, 775–783.
14. Janas,M.M., Khaled,M., Schubert,S., Bernstein,J.G., Golan,D., Veguilla,R.A., Fisher,D.E., Shomron,N., Levy,C. and Novina,C.D. (2011) Feed-forward microprocessing and splicing activities at a microRNA-containing intron. *PLoS Genet.*, **7**, e1002330.
15. Kataoka,N., Fujita,M. and Ohno,M. (2009) Functional association of the Microprocessor complex with the spliceosome. *Mol. Cell. Biol.*, **29**, 3243–3254.
16. Pawlicki,J.M. and Steitz,J.A. (2008) Primary microRNA transcript retention at sites of transcription leads to enhanced microRNA production. *J. Cell Biol.*, **182**, 61–76.
17. Kubo,T., Toyooka,S., Tsukuda,K., Sakaguchi,M., Fukazawa,T., Soh,J., Asano,H., Ueno,T., Muraoka,T., Yamamoto,H. *et al.* (2011) Epigenetic silencing of microRNA-34b/c plays an important role in the pathogenesis of malignant pleural mesothelioma. *Clin. Cancer Res.*, **17**, 4965–4974.
18. Toyota,M., Suzuki,H., Sasaki,Y., Maruyama,R., Imai,K., Shinomura,Y. and Tokino,T. (2008) Epigenetic silencing of microRNA-34b/c and B-cell translocation gene 4 is associated with CpG island methylation in colorectal cancer. *Cancer Res.*, **68**, 4123–4132.
19. Corney,D.C., Hwang,C.I., Matoso,A., Vogt,M., Flesken-Nikitin,A., Godwin,A.K., Kamat,A.A., Sood,A.K., Ellenson,L.H., Hermeking,H. *et al.* (2010) Frequent downregulation of miR-34 family in human ovarian cancers. *Clin. Cancer Res.*, **16**, 1119–1128.
20. Bae,Y., Yang,T., Zeng,H.C., Campeau,P.M., Chen,Y., Bertin,T., Dawson,B.C., Munivez,E., Tao,J. and Lee,B.H. (2012) miRNA-34c regulates Notch signaling during bone development. *Hum. Mol. Genet.*, **21**, 2991–3000.
21. Wei,J., Shi,Y., Zheng,L., Zhou,B., Inose,H., Wang,J., Guo,X.E., Grosschedl,R. and Karsenty,G. (2012) miR-34s inhibit osteoblast proliferation and differentiation in the mouse by targeting SATB2. *J. Cell. Biol.*, **197**, 509–521.
22. Bernardo,B.C., Gao,X.M., Winbanks,C.E., Boey,E.J., Tham,Y.K., Kiriazis,H., Gregorevic,P., Obad,S., Kauppinen,S., Du,X.J. *et al.* (2012) Therapeutic inhibition of the miR-34 family attenuates pathological cardiac remodeling and improves heart function. *Proc. Natl Acad. Sci. USA*, **109**, 17615–17620.
23. Gaughwin,P.M., Ciesla,M., Lahiri,N., Tabrizi,S.J., Brundin,P. and Bjorkqvist,M. (2011) Hsa-miR-34b is a plasma-stable microRNA that is elevated in pre-manifest Huntington's disease. *Hum. Mol. Genet.*, **20**, 2225–2237.
24. Minones-Moyano,E., Porta,S., Escaramis,G., Rabionet,R., Iraola,S., Kagerbauer,B., Espinosa-Parrilla,Y., Ferrer,I., Estivill,X. and Marti,E. (2011) MicroRNA profiling of Parkinson's disease brains identifies early downregulation of miR-34b/c which modulate mitochondrial function. *Hum. Mol. Genet.*, **20**, 3067–3078.
25. Cha,Y.H., Kim,N.H., Park,C., Lee,I., Kim,H.S. and Yook,J.I. (2012) MiRNA-34 intrinsically links p53 tumor suppressor and Wnt signaling. *Cell Cycle*, **11**, 1273–1281.
26. Pastor,T. and Pagani,F. (2011) Interaction of hnRNPA1/A2 and DAZAP1 with an Alu-derived intronic splicing enhancer regulates ATM aberrant splicing. *PLoS One*, **6**, e23349.
27. Goina,E., Skoko,N. and Pagani,F. (2008) Binding of DAZAP1 and hnRNPA1/A2 to an exonic splicing silencer in a natural BRCA1 exon 18 mutant. *Mol. Cell Biol.*, **28**, 3850–3860.
28. Deirdre,A., Scadden,J. and Smith,C.W. (1995) Interactions between the terminal bases of mammalian introns are retained in inosine-containing pre-mRNAs. *EMBO J.*, **14**, 3236–3246.
29. Morgan,M., Iaconcig,A. and Muro,A.F. (2010) CPEB2, CPEB3 and CPEB4 are coordinately regulated by miRNAs recognizing conserved binding sites in paralog positions of their 3′-UTRs. *Nucleic Acids Res.*, **38**, 7698–7710.
30. Qin,A.Y., Zhang,X.W., Liu,L., Yu,J.P., Li,H., Wang,S.Z., Ren,X.B. and Cao,S. (2012) MiR-205 in cancer: an angel or a devil? *Eur. J. Cell Biol.*, **92**, 54–60.
31. Kim,J.S., Yu,S.K., Lee,M.H., Park,M.G., Park,E., Kim,S.G., Lee,S.Y., Kim,C.S., Kim,H.J., Chun,H.S. *et al.* (2013) MicroRNA-205 directly regulates the tumor suppressor, interleukin-24, in human KB oral cancer cells. *Mol. Cells*, **35**, 17–24.
32. Pigazzi,M., Manara,E., Baron,E. and Basso,G. (2009) miR-34b targets cyclic AMP-responsive element binding protein in acute myeloid leukemia. *Cancer Res.*, **69**, 2471–2478.
33. Tsai,K.W., Wu,C.W., Hu,L.Y., Li,S.C., Liao,Y.L., Lai,C.H., Kao,H.W., Fang,W.L., Huang,K.H., Chan,W.C. *et al.* (2011) Epigenetic regulation of miR-34b and miR-129 expression in gastric cancer. *Int. J. Cancer.*, **129**, 2600–2610.
34. Corney,D.C., Flesken-Nikitin,A., Godwin,A.K., Wang,W. and Nikitin,A.Y. (2007) MicroRNA-34b and MicroRNA-34c are targets of p53 and cooperate in control of cell proliferation and adhesion-independent growth. *Cancer Res.*, **67**, 8433–8438.
35. Nohata,N., Hanazawa,T., Enokida,H. and Seki,N. (2012) microRNA-1/133a and microRNA-206/133b clusters: dysregulation and functional roles in human cancers. *Oncotarget*, **3**, 9–21.
36. Tran,M.N., Choi,W., Wszolek,M.F., Navai,N., Lee,I.L., Nitti,G., Wen,S., Flores,E.R., Siefker-Radtke,A., Czerniak,B. *et al.* (2013) The p63 protein isoform deltaNp63alpha inhibits epithelial-mesenchymal transition in human bladder cancer cells: ROLE OF MIR-205. *J. Biol. Chem.*, **288**, 3275–3288.
37. Li,D., Wang,Q., Liu,C., Duan,H., Zeng,X., Zhang,B., Li,X., Zhao,J., Tang,S., Li,Z. *et al.* (2012) Aberrant expression of miR-638 contributes to benzo(a)pyrene-induced human cell transformation. *Toxicol. Sci.*, **125**, 382–391.
38. Liu,N., Bezprozvannaya,S., Williams,A.H., Qi,X., Richardson,J.A., Bassel-Duby,R. and Olson,E.N. (2008) microRNA-133a regulates cardiomyocyte proliferation and suppresses smooth muscle gene expression in the heart. *Genes Dev.*, **22**, 3242–3254.
39. Hermeking,H. (2010) The miR-34 family in cancer and apoptosis. *Cell Death Differ.*, **17**, 193–199.
40. Chua,K. and Reed,R. (2001) An upstream AG determines whether a downstream AG is selected during catalytic step II of splicing. *Mol. Cell Biol.*, **21**, 1509–1514.
41. Kralovicova,J., Houngninou-Molango,S., Kramer,A. and Vorechovsky,I. (2004) Branch site haplotypes that control alternative splicing. *Hum. Mol. Genet.*, **13**, 3189–3202.
42. Cartegni,L., Chew,S.L. and Krainer,A.R. (2002) Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat. Rev. Genet.*, **3**, 285–298.
43. Manley,J.L. and Tacke,R. (1996) SR proteins and splicing control. *Genes Dev.*, **10**, 1569–1579.
44. Okamura,K., Hagen,J.W., Duan,H., Tyler,D.M. and Lai,E.C. (2007) The mirtron pathway generates microRNA-class regulatory RNAs in *Drosophila*. *Cell*, **130**, 89–100.
45. Ruby,J.G., Jan,C.H. and Bartel,D.P. (2007) Intronic microRNA precursors that bypass Drosha processing. *Nature*, **448**, 83–86.
46. Flynt,A.S., Greimann,J.C., Chung,W.J., Lima,C.D. and Lai,E.C. (2010) MicroRNA biogenesis via splicing and exosome-mediated trimming in *Drosophila*. *Mol. Cell*, **38**, 900–907.
47. Kim,V.N., Han,J. and Siomi,M.C. (2009) Biogenesis of small RNAs in animals. *Nat. Rev. Mol. Cell Biol.*, **10**, 126–139.
48. Lenasi,T., Peterlin,B.M. and Barboric,M. (2011) Cap-binding protein complex links pre-mRNA capping to transcription elongation and alternative splicing through positive transcription elongation factor b (P-TEFb). *J. Biol. Chem.*, **286**, 22758–22768.
49. Montes,M., Becerra,S., Sanchez-Alvarez,M. and Sune,C. (2012) Functional coupling of transcription and splicing. *Gene*, **501**, 104–117.

50. Bielewicz,D., Kalak,M., Kalyna,M., Windels,D., Barta,A., Vazquez,F., Szweykowska-Kulinska,Z. and Jarmolowski,A. (2013) Introns of plant pri-miRNAs enhance miRNA biogenesis. *EMBO Rep.*, **14**, 622–628.
51. Damgaard,C.K., Kahns,S., Lykke-Andersen,S., Nielsen,A.L., Jensen,T.H. and Kjems,J. (2008) A 5′ splice site enhances the recruitment of basal transcription initiation factors *in vivo*. *Mol. Cell*, **29**, 271–278.
52. Dye,M.J., Gromak,N. and Proudfoot,N.J. (2006) Exon tethering in transcription by RNA polymerase II. *Mol. Cell*, **21**, 849–859.
53. Lu,J., Kwan,B.C., Lai,F.M., Tam,L.S., Li,E.K., Chow,K.M., Wang,G., Li,P.K. and Szeto,C.C. (2012) Glomerular and tubulointerstitial miR-638, miR-198 and miR-146a expression in lupus nephritis. *Nephrology (Carlton)*, **17**, 346–351.

Advanced Review



Cross talk between spliceosome and microprocessor defines the fate of pre-mRNA

Chiara Mattioli, Giulia Pianigiani and Franco Pagani*

The spliceosome and the microprocessor complex (MPC) are two important processing machineries that act on precursor (pre)-mRNA. Both cleave the pre-mRNA to generate spliced mature transcripts and microRNAs (miRNAs), respectively. While spliceosomes identify in a complex manner correct splice sites, MPCs typically target RNA hairpins (pri-miRNA hairpins). In addition, pre-mRNA transcripts can contain pri-miRNA-like hairpins that are cleaved by the MPC without generating miRNAs. Recent evidence indicates that the position of hairpins on pre-mRNA, their distance from splice sites, and the relative efficiency of cropping and splicing contribute to determine the fate of a pre-mRNA. Depending on these factors, a pre-mRNA can be preferentially used to generate a miRNA, a constitutively or even an alternative spliced transcript. For example, competition between splicing and cropping on splice-site-overlapping miRNAs (SO miRNAs) results in alternative spliced isoforms and influences miRNA biogenesis. In several cases, the outcome of a pre-mRNA transcript and its final handling as miRNA or mRNA substrate can be frequently closely connected to the functional relationships between diverse pre-mRNA processing events. These events are influenced by both gene context and physiopathological conditions.

How to cite this article:
WIREs RNA 2014. doi: 10.1002/wrna.1236

miRNA BIOGENESIS AND SPLICING

MicroRNAs (miRNAs) are ~21-nucleotide (nt)-long noncoding RNAs and ~24,000 miRNA precursors belonging to 206 different species have been identified (www.mirbase.org; release: 20.0, June 2013). miRNAs have a fundamental role in gene expression[1] and their classical biosynthetic pathway begins in the nucleus (Figure 1), where the miRNA is transcribed by RNA polymerase II in a polyadenylated primary (pri-) miRNA transcript. The typical metazoan pri-miRNA contains an ~80-nt-long double-stranded RNA secondary structure, with a stem, a terminal loop, and single-stranded RNA flanking regions (referred here as pri-miRNA hairpin). The hairpin is recognized and cleaved, in a process defined as 'cropping', by the microprocessor complex (MPC), the macromolecular machinery whose basic components are the RNase III-type protein Drosha and its cofactor DGCR8.[2–4] DGCR8 recognizes the RNA substrate, whereas Drosha functions as an endonuclease. The resulting precursor (pre-) miRNA is exported to the cytoplasm by exportin-5.[5] There, the RNase III-type enzyme Dicer further processes pre-miRNA (dicing) to produce a 21- to 24-nt-long RNA duplex, which is then loaded on the RNA-induced silencing complex (RISC) containing the Argonaute (AGO) protein. The final single-stranded miRNA guides the AGO to the target mRNAs that will be downregulated either through mRNA translation inhibition or mRNA destabilization.[6] In addition, the MPC can have a function independent from its role in miRNA biogenesis (Figure 1). In this case, the cleavage of pri-miRNA-like hairpins is not associated with the production of a miRNA. One example is the MPC-dependent cleavage of

*Correspondence to: pagani@icgeb.org
Human Molecular Genetics, International Centre for Genetic Engineering and Biotechnology, Trieste, Italy

Conflict of interest: The authors have declared no conflicts of interest for this article.

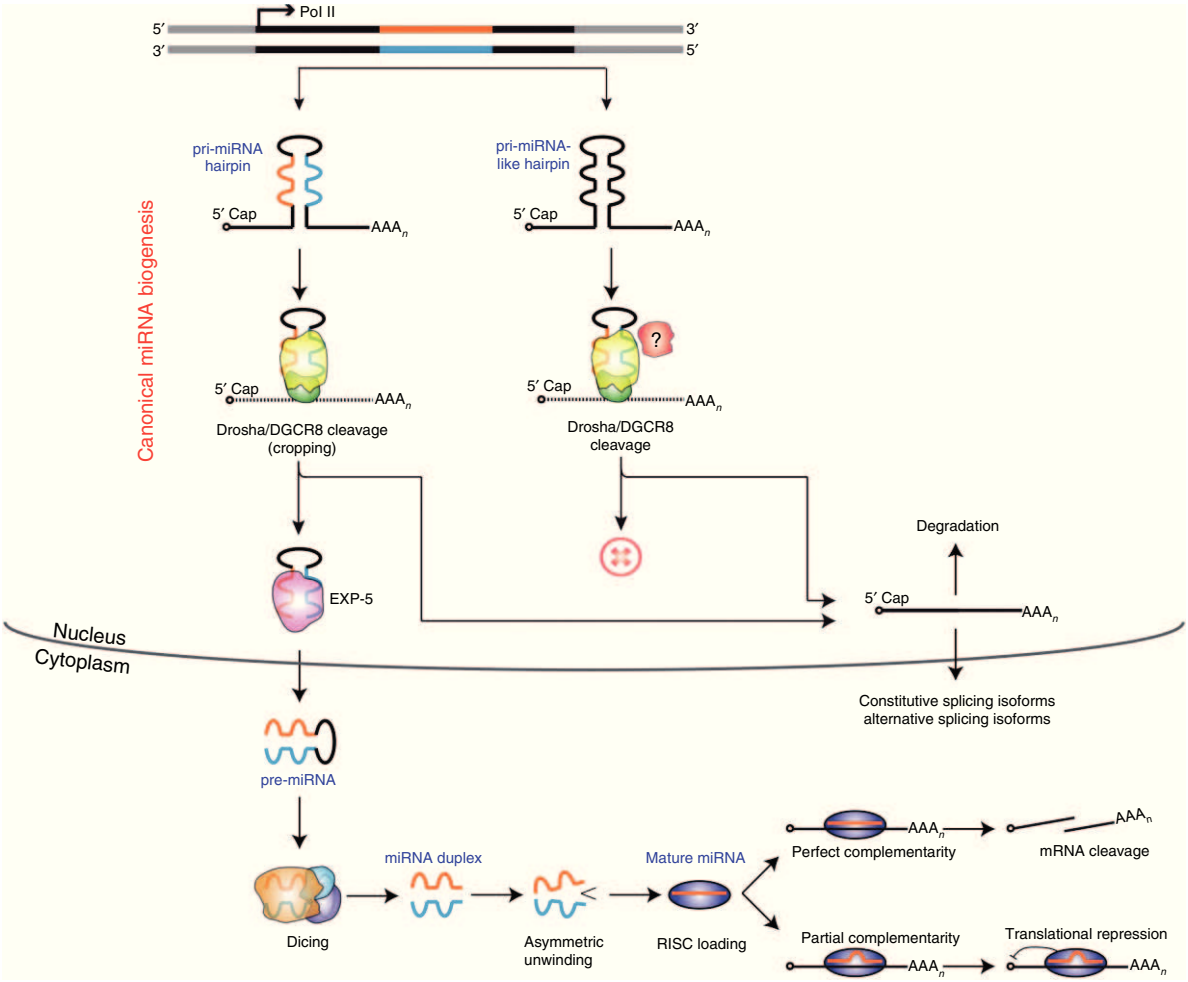


FIGURE 1 | Canonical microRNA (miRNA) biogenesis, processing of primary (pri)-miRNA-like hairpins, and fate of cleaved RNAs. Transcription of a coding or noncoding polymerase II gene generates capped and polyadenylated pre-mRNAs that contain stem–loop structures. In the canonical miRNA biosynthesis, the pri-miRNA hairpins are cleaved (cropping) by the Drosha/DGCR8 complex (microprocessor complex, MPC) to generate a ~70-nt pre-miRNA, which is recognized by exportin-5 (Exp-5). Following the export in the cytoplasm, Dicer catalyzes the second processing step—dicing—to produce a ~21-nt miRNA duplex. The miRNA duplex is unwound and one strand (orange) is selected as mature miRNA and incorporated in the RISC complex to function as guide molecule in cleavage or translational repression of target mRNAs, depending on the degree of complementarity between the miRNA and the target genes. The other strand (blue) of the miRNA duplex is degraded. In some cases, pri-miRNA-like structures are cleaved by the MPC or by unknown factors but do not produce a miRNA. The cleaved transcripts originated from pri-miRNA and pri-miRNA-like pre-mRNA can be either degraded in the nucleus or processed to generate normal or alternatively spliced isoforms, which are then exported to the cytoplasm.

the hairpin present in the 5′ UTR of *DGCR8* that directly regulates the mRNA (see below).[7] Furthermore, a recent global analysis of DGCR8 protein through HITS-CLIP (high-throughput sequencing of RNA isolated by cross-linking immunoprecipitation) experiments has identified predicted RNA secondary structures that resemble pri-miRNAs but are not associated with a miRNA.[8] In addition, Drosha binding to promoter-proximal regions of many human genes can regulate gene expression in an RNA cleavage-independent manner. This function of Drosha is dependent on its association with CPB80 and RNA polymerase II, probably requires imperfect hairpins on nascent RNA, and does not result in miRNA production or RNA cleavage.[9] The recognition and cleavage of the pri-miRNA by the MPC represents one of the key regulatory steps in miRNA biosynthesis, along with transcription and dicing.[10] The presence of a pri-miRNA hairpin on the nascent transcript may not be sufficient for the MPC-mediated cleavage and a large number of cofactors facilitates or inhibits the cropping activity. This can occur in different manners, either through modulation of components of the MPC or recognition of the miRNA

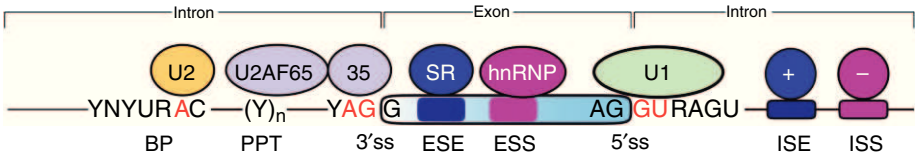


FIGURE 2 | Canonical splicing signal and auxiliary regulatory elements in precursor (pre)-mRNA splicing. Correct identification of exonic sequences requires classical 5′ and 3′ splice sites (5′ss and 3′ss) and a series of auxiliary regulatory elements. In the vast majority of splicing events (the so-called U1–U2 dependent), the 5′ss is composed of the invariant GU dinucleotide surrounded by partially conserved sequences and is recognized by U1 snRNP (U1). The 3′ss contains the AG dinucleotide, the polypyrimidine tract [(Y)n], and the branch point (BP), and interacts with U2 snRNP (U2) and proteins of the U2AF complex. Auxiliary *cis*-acting elements [enhancer and silencer, located in the exon (ESE, ESS) and/or in introns (ISE, ISS)] facilitate the exon recognition mainly through direct interaction with *trans*-acting factors: the SR proteins act on enhancers and have a positive effect on exon recognition by directly recruiting the splicing factors and/or by antagonizing the action of nearby silencer elements, while hnRNPs mediate splicing inhibition by sterically interfering with other splicing factors or antagonizing SR proteins. Light blue boxes represent the exons and thin line represents the intron. Y indicates pyrimidines, R purine, and N any nucleotide.

by accessory proteins, mainly RNA-binding proteins and/or splicing factors.[11] For example, binding of the splicing factors SRSF1 (SF2/ASF) and hnRNP A1 to the stem–loop of pri-miR-7 and pri-miR-18a, respectively, favors Drosha processing.[12–14] Similarly, the splicing protein KSRP (K-homology splicing regulatory protein) regulates the processing of a subset of miRNAs binding to a specific G-rich motif present in the terminal loop and stabilizing the interaction with Drosha.[15] Details on regulated cropping mechanisms have been extensively described in recent reviews.[16,17] In several cases, MPC and splicing act on the same transcript. During splicing the exon–intron junctions located on the pre-mRNA transcript are recognized and cut, the introns removed, and the exons joined to generate mature mRNA, which is then exported to the cytoplasm.[18] The 5′ and 3′ splice sites (5′ss and 3′ss) at the junctions have the nearly invariant GU and AG dinucleotides, respectively. The branch point and the polypyrimidine tract contribute to the 3′ss definition, whereas the 5′ss is complementary to the U1 snRNA (Figure 2). Auxiliary elements, classified as exonic or intronic splicing enhancers (ESEs or ISEs) and silencers (ESSs or ISSs), influence the recognition of the splice sites in a positive or negative manner. These elements are also important in alternative splicing, a mechanism that increases transcript diversity through the selection of alternative splice sites.[19,20] A large number of *trans*-acting factors recognize these auxiliary elements. In general, they have a positive or a negative role on exon definition, similar to the serine/arginine (SR)-rich proteins and the heterogeneous nuclear ribonucleoproteins (hnRNPs), but in several cases their effect is position and context dependent.[21] At the end, the final decision to include or exclude a sequence in the mature mRNA is the result of a complex combinatorial control between the strength of splicing regulatory elements in *cis*- and the different *trans*-acting factors that are assembled on the pre-mRNA.[22]

Both splicing and cropping are known to occur co-transcriptionally. In this review, we analyze the effect of the position of the hairpins targeted by the MPC in the pre-mRNA on the fate of the final transcript and on miRNA biosynthesis.

LOCALIZATION OF pri-miRNA AND pri-miRNA-LIKE HAIRPINS IN pre-mRNA TRANSCRIPTS

The available information regarding the position of the miRNA hairpins in the 'transcriptome' is mainly derived from bioinformatics analysis with gene and miRNA annotations and it is common that most of them are located in introns.[23] Analysis of 541 pri-miRNA hairpins in the human genome showed that ~62% of miRNAs are located in introns (43 and 19% in coding and noncoding genes, respectively), 9% in exons or UTRs, and 2% in alternatively spliced exons.[23] A small fraction of annotated miRNA hairpins (1–2%) are also located either on donor or acceptor splice sites (splice-site-overlapping miRNA, SO miRNA).[24,25] In addition, according to recent CLIP data in HEK 293T cells, 45% of DGCR8 targets are mapped to intergenic regions, 43% to protein-coding genes, and 5% to long noncoding RNAs. Thirty percent of the genomic clusters mapped to repetitive elements. In protein-coding genes, 26% of the clusters were in introns, followed by exons (12%). The exonic ones also included those located in 3′ and 5′ UTRs.[8] Most of these DGCR8 targets are real pri-miRNA hairpins but a significant amount does not have a corresponding miRNA in databases. These pri-miRNA-like hairpins have a major role in regulating the fate or quality of the transcript.[8]

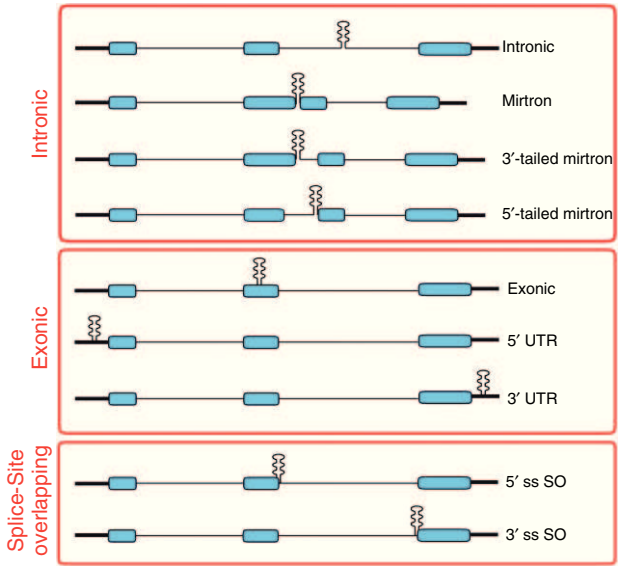


FIGURE 3 | Genomic location of primary (pri)-microRNA (miRNA) and pri-miRNA-like hairpins. Pri-miRNA and pri-miRNA-like hairpins are located in different positions relative to the splice sites and the gene in both coding and noncoding transcripts. The intronic group contains the intronic hairpins, the mirtrons, and the tailed mirtrons. The exonic group includes hairpins located in the central exons and in the 5′ and 3′ UTRs (or in corresponding first and terminal exons for noncoding transcripts). In splice-site overlapping (SO), the hairpins overlap with 5′ or 3′ splice sites. Blue boxes represent the exons, thin lines the introns, and thick lines the 5′ or 3′ UTRs.

According to their position in the nascent transcript, pri-miRNA and pri-miRNA-like hairpins can be located in introns (including mirtrons and tailed mirtrons), exons, 5′ or 3′ UTRs, or they can overlap with splice sites in both coding and noncoding genes (Figure 3).

Intronic Hairpins

Pure Intronic miRNAs Located at Distance from Splice Sites

The preferential location of miRNA hairpins in introns has been interpreted as a way to avoid the interference between the MPC and the spliceosome.[23,26] Accordingly, the most accepted view on intronic miRNAs, located far from splice sites, proposes that miRNAs are produced without altering the levels of the mature mRNA[26] and that miRNA production occurs co-transcriptionally before intron splicing.[23] According to the exon-tethering model, the C-terminal domain of polymerase II can tether an upstream exon before reaching the next one, even if the intervening sequences are cleaved by the MPC.[27,28] In this case, the MPC can excise the pre-miRNA co-transcriptionally from the intron, without affecting the recognition of the flanking splice sites by the spliceosome. This interesting model might be important for longer introns before the polymerase reaches the next exon, whereas in shorter introns miRNA cropping might occur when the two flanking exons are already engaged in splicing or even later on the excised lariat. In any case, as a common precursor generates both the miRNA and the mRNA, one pre-mRNA molecule will produce one miRNA precursor and one mRNA. However, in experimental systems, where the miRNA hairpins are relatively distant from splice sites, MPC-dependent cropping was shown to have a small influence on splicing, either in a positive[23,29] or negative[26] manner. Furthermore, spliceosome assembly can facilitate cropping.[29,30] In the case of the melanoma invasion suppressor miR-211, expressed from intron 6 of melastatin, mutation of the 5′ss was shown to reduce the biosynthesis of the miRNA and the recruitment of Drosha to the miRNA hairpin. Moreover, knockdown of the U1 snRNP-specific factor snRNP70 globally altered the production of intronic miRNAs,[29] suggesting that the 5′ss has a positive influence on cropping of intronic miRNAs.

Mirtrons Are MPC Independent and Require Splicing

Mirtrons originate from pri-miRNA-shaped introns; they were first described in *Drosophila melanogaster*[31,32] and validated recently in mammals.[33] Of 737 small RNAs in the human and mouse genome, 41 have been classified as *bona fide* mirtrons.[33] The main characteristic of mirtrons is that the 5′ and 3′ ends of their hairpins exactly correspond to splice-site junctions and therefore, they are not processed by the canonical MPC but directly by the splicing machinery. The spliced pre-mirtron is then exported to the cytoplasm where it follows the classical pathway. Other two variants of splicing-dependent miRNAs have been described, 3′- and 5′-tailed mirtrons. The hairpin of miR-1017 has the 5′ end that corresponds to the splice site and a 100-nucleotide-long tail that separates it from the 3′ss.[34] In this case, the entire intron is cleaved first to generate an intermediate in which the 3′ tail is processed by an unknown exonuclease to generate the pre-miRNA. Bioinformatic analysis has also identified other potential 3′-tailed mirtrons as well as 5′-tailed mirtrons.[33,35] In 5′-tailed mirtrons, the 3′ end of the hairpin end with the splice site and a tail separates the secondary structure from the 5′ss. For these cases experimental validation is lacking. To further complicate the picture, two predicted mirtrons (miR-1225 and miR-1228) are produced in the absence of splicing. The biogenesis of these so-called simtrons (splicing-independent mirtron-like

RNAs) is Drosha dependent but does not require DGCR8, Dicer, or the splicing process.[36] It is possible that the initial processing of simtrons is performed by an alternative Drosha complex in which DGCR8 is replaced by a yet unidentified adaptor for proper positioning and cleavage of the pri-simtron.

Exonic Hairpins

Exonic pri-miRNA or pri-miRNA-like hairpins constitute a small part of the known human miRNAs. Most of them are located in noncoding genes or in UTRs of coding transcripts, possibly to avoid the interference with the protein-coding sequences and splicing regulatory elements. However, there are interesting examples in which the exonic position of the hairpin influences the quality of the mRNA or its function.

Analysis of DGCR8 CLIP has identified some hairpin structures located in alternatively spliced exons of coding genes.[8] Detailed analysis of four cases showed that silencing of either Drosha or DGCR8, but not of Dicer, changes their pattern of splicing, suggesting that the MPC specifically cleaves and destabilizes the mRNA isoforms that contain the hairpin. This is an interesting and novel way to regulate alternative splicing and is similar to the effect described for one mouse SO miRNA (see below). It is not clear if the cleavage products lead to the generation of miRNAs and it would be interesting to understand if the MPC-dependent cleavage occurs co-transcriptionally or after the splicing decision.

A second intriguing example is the pri-miR-133b located in the terminal exon of the noncoding *linc-MD1* gene, whose biosynthesis is regulated by the splicing factor HuR.[37] linc-MD1 acts in the cytoplasm repressing target genes that promote myogenesis and this effect is partially mediated by its 'sponge' activity on its own producing miR-133b. If Drosha crops pri-miR-133b in the nucleus, the pre-mRNA linc-MD1 will be used to generate the miRNA; alternatively, if linc-MD1 is not processed by the MPC, it will act as a sponge in the cytoplasm for miR-133b and probably for other miRNAs of the same family.[37–39] Thus, the decision to cleave a pri-miRNA hairpin located in a terminal exon switches between the two alternative and antagonistic gene products: a cytoplasmic sponge RNA and a miRNA. This might not be the only case in which a hairpin can switch between a sponge and a miRNA, as several pri-miRNA hairpins are located in the terminal exon of noncoding transcript.

In plants, pri-miRNA hairpins are frequently located in exons of independently transcribed units and downstream introns enhance miRNA production.[40] According to Jarmolowski and coworkers,[41] a functional 5′ss has a major role in miRNA biosynthesis, suggesting that U1 snRNP can promote cropping. Interestingly, a similar enhancing role of U1 snRNP has been proposed for the intronic miR-211 in humans.[29] Moreover, dcl-1 mutants (Dicer-like 1, the enzyme that cleaves the stem–loop in plants) increase intron splicing, suggesting a competition between splicing and miRNA processing.[42] Even if it is not clear which are the intronic factors involved, the exonic position of the hairpins in plants is necessary to improve miRNA biogenesis.

The pri-miRNA-Like Hairpin in the 5′ UTR of DGCR8

The stem–loop in the 5′ UTR of *DGCR8* is the only example of a miRNA hairpin located in the 5′ UTR of a coding gene with an important effect on the fate of the transcript. This structure is a substrate for the MPC and, once it is cropped, the transcript cannot be processed further, with a decrease in both the DGCR8 mRNA and protein.[7,43] The presence of a second hairpin in the first exon also contributes to the DGCR8 regulation.[7,43] The 5′ UTR hairpin of DGCR8 can be considered a pri-miRNA-like hairpin, as the mature miRNA has never been detected, neither with northern blot nor with massive parallel sequencing. Therefore, this structure mainly promotes mRNA degradation and is a useful system to finely tune the levels of DGCR8. Moreover, it constitutes a feedback loop mechanism to control the levels of the MPC itself, because DGCR8 stabilizes Drosha through protein–protein interactions.[7]

The pri-miRNA Hairpin in the 3′ UTR of FSTL1

MPC-dependent cropping of hairpin structures in the 3′ UTR may have important consequences on the transcript and on miRNA biosynthesis, in particular when linked to an upstream open reading frame. A paradigmatic example of a 3′ UTR MPC-dependent regulation occurs in the follistatin-like 1 (*FSTL1*) gene where it switches a protein coding toward a miRNA-producing transcript in wound healing[44] (Figure 4). The *FSTL1* gene is composed of 11 exons and contains the primate-specific pri-miR-198 in the 3′ UTR. In normal conditions, human skin keratinocytes express miR-198 but not FSTL1. After wound healing, cells start to migrate, produce FSTL1, and progressively reduce miR-198 expression. This is not associated with significant changes in the FSTL1 pre-mRNA, suggesting a post-transcriptional regulation. Indeed, in normal conditions, the transcripts are retained in the nucleus where they are used to produce only the miRNA. After injury, the mRNA

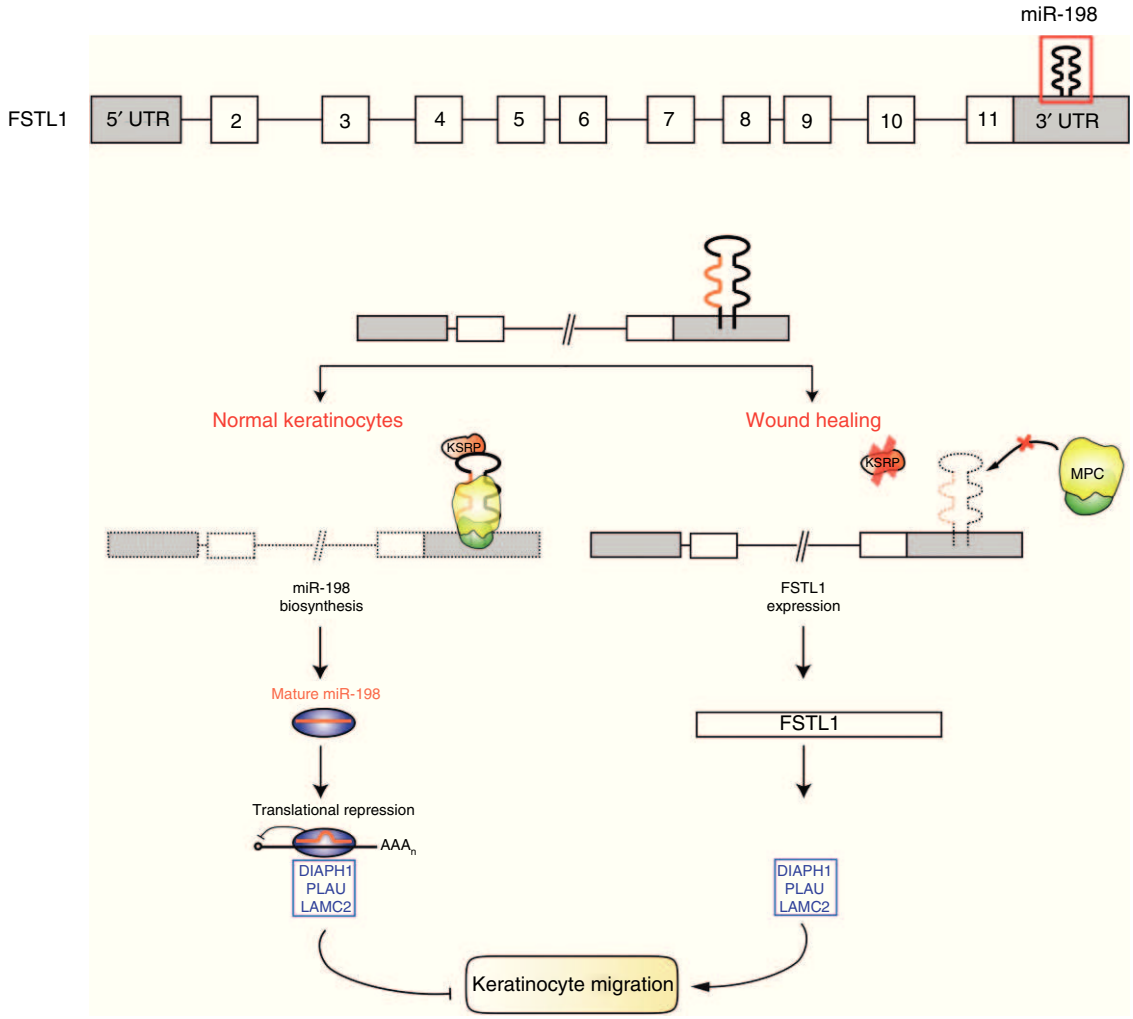


FIGURE 4 | Cropping of a primary (pri)-microRNA (miRNA) hairpin in the 3′ UTR of *FSTL1* regulates keratinocytes proliferation/migration. Pri-miR-198 is located in the 3′ UTR of the follistatin-like 1 (FSTL1), a protein involved in keratinocytes migration. In normal keratinocytes, binding of KSRP (K-homology splicing regulatory protein) to the terminal loop of the hairpin promotes microprocessor complex (MPC)-dependent cropping of miR-198 that in turn downregulates the expression of promigratory protein targets (DIAPH1, PLAU, and LAMC2).[44] After injury, a series of events induced by Transforming growth factor - β (TGF-β) result in downregulation of KSRP and inhibition of MPC-dependent cleavage. This reduces the amount of miR-198 and activates promigratory proteins. In parallel, the expression of FSTL1 protein is increased. Gray boxes indicate 5′ and 3′ UTR regions; white boxes and thin line represent exons and introns, respectively.

is not cleaved by the MPC and goes to the cytoplasm where it is translated into protein. Important controls exclude that miR-198 acts directly on FSTL1 mRNA and silencing of FSTL1 does not affect miR-198 expression. As FSTL1 and mir-198 have antimigratory and promigratory effects, respectively, the switch between the two alternative gene products contributes to wound re-epithelialization (Figure 4). Interestingly, the switch fails in nonhealing diabetes ulcers, where cells continue to express miR-198 and FSTL1 is absent. A major player in regulating this switch is the RNA-binding protein KSRP. This factor was previously shown to regulate cropping efficiency of a set of miRNAs through its interaction with G-rich sequences in the stem–loop.[15] KSRP binds to a GUG motifs within the terminal loop of pri-miR-198 and promotes its cropping. As silencing of this factor nearly completely inhibits miR-198 synthesis, it seems that KSRP is strictly required for miRNA processing. In the absence of KSRP, the pri-miR-198 hairpin is a poor substrate for the MPC. As the 3′ UTR MPC-dependent hairpin found in FSTL1 is proximal to the polyadenylation site, it is not clear if the miRNA cropping occurs before or after the polyadenylation process. However, it is possible that in particular conditions the miRNA cropping can function as an alternative 3′-end processing mechanism in competition with the classical one, as Drosha silencing was shown

TABLE 1 | List of Human SO miRNAs

miRNA	Chr	Conservation	Gene	References
3′ SO miRNAs				
205	1	Y	MIR205HG (nc)	45,59,60
943	4	N	WHSC2	
936	10	N	COL17A1	
1287	10	Y	PYROXD2	
1307	10	N	USMG5	
34b	11	Y	BC021736 (nc)	52–54,56–58,61,62
1178	12	N	CIT	
636[1]	17	N	SFRS2	
4315-2	17	Y	PLEKHM1P	
4321	19	N	AMH	
133a-2	20	Y	C20orf166	50,63
1292	20	N	NOP56	
5′ SO miRNAs				
4260	1	N	LAMB3	
761	1	Y	NRD1	
555	1	Y	ASH1L	
1204	8	N	MYC (nc)	
611	11	N	TMEM258	
3656	11	N	TRAPPC4	
638	19	Y	DNM2	64,65
3′ → 5′ SO miRNAs				
711	3	N	COL7A1	
937[1]	8	N	SCRIB	
SO miRNAs on ESTs				
3916	1	N		
1248	3	Y		66
302b	4	Y		67–69
3150b	8	N		
202	10	Y		70,71
675	11	N		72–75
3613	13	Y		
627	15	N		76
3614	17	N		
365-2	17	Y		77

nc, noncoding; SO miRNAs, splice-site-overlapping microRNAs; pri-miRNA, primary microRNA; EST, expressed sequence tags.
3′ SO miRNAs and 5′ SO miRNAs are pri-miRNA hairpins that overlap with the 3′ or the 5′ss, respectively.[24,25] In the 3′ → 5′ SO miRNAs group, the miRNA secondary structure includes both donor and acceptor sites. SO miRNAs on ESTs correspond to pri-miRNA hairpins that overlap with intron–exon junctions of spliced ESTs for which there is no clear gene annotation.
[1]miRNAs classified as potential tailed mirtrons.[33,35]

to increase the levels of FSTL1 protein. This peculiar type of post-transcriptional FSTL1/miR-198 regulation might be present in other genes that contain miRNA hairpins located on 3′ UTR. For example, the coding *DCP1A*, *HOXA7*, *INO80E*, *SNX12* and *NLS1* genes contain pri-miRNA-like secondary structures in their 3′ UTR, which are DGCR8-binding sites according to CLIP data.[8] These mRNAs are probably direct targets of the MPC, as depletion of Drosha and/or DGCR8 resulted in an increase of their mRNA levels and because they are cleaved *in vitro* by the MPC as normal pri-miRNAs. This MPC-dependent regulation also affects the corresponding protein levels.[8] In these cases, however, the putative factor(s) that regulate the cleavage efficiency are not known.

Splice-Site-Overlapping miRNA Hairpins

The position of the SO miRNA hairpins in the nascent transcript determines a direct competition between two pre-mRNA processing machineries: the MPC and the spliceosome. SO miRNAs are characterized by the presence of overlapping miRNA hairpins and splice sites on the nascent transcripts.[24,25] The hairpins can be located either on donor or acceptor splice sites (Table 1). They are present in vertebrates and nonvertebrates and represent a small but significant fraction of annotated miRNAs (1–2%). From the evolutionary point of view, in several cases their location on the splice sites is conserved and most are in coding transcripts. Some SO miRNAs have a clear physiopathological role in humans. For example, miR-205 is described both as a tumor suppressor and an oncogene, depending on the tumor considered and the target genes,[45] and has a protective role against oxidative stress and endoplasmic reticulum stress in renal tubular cells[46]; miR-133a-2 is a tumor suppressor deregulated in colorectal and bladder cancers[47,48] and is involved in maintaining the structure and biogenesis of skeletal muscle[49] and heart.[50] Two 3′ss SO miRNAs embedded in noncoding transcripts, the evolutionarily conserved miR-34b and the mouse-specific mmu-miR-412, have been analyzed in detail and their biological function is known. The SO miR-34b is part of a noncoding transcript composed of two exons and is located on the junction between intron 1 and exon 2. The same transcriptional unit includes miR-34c in exon 2 and the conserved elements of the gene are a p53-responsive promoter region,[51] the splice sites, the miRNA hairpins, and the polyadenylation site (Figure 5). Originally identified as a tumor suppressor miRNA,[52–54] miR-34b has roles in cardiomyocytes proliferation,[55] neurological disorders,[56,57] and

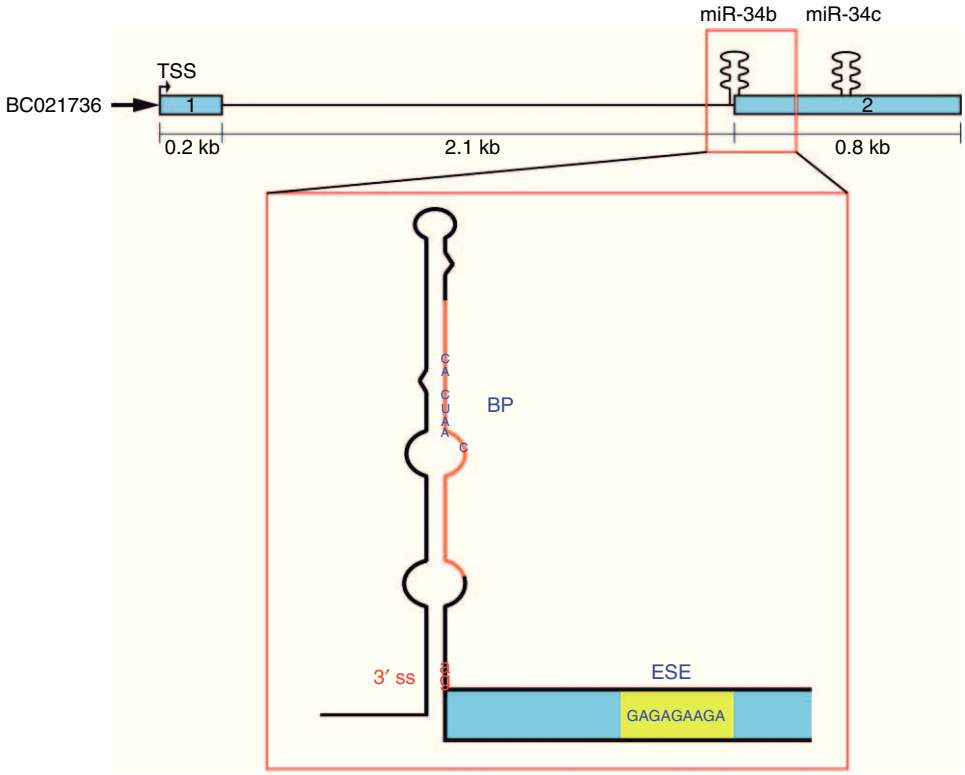


FIGURE 5 | miR-34b/c transcript and miR-34b secondary structure. BC021736 is the noncoding transcript that contains the pri-miR-34b and pri-miR-34c. The zoomed image of splice-site-overlapping (SO) miR-34b secondary structure shows the elements involved in splicing regulation: the AG dinucleotide of the 3′ss, which is embedded in the hairpin, the branch point (BP) consensus, and the downstream 9-bp purine-rich exonic enhancer (ESE). The mature form of the miRNA is shown in orange. Blue boxes and thin line represent exons and intron, respectively. The black arrow indicates the promoter and TSS is the transcription start site.

osteoblast differentiation,[58] and has been proposed as a serum marker for Huntington's disease.[56]

In common with other two 3′ SO miRNAs (miR-205 and miR-133a-2), miR-34b does not have a polypyrimidine tract and the splice-site strength of the acceptor site, embedded in the miR-34b hairpin, is below the threshold of detection using bioinformatics programs. However, the 3′ss is used and the spliced and unspliced forms are present *in vivo* in human and mouse tissues. Detailed molecular analysis showed that the definition of the 3′ss depends on two splicing regulatory elements: a branch point located in miR-34b hairpin and a purine-rich enhancer in exon 2. The *trans*-acting factor(s) binding to the ESE were not identified, and silencing of splicing factors known to bind to purine-rich sequences (SRSF1 and Tra2β) had no effect on miR-34b splicing. Interestingly, in minigene assay, splicing inhibition performed either through mutations of the ESE or of the AG in the 3′ss increased miR-34b levels, indicating that splicing influences miRNA biosynthesis. On the contrary, mutation of the 5′ss reduced the levels of miR-34b. The requirement of a functional 5′ss has also been observed for human intronic[29] and plant miRNAs,[41] suggesting that recruitment of U1 snRNP at the donor site influences the MPC activity. On the other hand, silencing of the MPC components Drosha and DGCR8 increased the splicing efficiency, whereas their overexpression had the opposite effect. Because of the overlap between the hairpin and the splice sites, the biosynthesis of SO miRNAs (and production of mature mRNAs) is regulated by the competition between the MPC and the spliceosome (Figure 6) and changes in the splicing efficiency, regulated by the ESE, will affect the levels of mature miRNAs.

When SO miRNAs are located in central exons, the competition between the MPC and the spliceosome can also occur on alternatively spliced isoforms. For example, the analysis of the 3′ SO mmu-miR-412,[25] which is located on an alternatively spliced exon, indicated that its tissue-specific alternative usage regulates miRNA biosynthesis. In minigene system, the induction of exon inclusion by the splicing factor TIA1 reduced the amount of pre-mRNA substrate for miRNA biosynthesis.

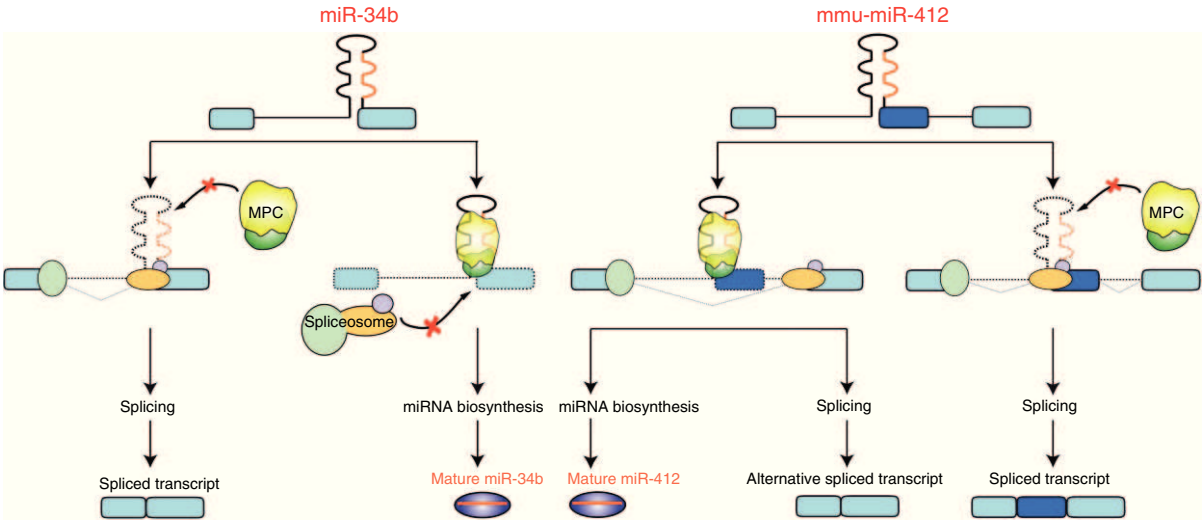


FIGURE 6 | Competition between spliceosome and microprocessor complex (MPC) on splice-site-overlapping (SO) microRNAs (miRNAs). The SO miR-34b and miR-412 hairpins are located on two different noncoding genes and overlap with corresponding 3′ss: the first is located in the last exon and the second in an internal exon. In miR-34b, when splicing processing prevails the transcript is preferentially spliced and the mature miR-34b is not produced. On the contrary, reduction in splicing or preferential recognition of the hairpin by the MPC leads to miR-34b maturation.[24,25] In miR-412, when the central exon is skipped, the MPC crops the pri-miRNA hairpin, promoting the production of the mature miRNA. On the other hand, when the alternative exon is included, the primary transcript cannot be processed by the MPC and the miRNA is not synthesized.[24,25] The continuous and dashed lines indicate the processed and unprocessed part of the transcript.

CONCLUSION

Pri-miRNA or pri-miRNA-like hairpins can be located in different positions on the nascent transcripts. In the most common intronic location, the pre-mRNA produces both the miRNA and the mature RNA through cropping and splicing, respectively. The general view regarding intronic miRNAs considers that splicing and cropping act in an independent manner. However, these two events may be subtly co-regulated in a context-dependent manner, according to the presence of intronic splicing regulatory signals or exon tethering. In mirtrons, an accurate superimposition of splice sites with the pre-miRNA tails makes the MPC useless, and therefore the initiation of miRNA biosynthesis will depend exclusively on splicing. On the contrary, in SO miRNAs the direct competition between the cropping and splicing machineries on the nascent transcript will affect miRNA production and/or alternative splicing. A detailed analysis of SO miRNAs behavior has been performed on a couple of interesting noncoding transcripts. However, most human SO miRNAs are located on intron or exon junctions of coding genes, hence their regulated cropping or splicing could affect final protein expression. In these cases, the competition between the MPC and the spliceosome might either change alternative splicing or result in a switch between mRNA and miRNA. In the *FSTL1* gene, the regulated cleavage of a pri-miRNA hairpin in the 3′ UTR allows a switch between a miRNA and an mRNA with antagonistic functions on keratinocyte proliferation. The recent identification of pri-miRNA-like hairpins in the 3′ UTR of other coding genes might have similar mRNA regulatory implications, which extend to the yet unidentified miRNAs that would be produced as a result. Lastly, exonic hairpins may regulate alternative splicing or act as miRNA sponges: these effects reported for few cases can be more widespread than currently envisaged. In conclusion, the position of pri-miRNA or pri-miRNA-like hairpins on nascent transcripts relative to splice sites depending on gene context can have a critical influence on gene expression through changes in miRNA biosynthesis and/or by determining the fate of the processed mature RNA. Future studies will clarify in more detail what mechanisms may coordinate primary transcript processing into both mature miRNA and spliced (or alternatively spliced) mRNA. Finally, it will be crucially important to establish the functional consequences of the processing regulation events discussed in the context of physiopathological conditions.

ACKNOWLEDGMENTS

We thank Emanuele Buratti, Rodolfo Garcia, and Francisco E Baralle for critical reading of the manuscript. This work was supported by research grants from Associazione Italiana per la Ricerca sul Cancro (AIRC) 10387.

REFERENCES

1. Pillai RS. MicroRNA function: multiple mechanisms for a tiny RNA? *RNA* 2005, 11:1753–1761.
2. Gregory RI, Yan KP, Amuthan G, Chendrimada T, Doratotaj B, Cooch N, Shiekhattar R. The microprocessor complex mediates the genesis of microRNAs. *Nature* 2004, 432:235–240.
3. Denli AM, Tops BB, Plasterk RH, Ketting RF, Hannon GJ. Processing of primary microRNAs by the microprocessor complex. *Nature* 2004, 432:231–235.
4. Han J, Lee Y, Yeom KH, Kim YK, Jin H, Kim VN. The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev* 2004, 18:3016–3027.
5. Lund E, Dahlberg JE. Substrate selectivity of exportin 5 and Dicer in the biogenesis of microRNAs. *Cold Spring Harb Symp Quant Biol* 2006, 71:59–66.
6. Fabian MR, Sonenberg N, Filipowicz W. Regulation of mRNA translation and stability by microRNAs. *Annu Rev Biochem* 2010, 79:351–379.
7. Han J, Pedersen JS, Kwon SC, Belair CD, Kim YK, Yeom KH, Yang WY, Haussler D, Blelloch R, Kim VN. Posttranscriptional crossregulation between Drosha and DGCR8. *Cell* 2009, 136:75–84.
8. Macias S, Plass M, Stajuda A, Michlewski G, Eyras E, Caceres JF. DGCR8 HITS-CLIP reveals novel functions for the microprocessor. *Nat Struct Mol Biol* 2012, 19:760–766.
9. Gromak N, Dienstbier M, Macias S, Plass M, Eyras E, Caceres JF, Proudfoot NJ. Drosha regulates gene expression independently of RNA cleavage function. *Cell Rep* 2013, 5:1499–1510.
10. Kim VN, Han J, Siomi MC. Biogenesis of small RNAs in animals. *Nat Rev Mol Cell Biol* 2009, 10:126–139.
11. Choudhury NR, Michlewski G. Terminal loop-mediated control of microRNA biogenesis. *Biochem Soc Trans* 2012, 40:789–793.
12. Guil S, Caceres JF. The multifunctional RNA-binding protein hnRNP A1 is required for processing of miR-18a. *Nat Struct Mol Biol* 2007, 14:591–596.
13. Michlewski G, Sanford JR, Caceres JF. The splicing factor SF2/ASF regulates translation initiation by enhancing phosphorylation of 4E-BP1. *Mol Cell* 2008, 30:179–189.
14. Wu H, Sun S, Tu K, Gao Y, Xie B, Krainer AR, Zhu J. A splicing-independent function of SF2/ASF in microRNA processing. *Mol Cell* 2010, 38:67–77.
15. Trabucchi M, Briata P, Garcia-Mayoral M, Haase AD, Filipowicz W, Ramos A, Gherzi R, Rosenfeld MG. The RNA-binding protein KSRP promotes the biogenesis of a subset of microRNAs. *Nature* 2009, 459:1010–1014.
16. Krol J, Loedige I, Filipowicz W. The widespread regulation of microRNA biogenesis, function and decay. *Nat Rev Genet* 2010, 11:597–610.
17. Finnegan EF, Pasquinelli AE. MicroRNA biogenesis: regulating the regulators. *Crit Rev Biochem Mol Biol* 2013, 48:51–68.
18. Matlin AJ, Clark F, Smith CW. Understanding alternative splicing: towards a cellular code. *Nat Rev Mol Cell Biol* 2005, 6:386–398.
19. Chen M, Manley JL. Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. *Nat Rev Mol Cell Biol* 2009, 10:741–754.
20. Nilsen TW, Graveley BR. Expansion of the eukaryotic proteome by alternative splicing. *Nature* 2010, 463:457–463.
21. Zhou Z, Fu XD. Regulation of splicing by SR proteins and SR protein-specific kinases. *Chromosoma* 2013, 122:191–207.
22. De Conti L, Baralle M, Buratti E. Exon and intron definition in pre-mRNA splicing. *Wiley Interdiscip Rev RNA* 2013, 4:49–60.
23. Morlando M, Ballarino M, Gromak N, Pagano F, Bozzoni I, Proudfoot NJ. Primary microRNA transcripts are processed co-transcriptionally. *Nat Struct Mol Biol* 2008, 15:902–909.
24. Mattioli C, Pianigiani G, Pagani F. A competitive regulatory mechanism discriminates between juxtaposed splice sites and pri-miRNA structures. *Nucleic Acids Res* 2013, 41:8680–8691.
25. Melamed Z, Levy A, Ashwal-Fluss R, Lev-Maor G, Mekahel K, Atias N, Gilad S, Sharan R, Levy C, Kadener S, et al. Alternative splicing regulates biogenesis of miRNAs located across exon-intron junctions. *Mol Cell* 2013, 50:869–881.
26. Kim YK, Kim VN. Processing of intronic microRNAs. *EMBO J* 2007, 26:775–783.
27. Dye MJ, Gromak N, Proudfoot NJ. Exon tethering in transcription by RNA polymerase II. *Mol Cell* 2006, 21:849–859.
28. Pastor T, Pagani F. Interaction of hnRNPA1/A2 and DAZAP1 with an Alu-derived intronic splicing enhancer regulates ATM aberrant splicing. *PLoS One* 2011, 6:e23349.
29. Janas MM, Khaled M, Schubert S, Bernstein JG, Golan D, Veguilla RA, Fisher DE, Shomron N, Levy C, Novina

CD. Feed-forward microprocessing and splicing activities at a microRNA-containing intron. *PLoS Genet* 2011, 7:e1002330.

30. Pawlicki JM, Steitz JA. Primary microRNA transcript retention at sites of transcription leads to enhanced microRNA production. *J Cell Biol* 2008, 182:61–76.
31. Okamura K, Hagen JW, Duan H, Tyler DM, Lai EC. The mirtron pathway generates microRNA-class regulatory RNAs in *Drosophila*. *Cell* 2007, 130:89–100.
32. Ruby JG, Jan CH, Bartel DP. Intronic microRNA precursors that bypass Drosha processing. *Nature* 2007, 448:83–86.
33. Ladewig E, Okamura K, Flynt AS, Westholm JO, Lai EC. Discovery of hundreds of mirtrons in mouse and human small RNA data. *Genome Res* 2012, 22:1634–1645.
34. Flynt AS, Greimann JC, Chung WJ, Lima CD, Lai EC. MicroRNA biogenesis via splicing and exosome-mediated trimming in *Drosophila*. *Mol Cell* 2010, 38:900–907.
35. Valen E, Preker P, Andersen PR, Zhao X, Chen Y, Ender C, Dueck A, Meister G, Sandelin A, Jensen TH. Biogenic mechanisms and utilization of small RNAs derived from human protein-coding genes. *Nat Struct Mol Biol* 2011, 18:1075–1082.
36. Havens MA, Reich AA, Duelli DM, Hastings ML. Biogenesis of mammalian microRNAs by a non-canonical processing pathway. *Nucleic Acids Res* 2012, 40:4626–4640.
37. Legnini I, Morlando M, Mangiavacchi A, Fatica A, Bozzoni I. A feedforward regulatory loop between HuR and the long noncoding RNA linc-MD1 controls early phases of myogenesis. *Mol Cell* 2014, 53:506–514.
38. Twayana S, Legnini I, Cesana M, Cacchiarelli D, Morlando M, Bozzoni I. Biogenesis and function of non-coding RNAs in muscle differentiation and in Duchenne muscular dystrophy. *Biochem Soc Trans* 2013, 41:844–849.
39. Cesana M, Cacchiarelli D, Legnini I, Santini T, Sthandier O, Chinappi M, Tramontano A, Bozzoni I. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell* 2011, 147:358–369.
40. Dhir A, Proudfoot NJ. Feed backwards model for microRNA processing and splicing in plants. *EMBO Rep* 2013, 14:581–582.
41. Bielewicz D, Kalak M, Kalyna M, Windels D, Barta A, Vazquez F, Szweykowska-Kulinska Z, Jarmolowski A. Introns of plant pri-miRNAs enhance miRNA biogenesis. *EMBO Rep* 2013, 14:622–628.
42. Schwab R, Speth C, Laubinger S, Voinnet O. Enhanced microRNA accumulation through stemloop-adjacent introns. *EMBO Rep* 2013, 14:615–621.
43. Triboulet R, Chang HM, Lapierre RJ, Gregory RI. Post-transcriptional control of DGCR8 expression by the microprocessor. *RNA* 2009, 15:1005–1011.
44. Sundaram GM, Common JE, Gopal FE, Srikanta S, Lakshman K, Lunny DP, Lim TC, Tanavde V, Lane EB, Sampath P. 'See-saw' expression of microRNA-198 and FSTL1 from a single transcript in wound healing. *Nature* 2013, 495:103–106.
45. Qin AY, Zhang XW, Liu L, Yu JP, Li H, Wang SZ, Ren XB, Cao S. MiR-205 in cancer: an angel or a devil? *Eur J Cell Biol* 2013, 92:54–60.
46. Muratsu-Ikeda S, Nangaku M, Ikeda Y, Tanaka T, Wada T, Inagi R. Downregulation of miR-205 modulates cell susceptibility to oxidative and endoplasmic reticulum stresses in renal tubular cells. *PLoS One* 2012, 7:e41462.
47. Yamasaki T, Yoshino H, Enokida H, Hidaka H, Chiyomaru T, Nohata N, Kinoshita T, Fuse M, Seki N, Nakagawa M. Novel molecular targets regulated by tumor suppressors microRNA-1 and microRNA-133a in bladder cancer. *Int J Oncol* 2012, 40:1821–1830.
48. Chen WS, Leung CM, Pan HW, Hu LY, Li SC, Ho MR, Tsai KW. Silencing of miR-1-1 and miR-133a-2 cluster expression by DNA hypermethylation in colorectal cancer. *Oncol Rep* 2012, 28:1069–1076.
49. Liu N, Bezprozvannaya S, Shelton JM, Frisard MI, Hulver MW, McMillan RP, Wu Y, Voelker KA, Grange RW, Richardson JA, et al. Mice lacking microRNA 133a develop dynamin 2-dependent centronuclear myopathy. *J Clin Invest* 2011, 121:3258–3268.
50. Liu N, Bezprozvannaya S, Williams AH, Qi X, Richardson JA, Bassel-Duby R, Olson EN. microRNA-133a regulates cardiomyocyte proliferation and suppresses smooth muscle gene expression in the heart. *Genes Dev* 2008, 22:3242–3254.
51. He L, He X, Lim LP, de Stanchina E, Xuan Z, Liang Y, Xue W, Zender L, Magnus J, Ridzon D, et al. A microRNA component of the p53 tumour suppressor network. *Nature* 2007, 447:1130–1134.
52. Kubo T, Toyooka S, Tsukuda K, Sakaguchi M, Fukazawa T, Soh J, Asano H, Ueno T, Muraoka T, Yamamoto H, et al. Epigenetic silencing of microRNA-34b/c plays an important role in the pathogenesis of malignant pleural mesothelioma. *Clin Cancer Res* 2011, 17:4965–4974.
53. Toyota M, Suzuki H, Sasaki Y, Maruyama R, Imai K, Shinomura Y, Tokino T. Epigenetic silencing of microRNA-34b/c and B-cell translocation gene 4 is associated with CpG island methylation in colorectal cancer. *Cancer Res* 2008, 68:4123–4132.
54. Corney DC, Hwang CI, Matoso A, Vogt M, Flesken-Nikitin A, Godwin AK, Kamat AA, Sood AK, Ellenson LH, Hermeking H, et al. Frequent downregulation of miR-34 family in human ovarian cancers. *Clin Cancer Res* 2010, 16:1119–1128.
55. Bernardo BC, Gao XM, Winbanks CE, Boey EJ, Tham YK, Kiriazis H, Gregorevic P, Obad S, Kauppinen S, Du XJ, et al. Therapeutic inhibition of the miR-34 family attenuates pathological cardiac remodeling and

improves heart function. *Proc Natl Acad Sci U S A* 2012, 109:17615–17620.

56. Gaughwin PM, Ciesla M, Lahiri N, Tabrizi SJ, Brundin P, Bjorkqvist M. Hsa-miR-34b is a plasma-stable microRNA that is elevated in pre-manifest Huntington's disease. *Hum Mol Genet* 2011, 20:2225–2237.
57. Minones-Moyano E, Porta S, Escaramis G, Rabionet R, Iraola S, Kagerbauer B, Espinosa-Parrilla Y, Ferrer I, Estivill X, Marti E. MicroRNA profiling of Parkinson's disease brains identifies early downregulation of miR-34b/c which modulate mitochondrial function. *Hum Mol Genet* 2011, 20:3067–3078.
58. Wei J, Shi Y, Zheng L, Zhou B, Inose H, Wang J, Guo XE, Grosschedl R, Karsenty G. miR-34s inhibit osteoblast proliferation and differentiation in the mouse by targeting SATB2. *J Cell Biol* 2012, 197: 509–521.
59. Kim JS, Yu SK, Lee MH, Park MG, Park E, Kim SG, Lee SY, Kim CS, Kim HJ, Chun HS, et al. MicroRNA-205 directly regulates the tumor suppressor, interleukin-24, in human KB oral cancer cells. *Mol Cells* 2013, 35:17–24.
60. Tran MN, Choi W, Wszolek MF, Navai N, Lee IL, Nitti G, Wen S, Flores ER, Siefker-Radtke A, Czerniak B, et al. The p63 protein isoform δNp63α inhibits epithelial-mesenchymal transition in human bladder cancer cells: role of MIR-205. *J Biol Chem* 2013, 288:3275–3288.
61. Pigazzi M, Manara E, Baron E, Basso G. miR-34b targets cyclic AMP-responsive element binding protein in acute myeloid leukemia. *Cancer Res* 2009, 69:2471–2478.
62. Tsai KW, Wu CW, Hu LY, Li SC, Liao YL, Lai CH, Kao HW, Fang WL, Huang KH, Chan WC, et al. Epigenetic regulation of miR-34b and miR-129 expression in gastric cancer. *Int J Cancer* 2011, 129:2600–2610.
63. Nohata N, Hanazawa T, Enokida H, Seki N. microRNA-1/133a and microRNA-206/133b clusters: dysregulation and functional roles in human cancers. *Oncotarget* 2012, 3:9–21.
64. Li D, Wang Q, Liu C, Duan H, Zeng X, Zhang B, Li X, Zhao J, Tang S, Li Z, et al. Aberrant expression of miR-638 contributes to benzo(a)pyrene-induced human cell transformation. *Toxicol Sci* 2012, 125:382–391.
65. Lu J, Kwan BC, Lai FM, Tam LS, Li EK, Chow KM, Wang G, Li PK, Szeto CC. Glomerular and tubulointerstitial miR-638, miR-198 and miR-146a expression in lupus nephritis. *Nephrology (Carlton)* 2012, 17:346–351.
66. Panganiban RP, Pinkerton MH, Maru SY, Jefferson SJ, Roff AN, Ishmael FT. Differential microRNA expression in asthma and the role of miR-1248 in regulation of IL-5. *Am J Clin Exp Immunol* 2012, 1:154–165.
67. Lee NS, Kim JS, Cho WJ, Lee MR, Steiner R, Gompers A, Ling D, Zhang J, Strom P, Behlke M, et al. miR-302b maintains "stemness" of human embryonal carcinoma cells by post-transcriptional regulation of cyclin D2 expression. *Biochem Biophys Res Commun* 2008, 377:434–440.
68. Zhang Y, Hu H, Song L, Cai L, Wei R, Jin W. Epirubicin-mediated expression of miR-302b is involved in osteosarcoma apoptosis and cell cycle regulation. *Toxicol Lett* 2013, 222:1–9.
69. Wang L, Yao J, Zhang X, Guo B, Le X, Cubberly M, Li Z, Nan K, Song T, Huang C. miRNA-302b suppresses human hepatocellular carcinoma by targeting AKT2. *Mol Cancer Res* 2014, 12:190–202.
70. Yu J, Qiu X, Shen X, Shi W, Wu X, Gu G, Zhu B, Ju S. miR-202 expression concentration and its clinical significance in the serum of multiple myeloma patients. *Ann Clin Biochem* 2013. doi: 10.1177/0004563213501155.
71. Zhao Y, Li C, Wang M, Su L, Qu Y, Li J, Yu B, Yan M, Yu Y, Liu B, et al. Decrease of miR-202-3p expression, a novel tumor suppressor, in gastric cancer. *PLoS One* 2013, 8:e69756.
72. Keniry A, Oxley D, Monnier P, Kyba M, Dandolo L, Smits G, Reik W. The H19 lincRNA is a developmental reservoir of miR-675 that suppresses growth and Igf1r. *Nat Cell Biol* 2012, 14:659–665.
73. Gao WL, Liu M, Yang Y, Yang H, Liao Q, Bai Y, Li YX, Li D, Peng C, Wang YL. The imprinted H19 gene regulates human placental trophoblast cell proliferation via encoding miR-675 that targets Nodal Modulator 1 (NOMO1). *RNA Biol* 2012, 9:1002–1010.
74. Kim NH, Choi SH, Kim CH, Lee CH, Lee TR, Lee AY. Reduced MiR-675 in exosome in H19 RNA-related melanogenesis via MITF as a direct target. *J Invest Dermatol* 2014, 134:1075–1082.
75. Zhuang M, Gao W, Xu J, Wang P, Shu Y. The long non-coding RNA H19-derived miR-675 modulates human gastric cancer cell proliferation by targeting tumor suppressor RUNX1. *Biochem Biophys Res Commun* 2014. doi: 10.1016/j.bbrc.2013.12.126.
76. Padi SK, Zhang Q, Rustum YM, Morrison C, Guo B. MicroRNA-627 mediates the epigenetic mechanisms of vitamin D to suppress proliferation of human colorectal cancer cells and growth of xenograft tumors in mice. *Gastroenterology* 2013, 145:437–446.
77. Zhou M, Liu W, Ma S, Cao H, Peng X, Guo L, Zhou X, Zheng L, Wan M, Shi W, et al. A novel onco-miR-365 induces cutaneous squamous cell carcinoma. *Carcinogenesis* 2013, 34:1653–1659.

Acknowledgments

Foremost, I would like to express my sincere gratitude to my supervisor, Prof. Franco Pagani for giving me the opportunity to do my PhD in his laboratory, at the International Centre for Genetic Engineering and Biotechnology in Trieste. I am very grateful for his encouragement, motivation, support and patience throughout these three years.

I also want to thank my tutor, Dr. Federica Benvenuti, the coordinator of the PhD program in Biochemistry, Molecular Biology and Biotechnology at the University of Ferrara, Prof. Francesco Bernardi and the members of my PhD committee for their general comments and suggestions.

A special thanks to my collaborators, Dr. Danilo Licastro (CBM, Trieste), Dr. Daniele Castiglia and Dr. Paola Fortugno (IDI, Rome), Dr. Dimitar Efremov and Dr. Stefania Gobessi (ICGEB, Monterotondo, Rome) for their helpful discussions and useful suggestions.

I want to thank present and past members of the Human Molecular Genetics group. I am especially grateful to the "original" HMG group: Chiara, Erica, Gosia, Andrea, Irving and Ify for their support always provided with enthusiasm and for the friendly atmosphere in the lab. I would also like to express a special thanks to Chiara and Erica for good time at work and outside. Their expertise and constant support, always provided with kindness and patience, were fundamental during my PhD. I feel incredibly lucky to have worked with both of them.

Finally, my deepest gratitude goes to my family and especially my mum, my dad and my brother Michele for their unconditional love and support. Thanks to all my friends for their support and friendship. To each of them I want to say: never ever stop dreaming, because "if you can dream it, you can do it".