UNIVERSITY OF TRENTO

Italy

International PhD Program in Biomolecular Sciences

Centre for Integrative Biology

XXIX Cycle

"Characterization of the hnRNP RALY in RNA transcription and metabolism"

Tutor:

Prof. Paolo Macchi

Cibio - University of Trento

Candidate:

Nicola Cornella

Cibio - University of Trento

Advisor:

Dr. Annalisa Rossi

Cibio - University of Trento

Academic Year 2016-2017

Index

ABSTRACT

The heterogeneous nuclear ribonucleoproteins (hnRNPs) form a large family of RNA-binding proteins (RBPs) that exert numerous functions in RNA metabolism. For example, soluble hnRNPs bind to RNAs to mediate their maturation, processing, and shuttling from the nuclear compartment to the cytoplasm. Additionally, hnRNPs might interact with chromatin to regulate the transcription and the post-transcriptional modification of nascent transcripts.

RALY is a member of the hnRNP family that binds poly-U rich elements within several RNAs and regulates the expression of specific transcripts. RALY is upregulated in different types of cancer and its downregulation has been shown to impair cell proliferation. In my PhD project, I characterized RALY to interact with transcriptionally active chromatin in a transcription-dependent manner and to cause a global decrease of RNA Polymerase II (RNAPII)-mediated transcription when downregulated, without affecting RNAPII elongation rate. Through microarray analysis of RALY-downregulated HeLa cells, I detected an altered expression of numerous genes involved in transcription promotion and cell cycle regulation, including the E2F transcription factors family. Due to its relevant role in regulating the cell cycle, I focused on the proliferation-promoting factor E2F1. I demonstrated that the stability of *E2F1* mRNA is reduced in cells lacking RALY expression, with a resulting reduction of E2F1 protein levels. As a consequence of RALY knock-out, HeLa cells present a slower cell proliferation compared to control cells. Finally, by crossing the list of RALY targets with the list of genes affected by RALY downregulation, I propose a positive role of RALY in regulating the fate of specific transcripts.

Taken together, my results highlight the importance of RALY expression for transcription and cell proliferation.

1. Introduction

The phenomenon of heredity and the capacity of organisms to reproduce themselves are central to the definition of life: it makes living beings distinguishable from other natural systems. Every cell possesses a genome, the "hard drive" harboring the information that will define the biological and phenotypical traits of the organism and that will be passed to the offspring to allow the pursuance of the species. The information is stored in the form of DNA, but to be allowed to free all of its potential, it has to be brought into the form of proteins and other molecules that will allow the biological machine to work correctly. In between the path from DNA to proteins, the genetic information passes through the intermediary form of RNA, that will be used as template to direct the synthesis of proteins. Proteins will ultimately administrate the information held by DNA and RNA, and the balance between different classes of proteins will determine the behavior of the entire cell, possibly leading it to adapt to different environments, to proliferate and even to die to preserve the fitness of the entire organism. The information flow from DNA to proteins holds the name of gene expression. Most of the genes of higher eukaryotes are composed by coding- and non-coding- regions, namely exons and introns, respectively. Only exons contain the information to give birth to the final protein, while introns will be lost during the processing of RNA.

Prior to the genome sequencing era, the idea that the complexity of organisms was directly proportional to the number of genes was fairly accepted. Nowadays we know that a high number of genes does not always correlate with complexity. In fact, it is the regulation of gene expression in space and time to be determinant for the formation of complex organisms. In addition, to make the system even more articulated, numerous genes will not end up in a protein form, giving rise instead to regulatory non-coding RNAs, and single genes can give rise to many different proteins through the process of alternative splicing.

There are multiple ranks of gene expression control, which can be grouped depending on the form of the information they target. In the first place, it is the activation or de-activation of a single gene to determine the beginning of the information flow. This ON/OFF switch is mainly exerted by factors that bind to the regulatory DNA sequences,

in particular to promoters, which are located upstream the sequence of the gene. However, a consistent part of gene expression control occurs at the level of RNA, during and after its synthesis.

1.1 A journey through RNA transcription

The process through which the information is moved from DNA to RNA is named transcription. It is a very complex procedure that involves numerous proteins, identified as the transcriptional machinery, and *trans*-acting factors that co-operate with it and interact with the newly synthesized RNA to ensure its correct processing [1–4]. During a transcriptional event, a long sequence of DNA will be copied into a complementary molecule of RNA, termed pre-messenger RNA (pre-mRNA).

The leading actor of transcription is the RNA Polymerase (RNAP) holoenzyme, the protein complex responsible for the reading of the DNA sequence and for the assembly of the complementary RNA copy. Eukaryotes present three different RNA Polymerases (I-III), which are responsible for the transcription of determined sets of RNAs. In particular, the complex responsible for the synthesis of protein-coding pre-mRNAs, but also of determined non-coding RNAs, is RNA Polymerase II (RNAPII). In human, RNAPII is composed by 12 subunits (RPB1-12), of which RPB1 is the largest and most characterized. RPB1 forms the DNA binding domain of RNAPII and contains a carboxy-terminal domain (CTD), composed by up to 52 heptapeptide repeats (Y_1-S_2-P_3-T_4-S_5-P_6-S_7), that will be chemically modified during transcription to regulate the activity of the whole holoenzyme and to recall *trans*-acting factors involved in the co-transcriptional processing of nascent RNAs [2,5–13].

The transcriptional activity of RNAPII can be divided into three main phases: recruitment of the holoenzyme on the regulatory sequence of the gene, initiation of transcription and productive elongation, and termination [14]. After 20-60 nucleotides from the transcription start site, RNAPII is driven in a paused condition by two proteins: the negative elongation factor (NELF) and the DRB-sensitivity inducing factor (DSIF) [15–17]. During this pausing, the newly synthesized RNA undergoes capping, one of the first modifications on the way to become mature, and the CTD of RNAPII is post-translationally modified to ensure a correct productive elongation [17,18]. The pausing of

RNAPII at promoter-proximal regions works as a control step for gene expression. In fact, at this stage, the fate of the RNA can be positively or negatively influenced [17]. The phosphorylation of the pausing inducing factors and of the CTD itself, enables the RNAPII to resume transcription [19]. In particular, the driving modification for switching from the paused state to productive transcription is the phosphorylation of serine residues 2 (pS2) of the CTD heptapeptide [20]. This modification is performed by the Positive Transcription Elongation Factor b (P-TEFb), a complex composed by the Cyclin Dependent Kinase 9 (CDK9) and one of the Cyclins T (CCNT1, CCNT2a, CCNT2b) or Cyclin K. In humans, about 80% of the cellular P-TEFb is composed by CDK9 bound to CCNT1. P-TEFb can act either alone or as part of different complexes [21–23].

After being released from the pause, RNAPII proceeds through elongation and finally arrives to termination, when the pre-mRNA will detach from the holoenzyme catalytic pocket. As the initiation phase, also the elongation part of the RNAPII activity is tightly regulated. In fact, RNA processing takes place in a co- and post-transcriptional manner and is tightly dependent on the elongation conditions of RNAPII. The pre-mRNA macromolecule will undergo to a series of modifications, namely RNA capping, splicing, A-to-I editing, 3'-end processing, addition of chemical groups and poly-adenylation, that will ultimately transform the pre-mRNA into a messenger RNA (mRNA) molecule [8–13].

In particular, the splicing of pre-mRNAs, the process during which introns are cut out from the pre-mRNA and exons are welded together, is highly influenced by RNAPII elongation speed. Exons and introns are recognized by splicing factors thanks to specific and conserved sequences at their extremities. Depending on their sequences, the splice sites can strongly or weakly attract splicing factors, determining the inclusion or exclusion of the exon in the final mRNA. Different studies support the so-called "kinetic model", where a slower elongation promotes exon inclusion, favoring the binding of splicing factors to weak splice sites, which would have been lost in favor of afterwards synthesized strong splice sites in a condition of faster RNAPII elongation. The slower RNAPII opens a wider "window of opportunity" for splicing factors to bind weak splice sites [24–28].

By including or skipping exons during splicing, a single genetic locus can give birth to multiple mRNAs, which will code for different isoforms of the same protein. This puzzle-like process holds the name of alternative splicing, and the majority (>95%) of higher

eukaryotes genes is alternatively spliced [3]. Both constitutive and alternative splicing are performed co-transcriptionally by the symphonic activity of numerous RBPs and ribonucleoprotein particles (RNP) assembled into a huge complex named spliceosome, which has the task to remove the introns from the pre-mRNA [3,26,29,30]. In higher eukaryotes, introns can be distinguished in U2- and U12-type depending on specific consensus sequences, and they are specifically spliced by two different spliceosomes, the major or the minor spliceosome, respectively. These two complexes are compositionally different but functionally analogous. U12-type introns are a minor subgroup of introns and only account for less than 0.5% of all introns in any given genome [29,31]. Both the spliceosomes are very dynamic RNP complexes, constituted by a fundamental building block of up to five small-nuclear RNAs (snRNAs), present as small-nuclear RNPs (snRNPs), and over 300 proteins. Most of the interactions and the reactions performed by the spliceosomes on RNA are the result of an intricate puzzle of bindings, associations and tertiary structures made by quickly exchanging components of these machineries, and of the contribution of numerous *trans*-acting RBPs [3].

The termination of transcription occurs differently between the mammalian RNA-Polymerases, with the RNAPII mechanism being the most complicate. Interestingly, in mammals, transcription termination can occur from a few bases till kilobases after the poly-adenylation site [32]. Transcript termination is intimately connected with the 3'-end processing of the pre-mRNA, and to be terminated, a protein-coding pre-mRNA needs an intact poly-adenylation site [33,34]. Defects in the termination process can cause the production of aberrant RNAs and can affect the transcription of downstream genes, possibly leading to the unwanted transcription of anti-sense RNAs and determining a shortage in the free RNAPII available for new rounds of transcription. Nevertheless, defective transcripts are usually degraded by different nucleases, however eventually generating unbalances in protein levels given the lower amount of correctly terminated RNAs. Notably, given the wide involvement of RNAs in the control of various cellular processes, RNA degradation is highly efficient and incredibly fast, so not to leave free RNA fragments floating inside the cell [35].

The newly formed mRNA molecule will be used as blueprint by ribosomes to synthesize proteins in the process of translation. Messenger RNAs that are not properly processed

will be considered as non functional by the cellular machinery and successively degraded.

The maturation steps bringing the pre-mRNA to mRNA, namely RNA processing and maturation, and the regulation of mRNA stability, localization, and translational state undergo the name of post-transcriptional gene regulation (PTGR) [36]. Phenomena tuning and affecting PTGR will ultimately determine the amount of every protein inside the cells. A representation of the genetic information flow is depicted in **Fig. 1**.

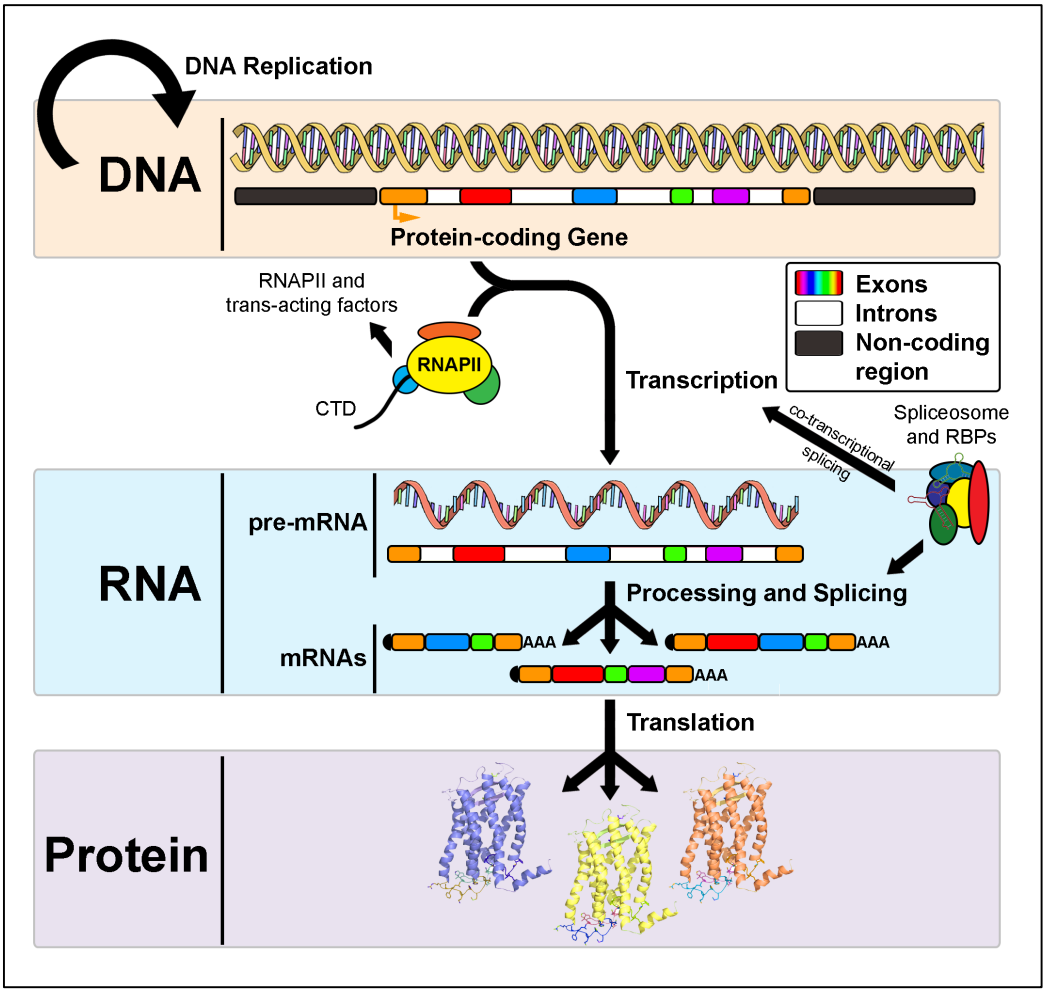

***Figure 1**. Schematic representation of the genetic information flow.*

1.2 RNA Binding Proteins

The proteins responsible for the maturation and the homeostasis of RNA are collectively known as RNA-binding proteins (RBP). In humans, a recent census identified 1542 RBPs, all of which are involved either directly or indirectly in protein synthesis and therefore

participate to PTGR [36]. Interestingly, almost all the identified RBPs had no tissue-specificity and the mRNAs coding for RBPs constituted more than the 20% of the total mRNAs of the cells, altogether highlighting the importance of RBPs in cellular physiology. Inside the cells, RBPs assemble with RNAs into dynamic RNPs that mature, process, regulate the translation and transport of the transcripts (**Fig. 2**). In addition, RBPs and RNPs operate as RNA chaperones helping the transcripts to assume their correct secondary structure, but also preventing the aggregation, misfolding and incomplete processing of RNAs. As a consequence, the abundance of RBPs differentially affects RNA regulation, determining the abundance of proteins and specific isoforms inside the cells and therefore indirectly influencing cellular processes [36–38].

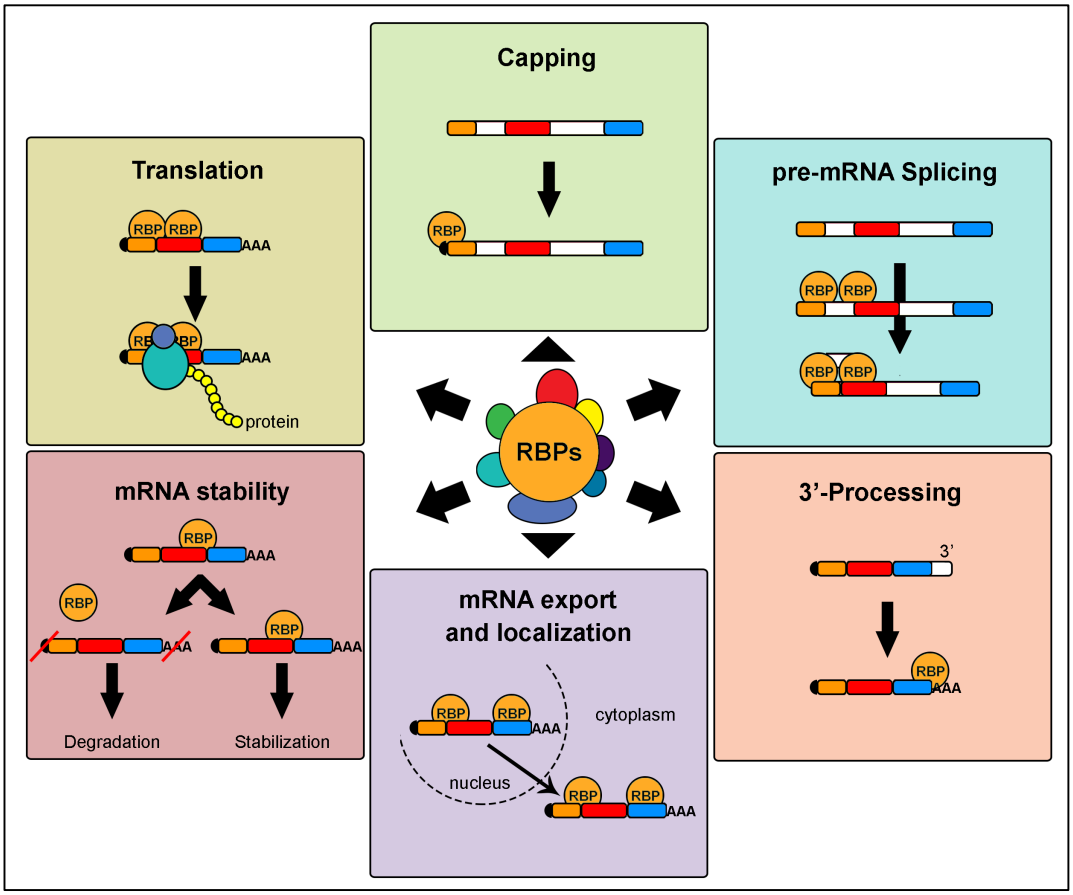

Figure 2. Summary of the main functions of RNA-binding proteins.

From the structural point of view, RBPs are constituted by a limited number of amino acidic modules, and the incredibly large substrate diversity recognition is achieved by the different arrangement of these recurrent building blocks in their RNA-binding domain (RBD) (**Fig. 3**). The modularity of the RBD is the key to both binding specificity and target versatility [39,40]. In fact, these singular elements usually weakly recognize short

stretches of nucleotides and the juxtaposition of multiple different or equal modules allows the fine tuning of both target binding strength and specificity. More modules permit to the protein to specifically recognize longer nucleotide stretches or sequences that are separated by an intervening nucleotide tract, but also target sequences localized on two different RNAs. The length and structure of the linker region between two modules also play a crucial role in the recognition capability of the RBD [39]. An RNA-binding protein will bind to its target RNA when a determined series of parameters, such as sequence, structure of the RNA and interactions with other proteins are satisfied. Depending on the RBD, an RBP can recognize single- and/or double-stranded RNA [39,41].

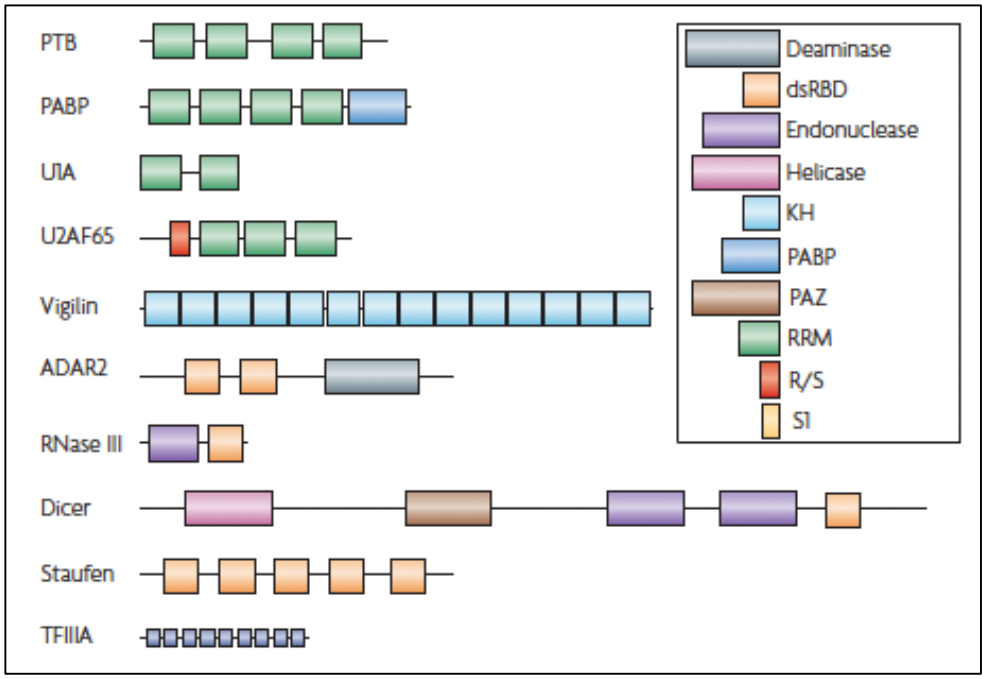

Figure 3. Representative examples of the variability in the organization of recurrent structural modules in some of the most characterized RBPs. The specific disposition of the modules allows the RBPs to perform specific functions. RRM=RNA-Recognition Motif, by far the most characterized RNA-binding protein module. KH=K-homology domain, able to bind both single-stranded RNA and DNA. dsRBD=double-stranded RNA-binding domain, a sequence-independent dsRNA-binding module. ZnF=RNA binding zinc-finger domain. Enzymatic domains and less common functional modules are also shown. PABP=poly(A)-binding protein. PAZ=Piwi Argonaut and Zwille domain. Picture taken and modified from Lunde, Moore and Varani, 2007 [39].

In addition, multiple different RBPs are simultaneously bound to every RNA, allowing RNPs to perform an even more various, specifically and sensitively tuned array of functions. A specific RNA-binding protein in fact, can have different targets and functions depending on the other RBPs interacting with it and with the target RNA. Often, RBPs create homodimers to interact with their targets, performing different

functions compared to the single RNA-binding protein. Virtually all the interactions ultimately result in the post-transcriptional regulation of gene expression [39,42].

The most common and characterized RBD is the RNA-recognition motif (RRM). Around ten thousand RRMs have been described, and, interestingly, around 0.5-1% of the human genes contain the sequence coding for an RRM, often in multiple copies [39,43,44]. RRMs are usually 80-90 amino acids long, containing two conserved sequences of eight and six amino-acids, respectively called RNP1 and RNP2, and recognize between four and eight nucleotides long sequences. Given this short sequence specificity, also in this case, multiple RRMs are often present in the same protein [39,45,46]. RRMs usually adopt a typical β1α1β2β3α2β4 topology that forms a four-stranded β-sheet packed against two α-helices. Interestingly, beyond protein-RNA interactions, RRMs have been described to mediate also protein-protein interactions, further enlarging the capabilities of RNA-binding proteins possessing an RRM [39,47].

RNA-binding proteins are grouped into different families, one of which is the *heterogeneous nuclear ribonucleoprotein* (hnRNP) family.

1.3 The hnRNP family

Historically, the term hnRNP was coined to indicate all the proteins interacting with nuclear high molecular weight RNA (the pre-mRNA, indicated also as heterogeneous nuclear RNA) synthesized by RNAPII [48]. Later on, the definition evolved, indicating all the proteins that could be cross-linked to hnRNA with UV radiation *in vivo*. Given this definition, it is no surprise to find consistent structural diversity among the members of the hnRNP family [49]. hnRNP A1, A2/B1, B2, C1 and C2 were the first to be identified given their abundance and somewhat constitutive association to hnRNAs, and were therefore called the hnRNP "core" or major hnRNPs. Minor hnRNPs are the ones not stably associated to pre-mRNAs [37,48,50]. To date, more than 30 hnRNPs have been identified and described to participate to different steps of RNA processing and metabolism, such as splicing, nucleo-cytoplasmic shuttling, RNA transport and mRNA stability maintenance, but also to other cellular processes (**Fig. 4**) [51,52]. From the structural point of view, hnRNPs are characterized by the presence of single or multiple RBD modules, which confer different sequence specificities, and of at least an auxiliary

domain determining the localization or mediating protein-protein interactions [51,53–57] .

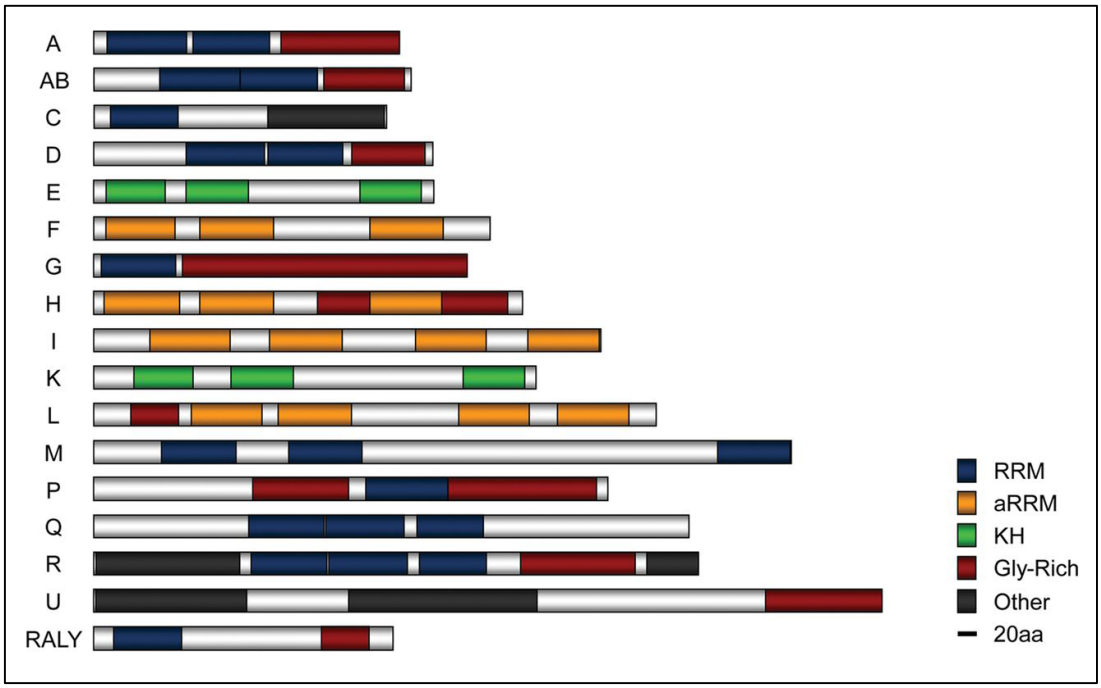

Figure 4. Schematic representation of the members of the hnRNP family. Modified from Han *et al*. [58]. RRM=RNA Recognition Motif, aRMM=atypical RNA Recognition Motif, KH=K Homology domain, Gly-Rich=Glycine Rich.

After the synthesis by RNAPII, every transcript will be associated to a specific array of hnRNP partners, as well as other RBPs, that will altogether form the messenger-ribonucleoprotein particle (mRNP). The specificity of the bound RBPs will influence the processing of the future mRNA, resulting in an "mRNP code" that will determine the fate of the mRNA. The hnRNPs associated with a single RNA exchange very quickly, making this code a very dynamic parameter regulating the phases of mRNA processing. Interestingly, core hnRNPs (such as hnRNP-A1 and -C1) are present in the nucleus at concentrations similar to fundamental nuclear components as histones, and in vast excess over their respective binding sites. This abundance allows these hnRNPs to establish also sequence non-specific interactions with transcripts, enlarging their array of functions [59,60].

1.3.1 hnRNPs in transcription and RNA processing

A function in constitutive and/or alternative splicing has been demonstrated or proposed for more than half of the major hnRNPs.

HnRNP-A1/A2 possess two RRMs at the N-terminus and a Glycine-rich region (RGG) at the C-terminus, which also contains a nuclear-cytoplasmic shuttling domain. They can both repress and/or promote splicing events ultimately determining exon skipping or inclusion. Moreover, they were described to mediate the transport and the stability of specific mRNAs, and also to regulate gene expression by influencing RNAPII elongation [49,61–65].

HnRNP-C1/C2 (hnRNP-C) are two strictly nuclear proteins presenting a N-terminal RRM, followed by a nuclear retention signal and a basic-leucine zipper-like motif (bZLM) typical of DNA-binding proteins. Interestingly, is the basic amino acidic region preceding the bZLM to provide the high affinity for RNA, while the RRM plays a minimal role in determining the binding to RNA [49,66,67]. HnRNP-C was detected to bind poly-U tracts in all nascent transcripts and to be involved in splicing and in regulating the stability of specific mRNAs. Interestingly, hnRNP-C was shown to exert a protective role in masking cryptic 3'-splice sites into repeated ALU sequences present in pre-mRNAs, preventing their recognition by the spliceosome and their successive inclusion into mature transcripts [49,68].

HnRNP-U, also known as Scaffold Attachment Factor A (SAF-A), presents a RGG domain capable of binding RNA at the C-terminus, a central SP1a and ryanodine receptor (SPRY) homology domain of unknown function, and a scaffold-associated region (SAR)-specific bipartite DNA binding domain (SAF) able to bind specific DNA sequences at the N-terminus [49,69]. HnRNP-U was described to be necessary for a correct heart development in mice, mediating the splicing of the fundamental excitation-contraction coupling protein Junctin, and to negatively regulate RNAPII elongation in human cells by inhibiting the phosphorylation of the CTD [70,71]. Moreover, hnRNP-U was shown to be required for the localization of the *Xist* RNA on the X chromosome and for the consequent formation of the inactivated-X chromosome. In this model, hnRNP-U interacts with the *Xist* RNA through the RGG domain and anchors it to the X chromosome binding DNA with its SAF domain [72].

HnRNP-K is a nucleic acid-binding protein that acts on its targets upon different cellular stimuli, integrating different signaling pathways. Historically, it was the first protein where a common RBD, the K-homology (KH) domain, was observed. This domain was later found in numerous other RNA-binding proteins. From the structural

point of view, hnRNP-K possesses three KH-domains, an RGG region and a nuclear shuttling signal. It has been observed to participate to the regulation of mRNA splicing, stability and translation. When bound to DNA, hnRNP-K can interact with transcription factors to direct gene expression or repression, and also with chromatin remodelers to influence DNA organization. Furthermore, hnRNP-K was shown to recruit the termination-exonuclease XRN2 on specific transcripts, regulating transcription termination [49,73–75].

As described, hnRNPs can have fundamental roles in the maturation of RNAs. Consequently, the misregulation of hnRNPs might result in the incorrect processing of important cellular factors, possibly leading to a loss of cellular homeostasis and to pathological conditions [76,77].

1.3.2 hnRNPs and diseases

Despite their ubiquitous expression, the deletion or mutation of single hnRNP can result in tissue-specific diseases. This specificity can be explained either by a tissue-specific expression of mRNA targets of the RBP or by a higher susceptibility of the tissue to changes in a specific PTGR pathway [36].

In general, a genetic mutation affecting the RBP-mRNA regulatory axis can hit either the RBP or the mRNA. In the first case, the mutation could occur in the DNA sequence coding for the RBD, rendering the protein unable to bind part or all of its target mRNAs, or in a position that will affect the behavior of the whole protein, for example modifying its localization or interactions with other proteins. In the second case, the mutation in the specific target mRNA could make it inaccessible to the RBP, therefore preventing the correct processing of the transcript [36]. The loss of interaction between the RBP and its target RNA might affect the splicing and processing of the latter, possibly leading to pathological conditions [77]. The misregulation of hnRNPs has been described in different pathologies, and in particular in tumors and neurodegenerative diseases (**Fig. 5**) [78].

hnRNP	Function	Link to disease
A1	Splicing mRNA stability Translational regulation	ALS/FTLD Cancer
A2/B1	Splicing mRNA stability	ALS/FTLD Alzheimer's Disease Cancer
C1/C2	Splicing Translational regulation Transcript sorting	Alzheimer's Disease Fragile X Syndrome Cancer
D (AUF1)	mRNA decay Telomere maintenance	–
E1/E2/E3/E4	Translational regulation Transcriptional regulation mRNA stability Splicing	Cancer
F	Splicing Telomere maintenance	ALS/FTLD Cancer
G	Splicing	SMA
H	Splicing	ALS/FTLD Cancer
I (PTB1)	Splicing mRNA stability Transcriptional regulation	–
K	Translational regulation Transcriptional regulation mRNA stability Splicing	ALS/FTLD Cancer
L	Splicing mRNA stability	–
M	Splicing	SMA Cancer
Q1/Q2/Q3	Splicing Translational regulation	SMA
R	Transcriptional regulation Translational regulation	SMA
U	Splicing Transcriptional regulation	–

Figure 5. List of the described functions and link to diseases of hnRNPs. Taken and modified from Geuens, Bouhy and Timmerman, 2016 [78].

HnRNP-A1 was found to be overexpressed in different tumor samples and to be associated with metastasis and tumor proliferation, directly interacting with the mRNAs of well-described oncoproteins as HRAS and KRAS [65,78–80]. Furthermore, hnRNP-A1 was observed to participate to the alternative-splicing of CD44, a transmembrane receptor described to be involved in tumorigenesis favoring epithelial-mesenchymal transition [81]. HnRNP-C was identified as a key player in the splicing of the mRNA of the oncogene *BRCA1*. In fact, loss hnRNP-C induced a decrease of BRCA1 inside the cells, together with other proteins involved in DNA repair, eventually favoring tumorigenesis [82]. Moreover, HnRNP-K was described to be deleted in several samples of hematologic malignancies, successively characterized as a haploinsufficient tumor suppressor, and to be involved in tumorigenesis affecting different steps of RNA maturation [83,84].

Neurons are the most polarized cells of the human body, demanding for a secure transport and a precise localization of mRNAs in space and time, and the nervous system is described as the tissue presenting the highest level of alternative splicing [85]. A misregulation of hnRNPs, and in general of RBPs, in neurons could therefore impair the correct balance between splicing isoforms and affect the correct localization and

translation of fundamental mRNAs, concurring to the development of neurodegenerative pathologies [86–89]. For example, hnRNP-A1 downregulation was observed to associate with the severity of various neurodegenerative disorders, such as Alzheimer's disease (AD), spinal muscular atrophy (SMA), fronto-temporal lobar degeneration (FTLD), amyotrophic lateral sclerosis (ALS) and multiple sclerosis (MS) [90,91]. Currently, five hnRNPs (A1, G, M, Q and R) are linked to SMA [78]. HnRNP-C was described to contribute to AD by increasing the stability of the Amyloid Precursor Protein, consequently increasing its expression [56,92,93].

FUS/TLS (FUS), also classified as hnRNP-P2, has been described to exert a vast array of functions in RNA metabolism and also to bind DNA [94]. Among its numerous functions, FUS is involved in splicing, interacts with RNAPII CTD regulating the phosphorylation on Serine 2, binds alternative poly-adenylation sites on pre-mRNAs to determine the final length of the transcripts, is involved in the repair of stress-induced DNA damages, binds transcriptionally active chromatin and regulates the gene transcription of the manganese superoxide dismutase [95–100]. Given the multiple roles inside the cells, FUS has been associated to multiple diseases, in particular affecting the nervous system. Mutations of FUS have been identified as causative or risk factors of ALS and FTLD. The accumulation and aggregation of FUS in the cytoplasm, and less frequently in the nucleus, is the major feature of FUS induced ALS and FTLD, but FUS inclusions have been observed in different neurodegenerative diseases. The FUS insoluble aggregates inside and outside neuronal cells induce cellular stress, ultimately leading to neuronal degeneration. Moreover, inside the cytoplasm, the inclusion sequester RNAs of which misexpression could further impair cellular activities [101,102].

HnRNPs interact with numerous RNAs, exerting a vast plethora of roles, and can modify their associations also in response to determined cellular and environmental stimuli. Given these characteristics, is no surprise that they have been associated with different pathologies. Future investigations will for sure find hnRNPs involved in other cellular processes and pathologies.

1.4 The hnRNP RALY

RALY, the RNA-binding protein Associated with Lethal Yellow mutation, is a member of the hnRNP family that was described for the first time in mice affected by the "lethal yellow mutation". In these animals, a deletion of 170 kb on chromosome 2 eliminates the coding sequence of RALY, the entire Eukaryotic Initiation Factor 2B and the 5' region of the ASIP/Agouti gene, encoding for a protein involved in regulating the pigmentation of the coat. Ultimately, this deletion leads to the generation of a chimeric gene composed by the first non-coding part of RALY and the coding sequence of ASIP/Agouti. This chimera will be under the control of RALY promoter, therefore resulting in an aberrant expression control during the embryogenesis. In heterozygosis, the deletion induces a yellow fur in the animals, while in homozygosis determines the death of the embryo [103,104].

In humans, RALY is a 306 amino acids protein containing an RRM at the N-terminus, two predicted nuclear localization signals (NLSs) in the middle, and a non canonical RGG at the C-terminus (**Fig. 6A**) [105]. This latter auxiliary domain was proposed to mediate protein-protein interactions, but since RGGs have been described to have also nucleic-acids binding properties, its function remains elusive [39,42].

RALY was classified as an hnRNP thanks to the similarity with hnRNP-C, particularly high in the sequence coding for the RRM [106,107]. The gene coding for RALY is located on chromosome 20 and produces two isoforms of the protein: a longer one, RALY, and a shorter one, named P542, which lacks 16 amino acids downstream the RRM compared to RALY [108]. RALY is mainly localized inside the nucleus, but is also present in the cytoplasm (**Fig. 6B**), where it has been observed to interact with the translational machinery [109].

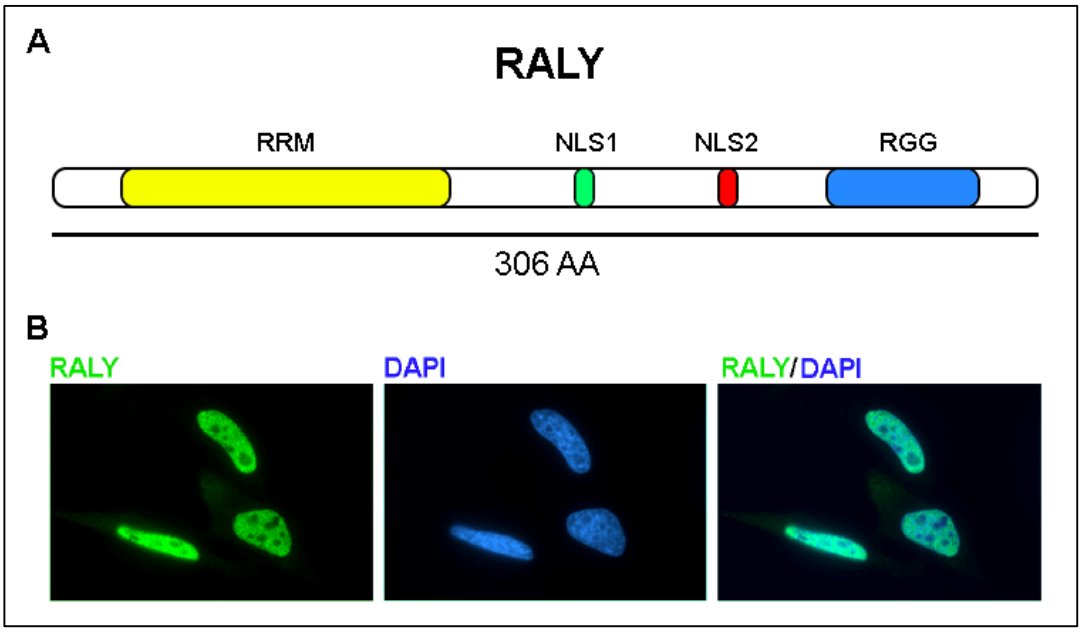

Figure 6. (**A**) Graphic representation of RALY. (**B**) HeLa cells were immunostained with antibody against RALY and nuclei were stained with DAPI. RALY is mainly localized in the nuclei, but is excluded from the nucleoli.

In human, RALY is ubiquitously expressed in all tissues, even if at different levels. Initially, RALY was identified in human as a cross-reacting antigen with the Epstein-Barr nuclear antigen 1 (EBNA1), a viral protein encoded by the Epstein-Barr virus [110]. Successively, RALY was found associated with components of the spliceosome and of the Exon Junction Complex (EJC), a protein machinery responsible to control the correct splicing of transcripts. However, its downregulation did not significantly affect the splicing of pre-mRNAs [111,112]. Thus, the role of RALY in the post-transcriptional modification of RNAs remains still elusive. Through proteomic analysis, RALY was observed to interact with the CTD of RNAPII and with proteins involved in RNA metabolism, such as the splicing factors U2AF65, PRP19, MATR3 and EIF4A3, and in translational control [105,113]. In addition, RALY was recently characterized as a transcriptional co-factor, together with the long-non coding RNA *LeXis*, together involved in the transcriptional regulation of genes belonging to the cholesterol biosynthesis pathway in mouse liver [114]. Through proteomic analysis, RALY was also observed to interact with FUS, but this association is still uncharacterized [105]. However, given the many functions exerted by FUS inside the cells, the interaction of RALY with FUS opens the interesting possibility of shared roles and cooperation in RNA metabolism.

RALY mRNA was found to be upregulated in different cancerous tissues such as ovarian, lung, bladder, brain and breast cancers as well as in multiple myelomas and melanomas. In particular, adenocarcinoma cell lines presented a high expression of RALY. In these cells, RALY was found to interact with YB-1, an RBP involved in alternative splicing, transcription and translational regulation of specific mRNAs, and in the repair of nicks and breaks into double-stranded DNA [115–118]. RALY was observed to be required for the YB-1 induced resistance to platinum-based drugs in adenocarcinoma cell lines [115]. Altogether, these observations suggest a potential involvement of RALY in tumorigenesis.

Even if RALY was described in multiple occasions, its biological and molecular roles remain elusive.

2. Aim of My PhD Project

RNA-binding proteins are involved in all the steps of RNA maturation, virtually regulating the expression of all proteins at the level of post-transcriptional gene regulation. The characterization of specific RBPs is therefore of central importance for the understanding of physiological cellular processes, but also of pathological dynamics. The RBP RALY is still largely uncharacterized, but its protein interactors and target RNAs suggest an involvement at different steps of RNA metabolism. My PhD project aimed to the expansion of the characterization of RALY, with a particular focus on its relationship with nuclear components, as chromatin, and with higher order cellular processes as transcription and cell proliferation.

The backstory of RALY

During the first year of my PhD, I mainly participated to the ongoing characterization of the interaction of RALY with RNA. In fact, RALY was classified as an RBP, given the high similarity with hnRNP-C, but its ability to effectively bind RNA and the identity of its RNA targets were not yet described. These results were published in Rossi *et al.* and the experiments I contributed to served as a hint to start my own project [109]. We demonstrated that RALY binds RNA through its RRM, preferentially in the 3' UTR of transcripts and specifically on poly-U stretches. We identified 2929 RNAs bound by RALY through RNA immunoprecipitation coupled with sequencing (RIP-seq) in MCF7 cells. Gene ontology and pathway annotation enrichment analyses revealed that proteins coded by the RNAs interacting with RALY were significantly associated with different RNA metabolism related processes such as RNA processing, transport, splicing and translation. These results immediately appeared interesting given the enrichment of RNA metabolism and splicing related proteins in the interactome of RALY, further suggesting the involvement of RALY in RNA processing [105]. Interestingly, regulation of cell cycle and apoptosis were cellular processes also significantly enriched in the proteins coded by RALY mRNA targets.

We analyzed the impact of RALY downregulation on global gene expression through microarray analysis. We found 217 misregulated genes in MCF7 cells transfected with

siRNA against RALY compared to control siRNA transfected cells. Gene Ontology enrichment analysis identified extracellular matrix organization and cell adhesion as biological processes significantly associated with upregulated genes, while cell growth and intracellular signal transduction were significantly associated with downregulated genes. Interestingly, both the RIP-seq and the microarray Gene Ontology enrichments suggested a connection between RALY and the regulation of cell proliferation. These results were supported by FACS analysis, which showed how the silencing of RALY contributed to the alteration of cell cycle progression in Panc-1 cells, with a partial arrest in G0-G1 phase. Finally, we also showed a connection of RALY with translation, identifying RALY on polysomes [109]. The results of this work are presented in details in the Publications section.

While we were characterizing RALY, I was interested in understanding the molecular function exerted by RALY on RNA. As described in the introduction, so far, RALY was found to be part of the EJC and to interact with factors exerting prominent roles in splicing and in RNA processing, but the meaning of these associations was not understood [105,111,112]. Different RBPs, including other hnRNPs such as hnRNP-C, hnRNP-U and hnRNP-K, were observed to bind chromatin to use it as a docking platform to exert their role in RNA processing [51,66,70,75,119]. Interestingly, these hnRNPs, together with other RBPs described to associate with chromatin, like FUS and Matrin 3, were included in the interactome of RALY [100,105,120]. These connections between the interactors of RALY and chromatin prompted me to investigate the possible association between RALY and chromatin. If subsisting, this association could have helped to understand the role of RALY in RNA metabolism. Starting from this point, I successively studied the effects of RALY downregulation on transcription and cell proliferation.

In parallel, I tried to better understand the involvement of RALY in splicing, studying the topic with multiple approaches. The results regarding this part are presented after the main project of the thesis, in the "RALY and Splicing" section.
I have been also involved in MaDEleNA, a very ambitious and challenging project aimed to develop an innovative device, based on the physics of "memristors", able to interface

and sense the activity of neurons. More details regarding this project can be found in the MaDEleNA section.

3. RESULTS

3.1 RALY interacts with chromatin in both an RNA-dependent and -independent manner

As a preliminary observation of the possible interaction of RALY with chromatin components, HeLa cells were fractionated into soluble and chromatin-bound protein fractions, and the presence of RALY was evaluated through Western blot. In parallel, the extracts were also treated with RNase A before fractionation to assess a possible involvement of RNA in determining the interaction, given the RNA-binding activity of RALY. Histone H3 (H3) and Actinin were used as chromatin-bound and soluble fraction markers, respectively. RALY was found to interact with chromatin, and, interestingly, the association showed a partial RNA-dependency (**Fig. 1A**). In fact, upon RNase A treatment, the signal of RALY decreased, but not disappeared, in the chromatin-bound fraction and in parallel increased in the soluble fraction (**Fig. 1A**).

To better characterize the RALY-chromatin interaction, the cells were fractionated into four different fractions: cytosolic fraction, nuclear-soluble fraction, low-salt soluble chromatin and high-salt soluble chromatin fractions [100,121]. The cytosolic fraction (C) contains all the proteins outside of the nucleus. The nuclear-soluble fraction (NS) comprises the proteins that localize inside the nucleus but are not attached to chromatin. The low-salt soluble chromatin fraction (LS) contains proteins associated to transcriptionally active chromatin and to transcription start sites (TSSs). The high-salt soluble chromatin fraction (HS) contains proteins associated to transcriptionally inactive chromatin and bulky complexes strongly associated to chromatin [121]. The enrichment of RALY in the different fractions was analyzed by Western blot. To verify the successful of the fractionation, beta-Tubulin (Tubulin), the linker histone H1X (H1X), and histone H3 (H3), were respectively used as cytosolic, inactive chromatin and whole chromatin markers. The RNA-binding protein FUS was used as a fractionation control, given its already described localization, and as a comparison for RALY with another hnRNP [94,100]. As expected, the linker histone H1X was found in the HS but not in the LS fraction, while histone H3 was present in both the LS and HS fractions. By contrast, Tubulin was

enriched in the cytosolic fraction (**Fig. 1B**). FUS confirmed the correctness of the protocol behaving as described by Yang *et al.,* showing a strong enrichment in the LS, a weaker presence in the NS and HS, and being absent from the C (**Fig. 1B**). RALY was observed to be faintly present in the cytoplasmic and in the nuclear-soluble fractions, and strongly enriched in the fractions containing both transcriptionally active and inactive chromatin bound proteins (**Fig. 1B**). Interestingly, RALY and FUS shared the localization in the transcriptionally active fraction of chromatin. To better characterize the involvement of RNA in the RALY-chromatin interaction, the samples were treated with RNase A prior to fractionation. Upon RNase A treatment, a statistically significant amount of RALY shifted from the LS fraction to the cytosolic fraction (**Fig. 1C, C and LS fractions**). On the contrary, the pool of RALY associated with transcriptionally inactive chromatin was not significantly affected by RNA degradation (**Fig. 1C, HS fraction**).

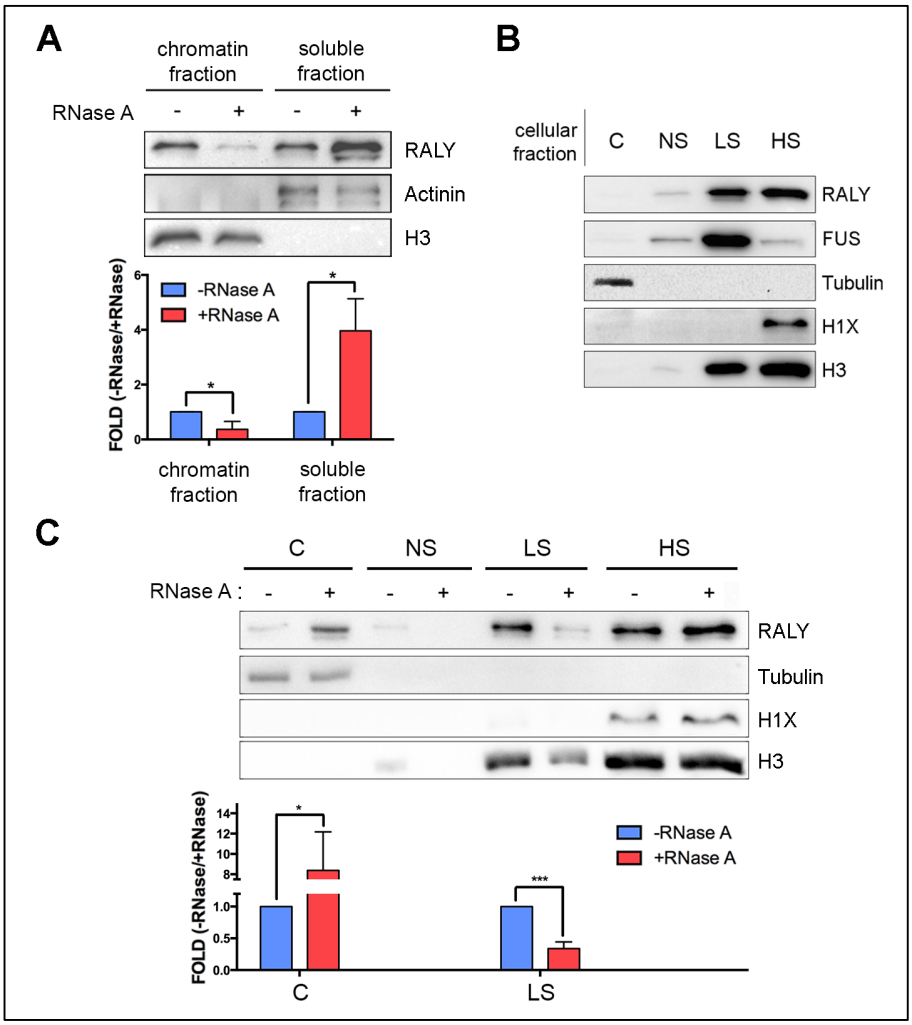

Figure 1. **RALY is associated with chromatin.** (**A**) HeLa cells were fractionated into chromatin-bound and soluble protein fractions. Each fraction was subjected to SDS/PAGE and Western blot with the indicated antibodies. Endogenous RALY is present in both the chromatin-bound and soluble fractions of proteins. (**B**) HeLa cells were fractionated into cytosol (C), nuclear soluble (NS), low-salt soluble chromatin (LS) and high-salt soluble chromatin (HS) fractions. Each fraction was subjected to SDS/PAGE and Western blot with the indicated antibodies. Endogenous RALY is present in both the LS and HS fractions. (**C**) The distribution of endogenous RALY was examined by fractionation, as in **B**, in presence (-RNase) or absence of RNA (+RNase). In absence of RNA, RALY decreased in the LS fraction with a consequent increase in the C fraction. In **A** and **C**, the levels of RALY were quantified by band densitometry analysis. The graphs show the mean values of three independent experiments and compare the level of RALY in presence and absence of RNA in the chromatin and soluble fractions (**A**), and in the C and LS fractions (**C**). Bars represent mean ± S.D. P-value was calculated using unpaired two-tailed t-test. *P< 0.05; ***P<0.001.

Taken together, these results show that RALY binds to both transcriptionally active and inactive chromatin and that part of the interaction with transcriptionally active chromatin is dependent on the presence of RNA.

3.2 The interaction with chromatin is mediated by different domains of RALY

To identify the region of RALY involved in the interaction with chromatin, we designed constructs coding for different domains of RALY and tagged them with a C-terminal c-Myc tag. The constructs coded for the full length protein (FL), the N-terminal region (amino acids 1-225), the C-terminal region (amino acids 143-306) and the unstructured Glycine-rich region (amino acids 225-306) of RALY (**Fig. 2A**). An empty expression vector (EV) was used as negative control. The expression level and the subcellular localization of the recombinant proteins were examined by Western blot and immunofluorescence analyses, respectively (**Fig. 2B-C**). All the fragments were expressed by the cells and all of them were mainly localized into the nucleus. Notably, the absence of the predicted NLSs did not preclude the construct 226-306 to accumulate inside the nucleus (**Fig. 2C**).

The distribution of the different domains of RALY was then analyzed upon cell fractionation, both in presence and absence of RNA (-RNase and +RNase, respectively). Although all the constructs were enriched in the nuclear fractions, only some of them changed their localization upon RNase A treatment (**Fig. 2D**). In particular, in absence of RNA, the RMM-containing N-terminal part of RALY (1-225) disappeared from the nuclear-soluble and low-salt soluble chromatin fractions (**Fig. 2D, NS and LS**). This was not observed instead for the C-terminal part of the protein (143-306) (**Fig. 2D, NS and LS**). These results suggest that RALY can interact with nuclear soluble components and transcriptionally active chromatin using either the N- or C-terminal region, and that the RRM-mediated interaction is RNA-dependent. To strengthen the hypothesis of two different types of binding, the FL RALY only partially decreased in both the NS and LS fractions upon RNase A treatment. This suggests that the FL RALY pool was involved in either a N-terminus- or C-terminus-mediated interaction with the proteins of the NS and LS fractions, and that the RNase A treatment only affected the pool of RALY binding chromatin through the N-terminal RRM (**Fig. 2D, NS and LS**).

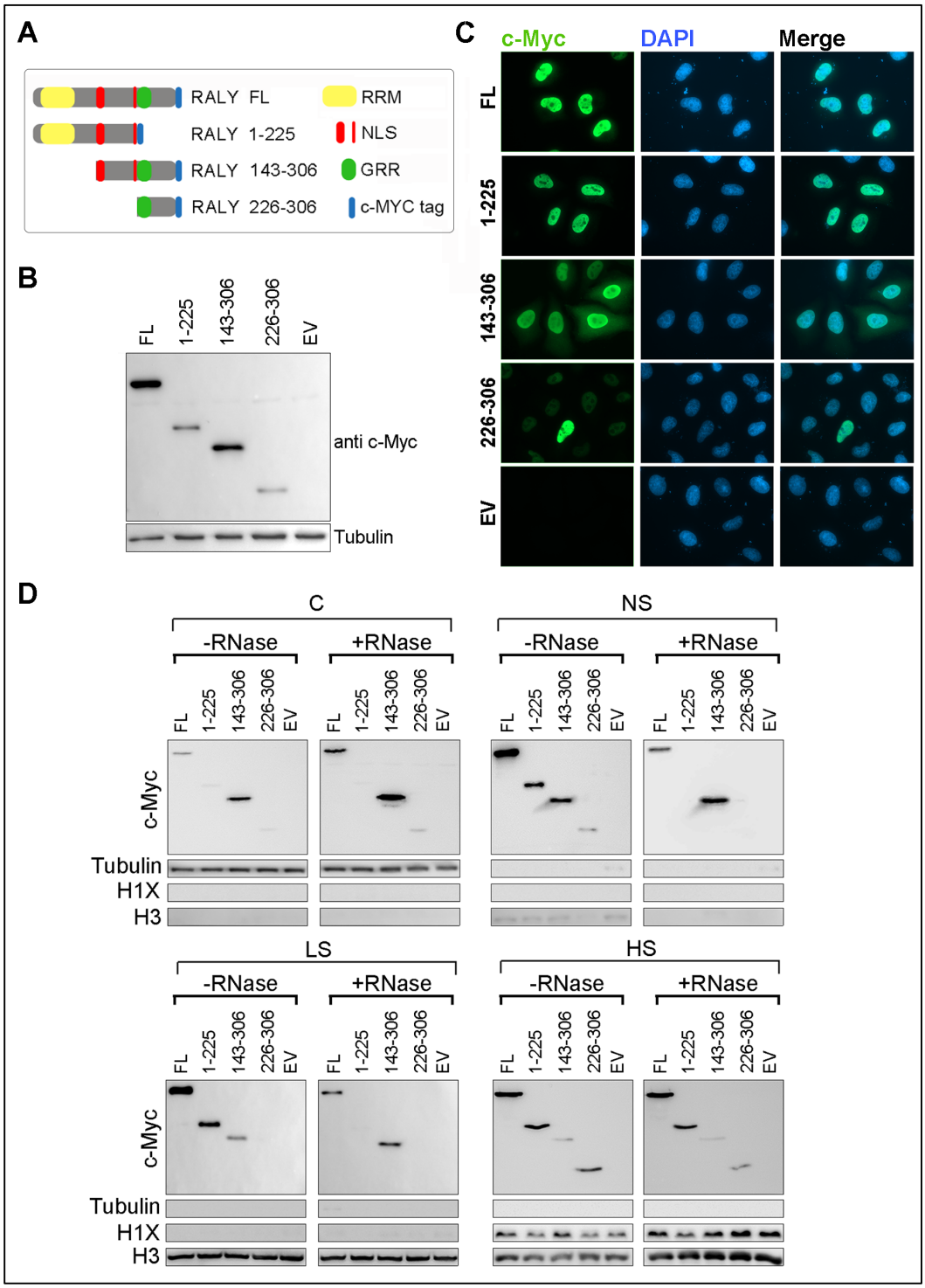

Figure 2. RALY interacts with nuclear components using either the N- or C-terminal region. (**A**) Schematic representation of RALY-c-Myc tagged constructs. (**B-D**) RALY-c-Myc tagged constructs were transfected in HeLa cells for 24 hours. (**B**) Total lysates were subjected to SDS/PAGE and Western blot with the indicated antibodies. All the constructs are expressed by HeLa cells. (**C**) The cellular localization of RALY-c-Myc tagged mutants was verified by immunofluorescence analysis with anti-c-Myc antibody and DAPI. All RALY fragments localize inside the nucleus, with the construct 143-306 being appreciably detected also in the cytoplasm. (**D**) The cells were fractionated, in presence or absence of RNA, into cytosolic (C), nuclear-soluble (NS), low-salt soluble chromatin (LS) and high-salt soluble chromatin (HS) fractions. Each fraction was subjected to SDS/PAGE and Western blot with the

indicated antibodies. The N- and C-terminal regions are both involved in the interaction of RALY with transcriptionally active chromatin (LS). In particular, the association of the RRM-containing mutant of RALY (1-225) with the LS fraction is dependent on RNA.

Regarding the association with transcriptionally inactive chromatin, all the domains of RALY localized in the HS fraction, but none of them was affected by RNase A treatment (**Fig. 2D, HS**). Further experiments are needed to better characterize the enrichment of RALY in this fraction.

These experiments highlighted the importance of RNA in the association of RALY with nuclear components and prompted me to focus on the interaction of RALY with transcriptionally active chromatin. In fact, different RBPs previously identified in RALY interactome are tightly associated with active transcription and exert their role in RNA processing in a co-transcriptional manner, interacting with the newly synthesized transcript, chromatin and also the transcriptional machinery [3,4,69,71,75,98,105,113,119,122]. To better understand the link of RALY with active transcription, I studied their relationship in the dynamic and physiological environment of living cells.

3.3 The localization of RALY is dependent on active transcription.

To study of the association of RALY with active chromatin *in vivo*, I analyzed the behavior of RALY upon the block of transcription. In fact, *in vivo*, the majority of the RNA present on chromatin is there because of active transcription. Theoretically, an RBP involved in active transcription might be affected by the absence of transcription inside the nucleus. RNA synthesis was blocked by treating the cells with the transcription blocker Actinomycin D (ActD, 5 μg/ml) for different periods of time and the nuclear amount of RALY was measured by immunostaining using the High Content Imaging System "Operetta" (PerkinElmer) [123]. The inhibition of transcription was verified by measuring the newly synthesized RNA through the incorporation of the Uracil-analogue 5-Ethynyl Uridine (5EU) into new transcripts, followed by staining with the fluorescent dye 5FAM via Click reaction [124]. A scheme of the procedure and a representative staining experiment are depicted in **Fig. 3A**.

The nuclear signal of RALY was observed to decrease over time during the ActD treatment (**Fig. 3B, left**), suggesting that the block of transcription activated either RALY

degradation into the nucleus and *de novo* synthesis into the cytoplasm, or RALY translocation from the nucleus to the cytoplasm. No signal of new RNA was detected within the nuclei upon ActD treatment (**Fig. 3B, right**).

The analysis of the samples with a more sensitive microscope showed an increase of RALY into the cytoplasm upon ActD treatment (**Fig. 3C**), but the intensity was too dim to allow for a reliable quantitative measurement. Again, the absence of 5FAM signal confirmed the effectiveness of the ActD treatment in blocking transcription (**Fig. 3C**).

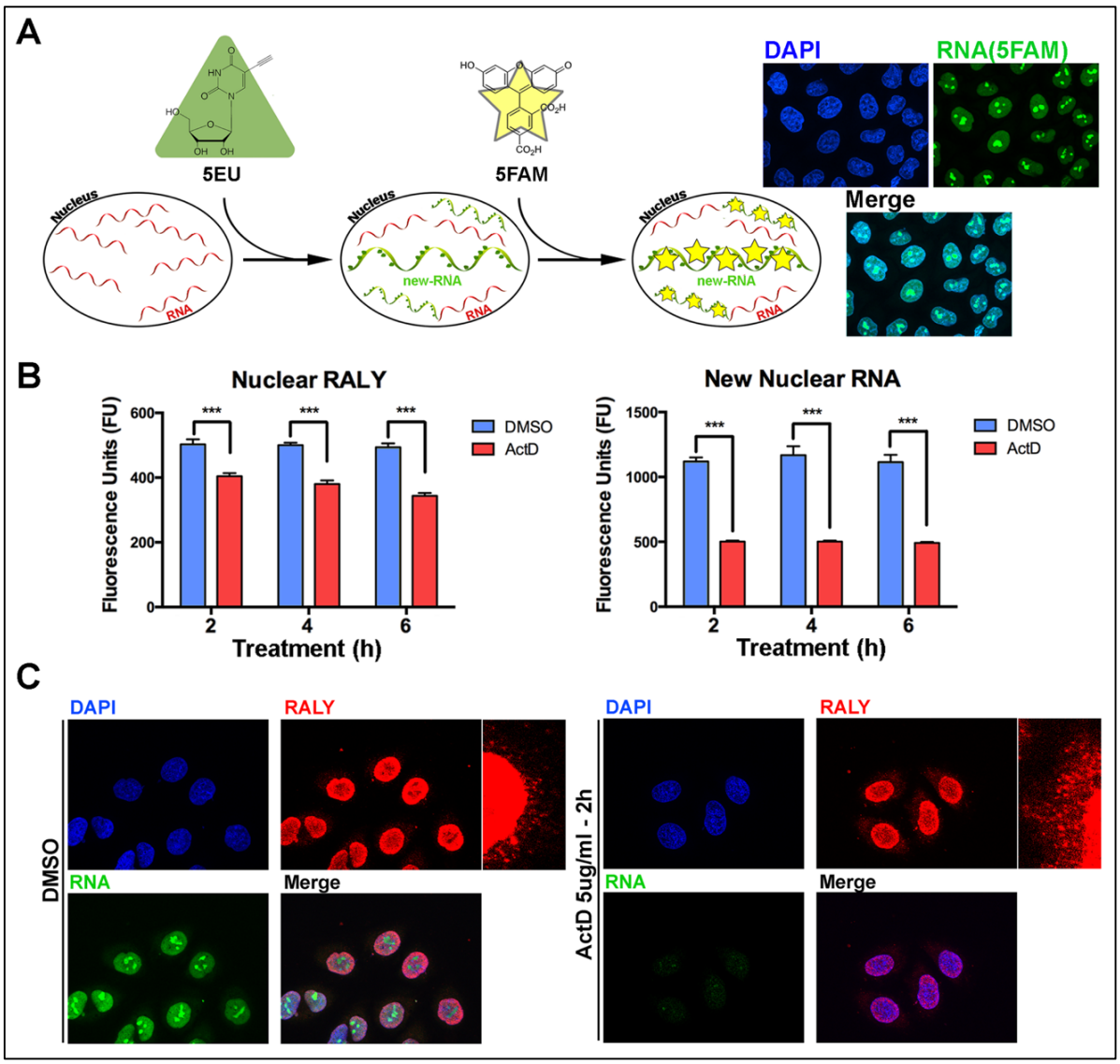

Figure 3. The localization of RALY is dependent on active transcription. (**A**) On the left, scheme of the staining procedure of newly synthesized RNA through Click reaction. On the right, representative experiment of newly synthesized RNA visualization through 5EU incorporation and successive 5FAM staining, together with DAPI to highlight nuclei. (**B**) HeLa cells were incubated with 5EU and treated with ActD (5 μg/ml) or vehicle (DMSO) for 2, 4, 6 hours and the amount of nuclear RALY was assessed by immunofluorescence with an anti-RALY antibody by the High Content Imaging System "Operetta". 5EU was stained with 5FAM through Click reaction. The amount of nuclear RALY significantly

decreases over time during the ActD treatment (left). The synthesis of new transcripts is blocked during ActD treatment (right). (**C**) HeLa cells were incubated with 5EU and treated with ActD or DMSO for 2 hours. New RNA species were stained through Click reaction with 5FAM while RALY was visualized by immunofluorescence with an anti-RALY antibody. Nuclei were highlighted with DAPI staining. After ActD treatment, there is no synthesis of new RNAs and RALY increases in the cytoplasm. The graphs in **B** represent the mean of the fluorescent signals detected over three independent experiments. Bars represent mean ± S.D. P-value was calculated using unpaired two-tailed t-test (***P<0.001).

To discriminate between the degradation and the nucleus-cytosol shift hypotheses, and to understand if a restricted population of RALY was affected by the block of transcription, the localization of RALY was studied by fractionation of cells treated with ActD, together with or without the proteasome inhibitor MG132. Treating with ActD, the level of RALY in the low-salt soluble chromatin fraction significantly decreased over time with a parallel progressive increase in the cytosolic fraction (**Fig. 4A**). The same result was observed when the cells were incubated with both ActD and MG132, while no changes were observed upon treatment with MG132 alone. These results prove that, upon transcription block, RALY shifts from the nucleus to the cytoplasm instead of being degraded by the proteasome and *de novo* synthesized to remain outside of the nucleus (**Fig. 4B**). The accumulation of Ubiquitin in the total protein extract was used as a control for MG132 activity. As expected, the MG132 treated samples showed higher levels of Ubiquitin compared to the untreated sample (**Fig. 4C**).

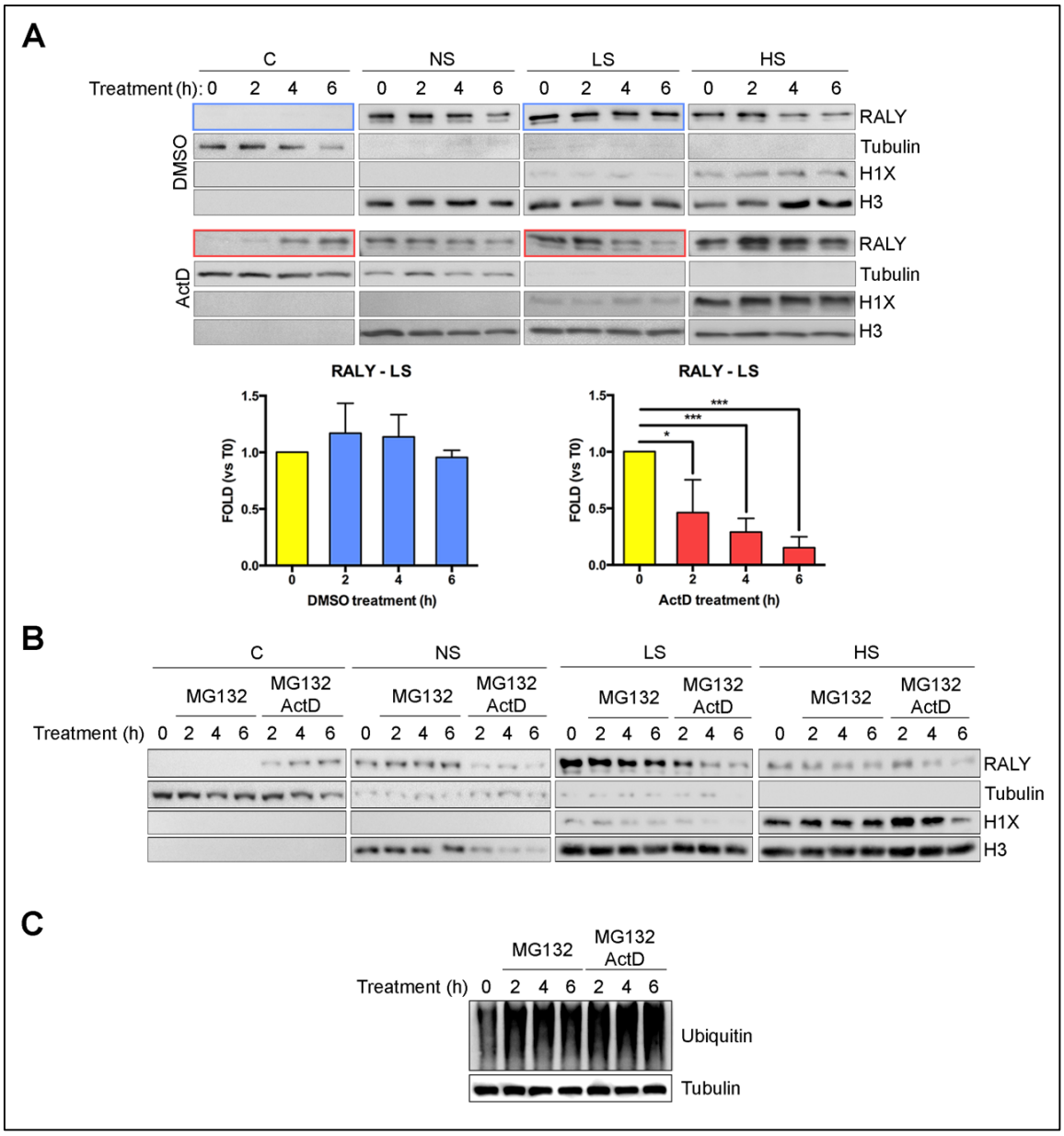

Figure 4. In absence of transcription, RALY migrates from transcriptionally active chromatin to the cytoplasm. (**A**) HeLa cells were treated with ActD (5 μg/ml) or vehicle (DMSO) for 2, 4, 6 hours and afterwards fractionated into cytosol (C), nuclear-soluble (NS), low-salt soluble chromatin (LS) and high-salt soluble chromatin (HS) fractions. Each fraction was subjected to SDS/PAGE and Western blot with the indicated antibodies. Endogenous RALY decreases in the LS fraction upon ActD treatment and progressively increases in the C fraction. (**B**) HeLa cells were treated with ActD (5 μg/ml) and/or MG132 (5 μM) for 2, 4 or 6 hours and afterwards fractionated as in **A**. Each fraction was subjected to SDS/PAGE and Western blot with the indicated antibodies. Endogenous RALY decreases from the LS fraction only in presence of ActD, while is not affected by MG132 alone. (**C**) Total cell lysates of HeLa cells treated with ActD (5 μg/ml) and/or MG132 5 μM were subjected to SDS/PAGE and Western blot with the Ubiquitin antibody. The treatment with MG132, blocking the proteasomal activity, induces an accumulation of Ubiquitin inside the cells compared to the untreated condition. In **A**, the levels of RALY in the LS fractions were quantified by band densitometry analysis. The graphs show the mean values of three independent experiments and compare the level

of RALY during the different treatments normalizing on the untreated condition (T0). Bars represent mean ± S.D. P-value was calculated using unpaired two-tailed t-test (*P< 0.05; ***P<0.001).

These results demonstrate that RALY localizes on transcriptionally active chromatin in a transcription-dependent manner and that without ongoing transcription, RALY migrates towards the cytoplasm. These observations further connect RALY to active transcription.

3.4 The downregulation of RALY affects transcription

Given the strong effect that the block of transcription had on RALY localization, I wondered whether the absence of RALY could somehow affect global transcription. To this aim, I developed a procedure to measure the synthesis of new RNA based on the incorporation of 5EU for different intervals of time and successive staining with 5FAM (**Fig. 5A**). The technique was tested and optimized with preliminary experiments on HeLa cells treated with either DMSO, as a control, or Trichostatin A (TSA), a drug described to enhance transcription by inhibiting the de-acetylation of chromatin [24,125]. In both the conditions, the new RNA species were observed to accumulate inside the nuclei along the incubations with 5EU, with a faster increase in TSA treated cells (**Fig. 5B**).

Therefore, I proceeded with the measurement of RNA synthesis in RALY downregulated HeLa cells. RALY was downregulated using siRNAs against RALY (si-RALY), while control cells were treated with non-targeting siRNAs (si-CTRL). The silencing of RALY was verified by immunostaining (**Fig. 5C, left panel**) and newly synthesized RNA was quantified only in the nuclei where RALY fluorescence signal was below a determined threshold, so to consider only successfully RALY downregulated cells. Interestingly, the levels of newly synthesized RNA showed a slower increase over time in RALY downregulated cells compared to controls (**Fig. 5C, right panel**).

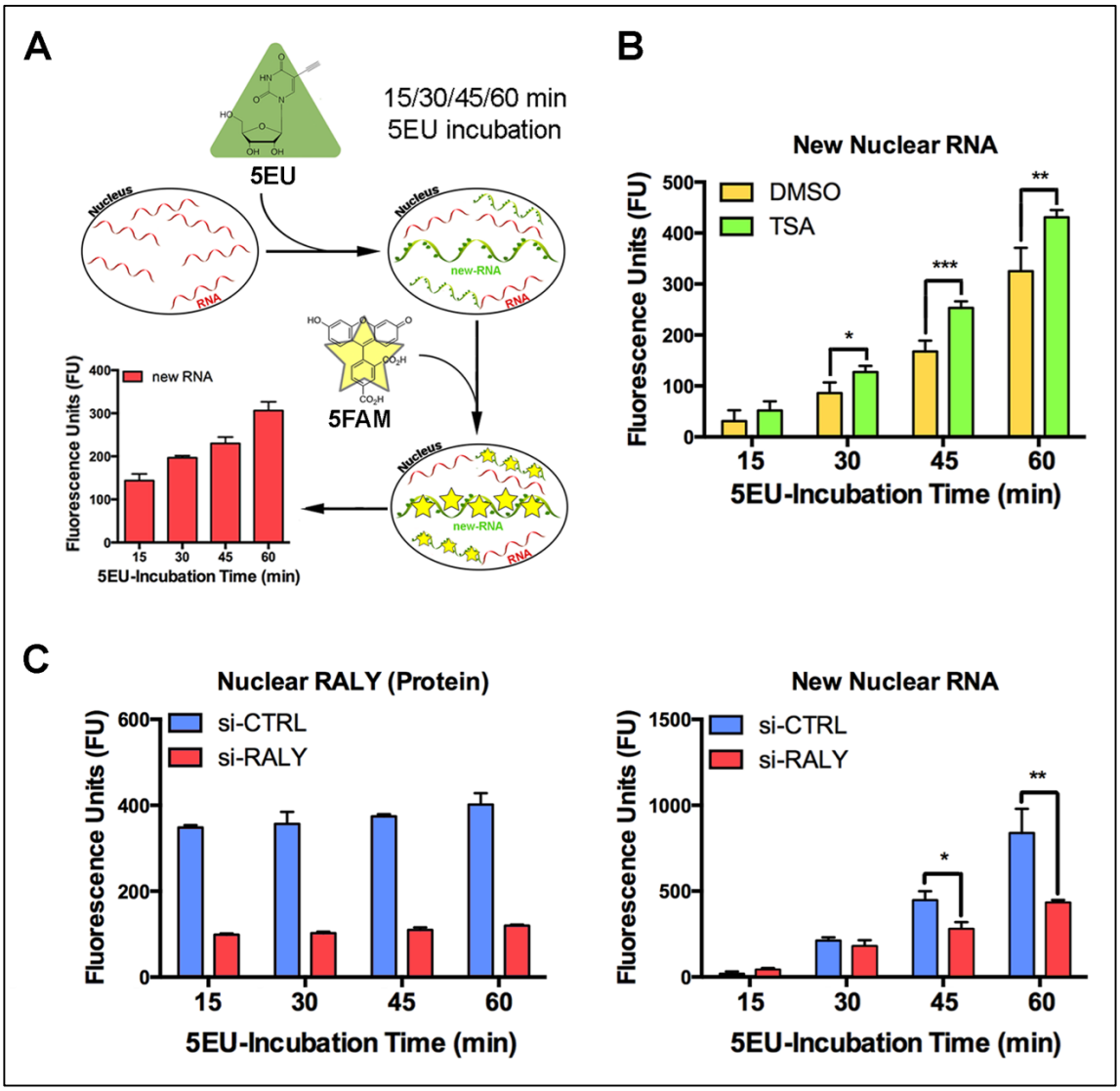

Figure 5. The downregulation of RALY determines a slower increase of new RNA species inside nuclei. (**A**) Representative scheme of the procedure to measure the progressive synthesis of new nuclear RNA species. (**B**) HeLa cells were incubated with 5EU and were then treated with either Trichostatin A (TSA) or vector (DMSO) for 15, 30, 45 and 60 minutes. 5EU was successively stained with 5FAM to visualize the transcripts synthesized over every time frame. The amount of new RNA increases over time in both the conditions, but the increase is faster in the TSA-treated cells, coherently with the drug biological effects. (**C**) HeLa cells were transfected for 72 hours either with si-CTRL or si-RALY and afterwards incubated for 15, 30, 45 or 60 minutes with 5EU to monitor the synthesis of new RNA species. 5EU was successively stained with 5FAM, while RALY was stained with an anti-RALY antibody. The images were acquired with the High Content Imaging System Operetta (Perkin Elmer) and analyzed with the software Harmony (Perkin Elmer). On the left, detection of the average signal of RALY in si-CTRL and si-RALY transfected cells. On the right, the increasing amount over time of newly synthesized RNA. In si-RALY cells, newly synthesized RNA species accumulate more slowly inside the nucleus compared to si-CTRL cells. **B** and **C** show the mean of the signal of four distinct samples processed during the same experiment. The whole experiment was repeated 3 times, giving the same result. Bars represent mean ± S.D. P-value was calculated using unpaired two-tailed t-test (*P< 0.05; **P<0.01; ***P<0.001).

The most abundant RNA species inside the cells are ribosomal RNAs, which are transcribed inside the nucleoli by RNA Polymerase I. Because of this, following 5EU incorporation into new RNAs and staining with 5FAM, nucleoli emit an intense fluorescence signal (**Fig. 3A and 3C)**. Since RALY is excluded from the nucleoli and the majority of the RBPs participating to transcription were referred to RNAPII-dependent transcription, I specifically studied the newly RNAPII-synthesized RNAs in absence of RALY. To this aim, the measurement of RNA synthesis was performed by treating the cells with a dose of ActD (125 ng/ml) sufficient to block only RNA Polymerase I transcription [124]. This treatment sensibly reduced the nucleolar staining (**Fig. 6A**). Under these conditions, the nuclear level of newly synthesized RNA in si-RALY transfected cells still increased more slowly over time compared to si-CTRL treated cells (**Fig. 6B, right**).

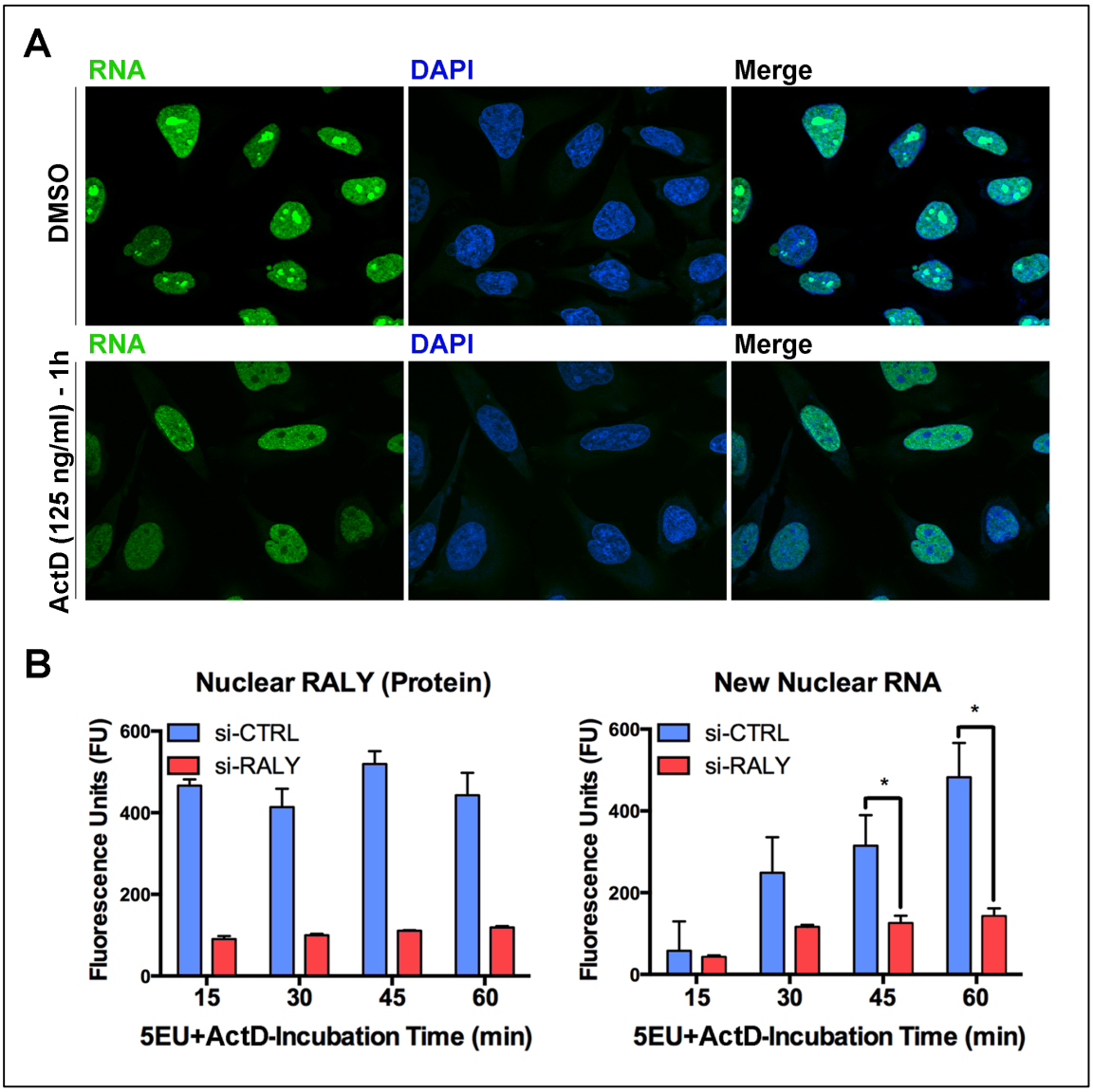

Figure 6. **The downregulation of RALY determines a slower increase over time of new RNA Polymerase II-transcribed RNAs inside nuclei.** (**A**) HeLa cells were incubated for 1 hour with 5EU and either with ActD (125 ng/ml) or vector (DMSO). 5EU was successively stained with 5FAM to visualize newly synthesized RNA species. The treatment with a mild dose of ActD only blocks RNA Polymerase I-dependent transcription, greatly diminishing the fluorescence signal of the nucleoli. (**B**) HeLa cells were transfected for 72 hours either with si-CTRL or si-RALY and afterwards incubated for 15, 30, 45 or 60 minutes with ActD (125 ng/ml) and 5EU. 5EU was successively stained with 5FAM, while RALY was stained with an anti-RALY antibody. The images were acquired with the High Content Imaging System Operetta (Perkin Elmer) and analyzed with the software Harmony (Perkin Elmer). On the left, detection of the average signal of RALY in si-CTRL and si-RALY transfected cells. On the right, the increasing amount over time of newly synthesized RNA. Also in a condition where RNAPI-dependent transcription was blocked, in si-RALY cells newly synthesized RNA species accumulate more slowly inside nuclei compared to si-CTRL cells. **B** shows the mean of the signal of four distinct samples processed during the same experiment. The whole experiment was repeated 3 times, giving the same result. Bars represent mean ± S.D. P-value was calculated using unpaired two-tailed t-test (*P< 0.05).

To confirm the involvement of RALY in the downregulation of transcription and to eliminate the possible contribution of off-target effects of siRNAs, stable RALY knockout HeLa cells were generated by CRISPR/Cas9 technology. Five RNA-guides targeting either exon 1 or exon 2 of the genomic sequence of RALY were inserted into a PX330 vector, which contained the sequence coding for the enzyme Cas9 and the resistance to Puromycin. After transfection, positively transfected cells were selected by incubation with Puromycin and single cell clones were isolated by serial dilutions (RALY KO). In parallel, as a control, single cell clones of HeLa cells transfected with an empty PX330 vector were also isolated (EV). The level of RALY was assessed in a RALY KO and an EV clone both by Western blot and immunofluorescence analyses (**Fig. 7A**). Then, the synthesis of new nuclear RNA was measured over time both in the untreated and mild ActD (125 ng/ml) treated condition. In both the cases, RALY KO cells showed a slower increase of newly synthesized RNA inside the nuclei compared to control cells (**Fig. 7B**).

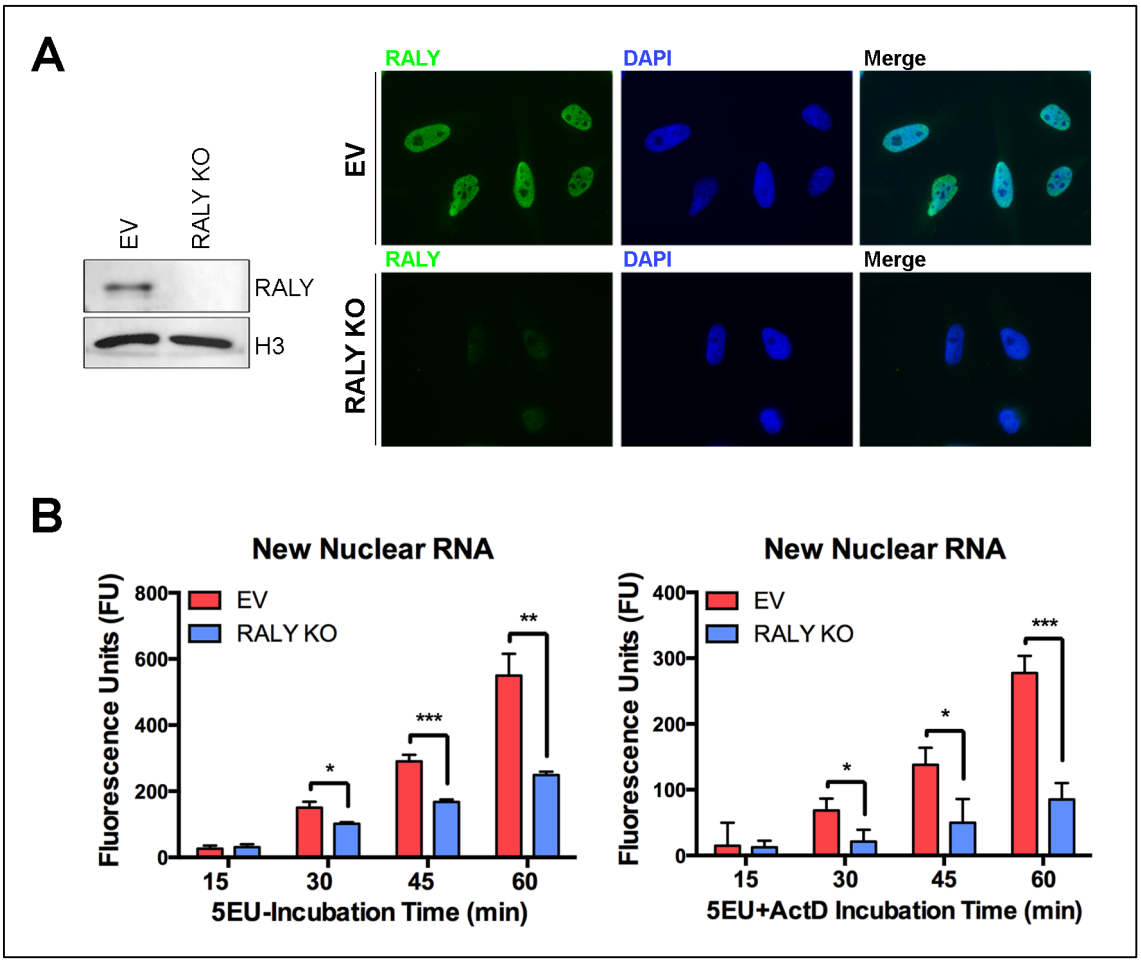

Figure 7. The knock-out of RALY determines, as for the RNAi experiments, a slower increase over time of new RNA Polymerase II-transcribed RNAs inside the nuclei. (**A**) RALY was knocked-out in HeLa cells through CRISPR/Cas9 technology. On the left, cell lysates of a RALY KO and an EV clone were subjected to SDS/PAGE and Western blot with the indicated antibodies. On the right, RALY KO and EV control cells were stained with an anti-RALY antibody and processed for immunofluorescence analysis. With both the techniques, RALY was not detected in the cells after CRISPR/Cas9 mediated knock-out. (**B**) RALY KO and EV control cells were incubated for 15, 30, 45 or 60 minutes with 5EU in absence (left) or presence (right) of ActD (125 ng/ml). 5EU was successively stained with 5FAM. The images of the samples were acquired with the High Content Imaging System Operetta (Perkin Elmer) and analyzed with the software Harmony (Perkin Elmer). The graphs show the increasing amount over time of newly synthesized RNA. In both the conditions, RALY KO cells showed a slower accumulation of newly synthesized RNA species inside the nucleus compared to EV control cells. **B** shows the mean of the signal of four distinct samples processed during the same experiment. The whole experiment was repeated 3 times, giving the same result. Bars represent mean ± S.D. P-value was calculated using unpaired two-tailed t-test (*P< 0.05; **P<0.01; ***P<0.001).

These results confirmed the RNAi experiments (**Fig. 5B and 6B**) and showed that both the downregulation and absence of RALY determine a slower increase over time of newly RNAPII-synthesized RNA species inside the nuclei.

The decrease in the amount of newly synthesized RNA in RALY downregulated cells could be caused either by a diminishment in RNAPII productivity, specifically in the elongation phase, or by a decreased number of transcriptional events.

3.5 The absence of RALY does not impair RNAPII elongation

Between the two possible reasons that could have explained the diminished synthesis of RNAs upon RALY downregulation, I decided to first study RNAPII elongation rate. In fact, previous experiments gave the hint for a possible influence of RALY on RNAPII dynamics. We recently observed that the silencing of RALY induced a strong downregulation of the linker histone H1X in both its RNA and protein forms [109]. Linker histones bind to nucleosomes near the entry/exit sites of the linker DNA and their commonly attributed function is to stabilize highly condensed chromatin [126]. However, they have been described to participate to different processes, including gene expression regulation, indicating different roles beyond mere structural components of chromatin [127–130]. The histone H1 family is composed by 11 members, of which H1X is the least characterized. H1X has been shown to accumulate inside nucleoli during the G1 phase of the cell cycle and to be instead distributed across the entire nucleus in the S phase [131]. Furthermore, H1X was observed to be preferentially located at the chromosome periphery in mitosis, and its ablation from the cells induced defects in chromosome alignment and segregation [132]. Mayor *et al.* observed H1X to co-localize with RNAPII on actively transcribed genes, proposing it to regulate RNAPII elongation speed to influence splicing [133]. Given this association and prompted by the effects of the downregulation of RALY on H1X expression, I measured RNAPII elongation speed in si-RALY transfected HeLa cells [134].

RNAPII-dependent transcription was blocked by treating si-CTRL and si-RALY transfected cells with the specific and reversible RNAPII inhibitor DRB for 3 hours [135]. During this incubation, the pre-mRNAs inside the cells are either processed or degraded, but no new RNA is transcribed by RNAPII. The cells were successively washed with PBS to eliminate DRB and were then incubated with warm medium to recover transcription in a synchronized manner. Total RNA was extracted at 5 minute intervals till reaching 120 minutes after transcription recovery. Along the increasing intervals, RNAPII will localize

at different positions along transcribed genes, depending on its elongation speed. Using different exon-intron primer pairs on long genes, it is possible to detect its position, and therefore speed, through qRT-PCR. Every measured signal was normalized on the respective result obtained in si-CTRL and si-RALY cells not treated with DRB (UT). A resuming scheme of the technique is depicted in **Fig. 8**.

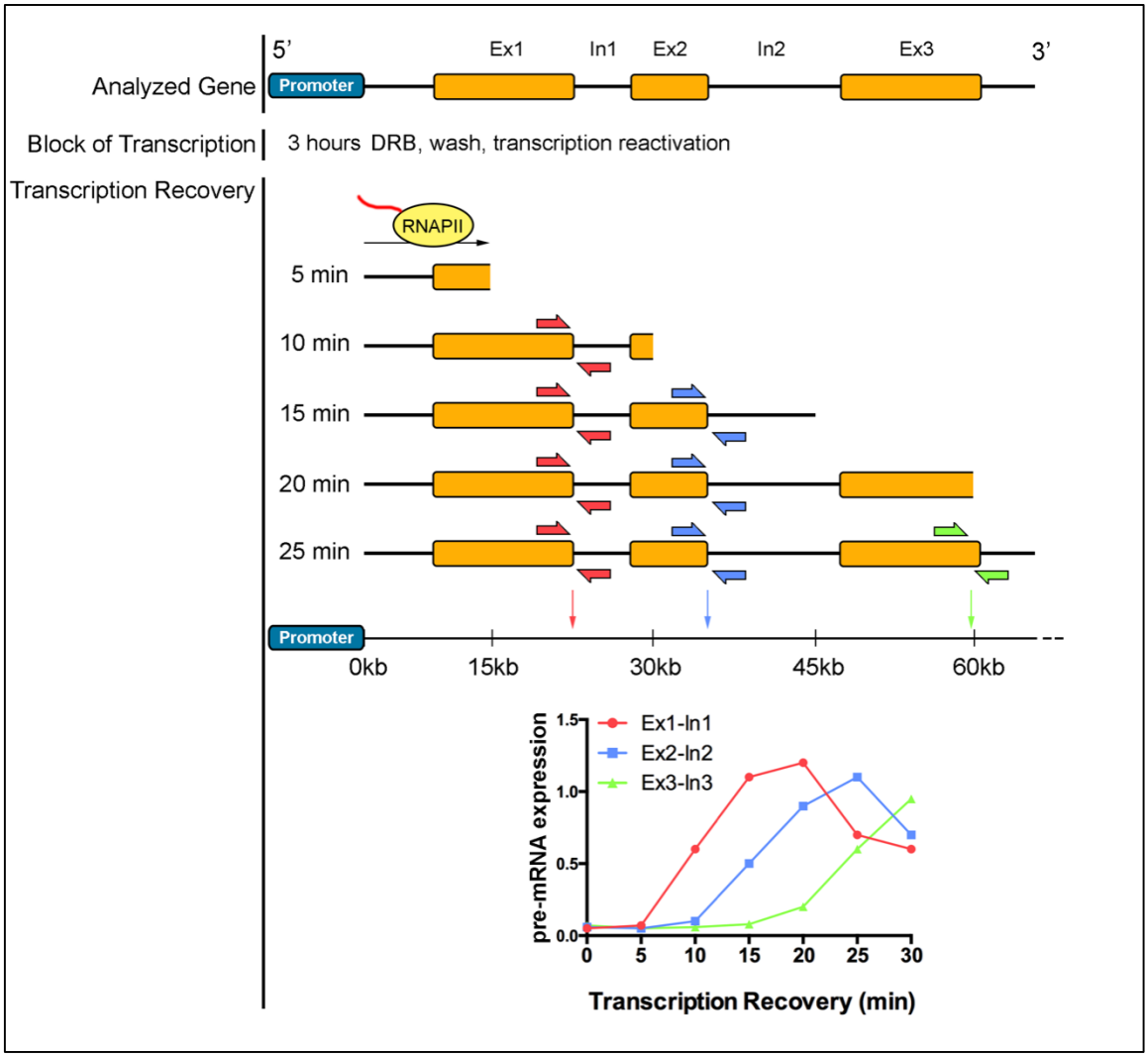

Figure 8. Measurement of RNAPII elongation speed. RNAPII-dependent transcription is blocked by treatment with DRB for 3 hours and synchronously re-activated by incubation of the cells in fresh medium. After re-activation, RNAPII starts again to transcribe genes. By extracting total RNA at different times after re-activation is possible to detect the position of RNAPII through qRT-PCR using exon-intron spanning primers on long genes. The procedure is described in details by Singh and Padgett [134].

RNAPII elongation speed was measured on the pre-mRNAs of *ITPR1*, *OPA1* and *CTNNBL1*, which were selected due to their length. In fact, long genes are needed for this technique in order to distinguish the localization of RNAPII at 5 minute intervals of

transcription. In addition, these genes showed different expression levels upon RALY downregulation, allowing a more detailed study of the connection between RALY and RNAPII-dependent transcription. More specifically, in si-RALY transfected cells, *ITPR1* level was found to be unchanged, while *OPA1* and *CTNNBL1* were measured to be respectively upregulated and downregulated compared to si-CTRL cells (**Fig. 9A**).

For the analysis of the elongation rate, the first time point presenting an average signal above the 10% of the UT condition was considered as the moment in which RNAPII reached that specific sequence. For all the analyzed genes, RNAPII elongation rate did not significantly differ between si-CTRL and si-RALY transfected HeLa cells (**Fig. 9B-D, Table 1**). However, in some cases, a hint of transcription slow down was noticeable in si-RALY cells: further experiments with more sensitive approaches could analyze more precisely the elongation rate of RNAPII in response to RALY downregulation.

Interestingly, in our experiments, the amount of synthesized pre-mRNA was constantly lower in RALY downregulated cells in respect to control cells, suggesting that even if RNAPII elongation rate was not significantly affected, a lower number of transcriptional events was taking place.

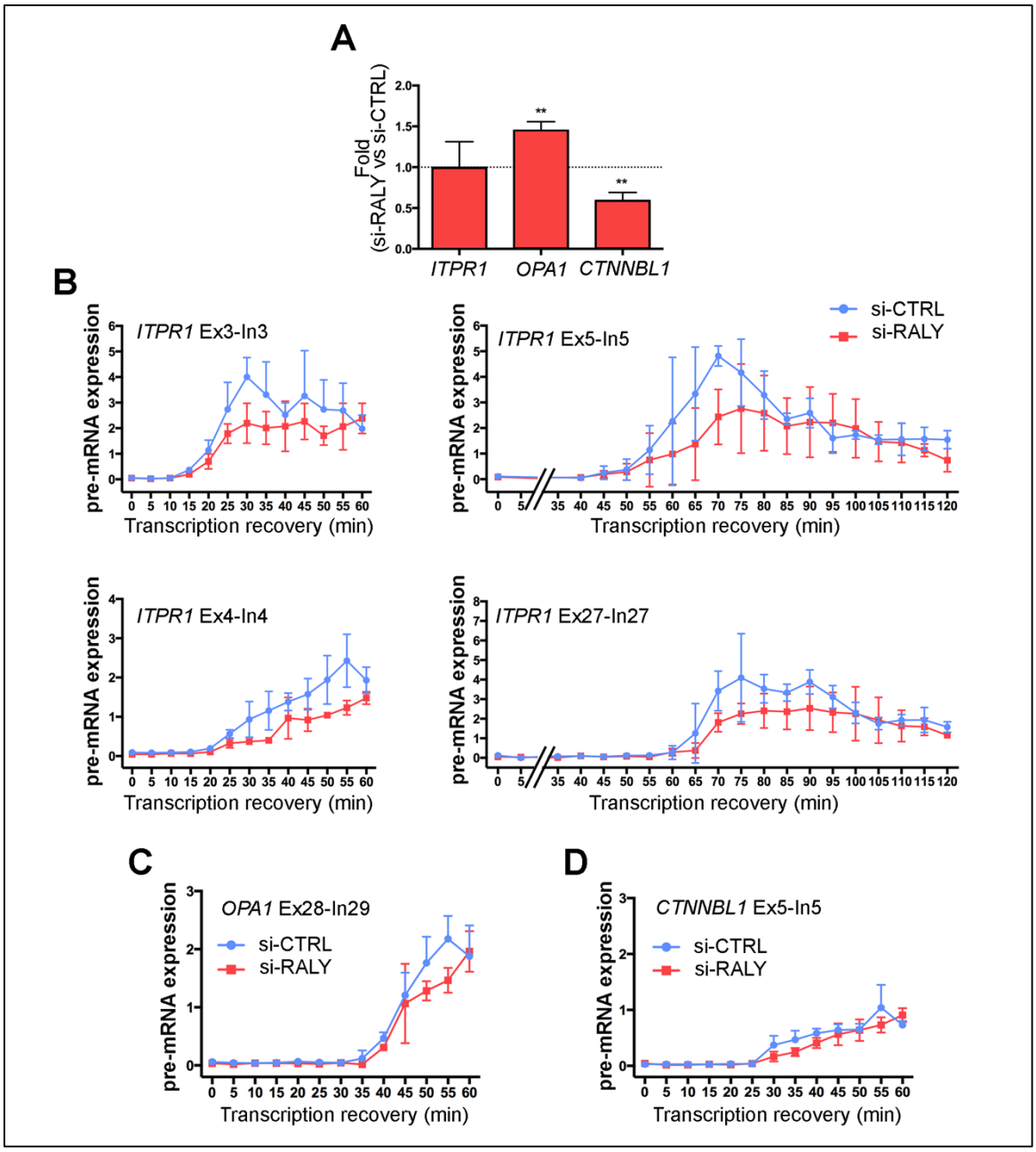

Figure 9. **The downregulation of RALY does not affect RNAPII elongation rate.** (**A**) HeLa cells were transfected for 72 hours either with si-CTRL or si-RALY. The expression level of *ITPR1*, *OPA1*, and *CTNNBL1* mRNAs was quantified by qRT-PCR and normalized on *GAPDH*. The level of *ITPR1* mRNA does not significantly change, while *OPA1* and *CTNNBL1* levels respectively increase and decrease in si-RALY cells compared to si-CTRL cells. The graph shows the mean of three independent experiments ± S.D. P-value was calculated using an unpaired two tailed t-test (**P<0.01) between si-RALY and si-CTRL transfected cells. (**B-D**). HeLa cells were transfected for 72 hours either with si-CTRL or si-RALY and successively treated with 100µM DRB for 3 hours to block RNAPII-dependent transcription. The cells were then incubated in fresh medium after DRB removal to recover transcription. Total RNA was extracted at 5 minute intervals and analyzed by qRT-PCR. The expression values are plotted relative to the expression level of the no DRB treated control, which is set to 1 in all experiments. RALY does not impair RNAPII elongation rate. The graphs show the mean of the signal of three independent experiments. Bars represent mean ± S.D.

Gene Region	Length (kb)	Time to transcribe (min)		Elongation rate (kb/min)	
		si-CTRL	si-RALY	si-CTRL	si-RALY
ITPR1 Ex1-Ex3	23.5	10	10	2.32	2.32
ITPR1 Ex1-Ex4	28	15	15	1.86	1.86
ITPR1 Ex1-Ex5	133	40	40	3.33	3.33
ITPR1 Ex1-Ex27	190	55	55	3.46	3.46
OPA1 Ex1-Ex29	100	30	35	3.33	2.86
CTNNBL1 Ex1-Ex5	64	25	25	2.5	2.5
		Average elongation rate		**2.8±0.66**	**2.72±0.61**

Table 1. Calculation of the average RNAPII elongation rate in si-CTRL and si-RALY transfected HeLa cells.

Transcription is the initial step of the genetic information flow, which will ultimately lead to the synthesis of proteins. For this reason, it is at the basis of the regulation of higher order cellular processes, as for example the renewal of cellular components, cell proliferation and the response to external stimuli. Vice versa, these processes regulate transcription of specific sets of genes to determine cellular behavior. The decrease in transcription observed upon RALY downregulation/KO could therefore be connected to the misregulation of other cellular processes, such as cell proliferation. In fact, we recently described how the silencing of RALY determined the downregulation of different genes involved in the regulation of cell cycle and how RALY bound RNAs were enriched in transcripts coding for proteins involved in cell cycle progression [109]. Transcription is tightly associated with cell growth: highly proliferating cells show high levels of transcription to sustain the intense demand of gene products necessary to maintain a quick growth rate. To this aim, these cells tend to overexpress versatile transcription factors, such as members of the MYC and EF2 family, which can supply the necessary gene products to promote their growth and division [136]. Taken all these considerations together, I investigated the possible involvement of RALY in the regulation of different cell cycle-related genes. In fact, the absence of RALY could alter the processing of the mRNA coding for factors involved in cell proliferation, ultimately affecting cell growth and transcription.

3.6 The downregulation of RALY impairs the expression of cell proliferation related genes

To study the effects of the absence of RALY on gene expression, we analyzed the gene expression profile of si-RALY transfected HeLa cells through microarray. We used a HTA2.0 microarray (Affymetrix Human Transcriptome Array) to analyze three different biological replicas of si-CTRL and si-RALY transfected HeLa cells. The levels of RALY were measured by qRT-PCR and were detected to be lower than the 10% in si-RALY transfected cells compared to si-CTRL cells (**data not shown**). Upon RALY downregulation, we found 1971 differentially expressed genes (**Table 2 in Appendix**), of which 919 were upregulated and 1052 were downregulated (**Fig. 10A**). The majority of the variations were observed in protein coding transcripts (93.5%), with a small percentage of non-coding RNAs, generally higher among upregulated genes (3% of long non-coding RNAs) (**Fig. 10B**). Functional annotation enrichment analysis by Gene Ontology and pathway databases identified extracellular matrix organization, regulation of cell motility and cell adhesion as biological processes significantly associated with the upregulated genes. On the other hand, cell cycle control and RNA splicing were the two broad themes more strongly associated with the downregulated genes upon RALY silencing (**Fig. 10C and Table 3 in Appendix**). Both in HeLa and MCF7 cells, a cell proliferation related class resulted to be an enriched class among downregulated genes in absence of RALY.

Interestingly, several genes coding for transcription-promoting factors were also present in the downregulated genes upon RALY silencing (**Table 2 in Appendix**).

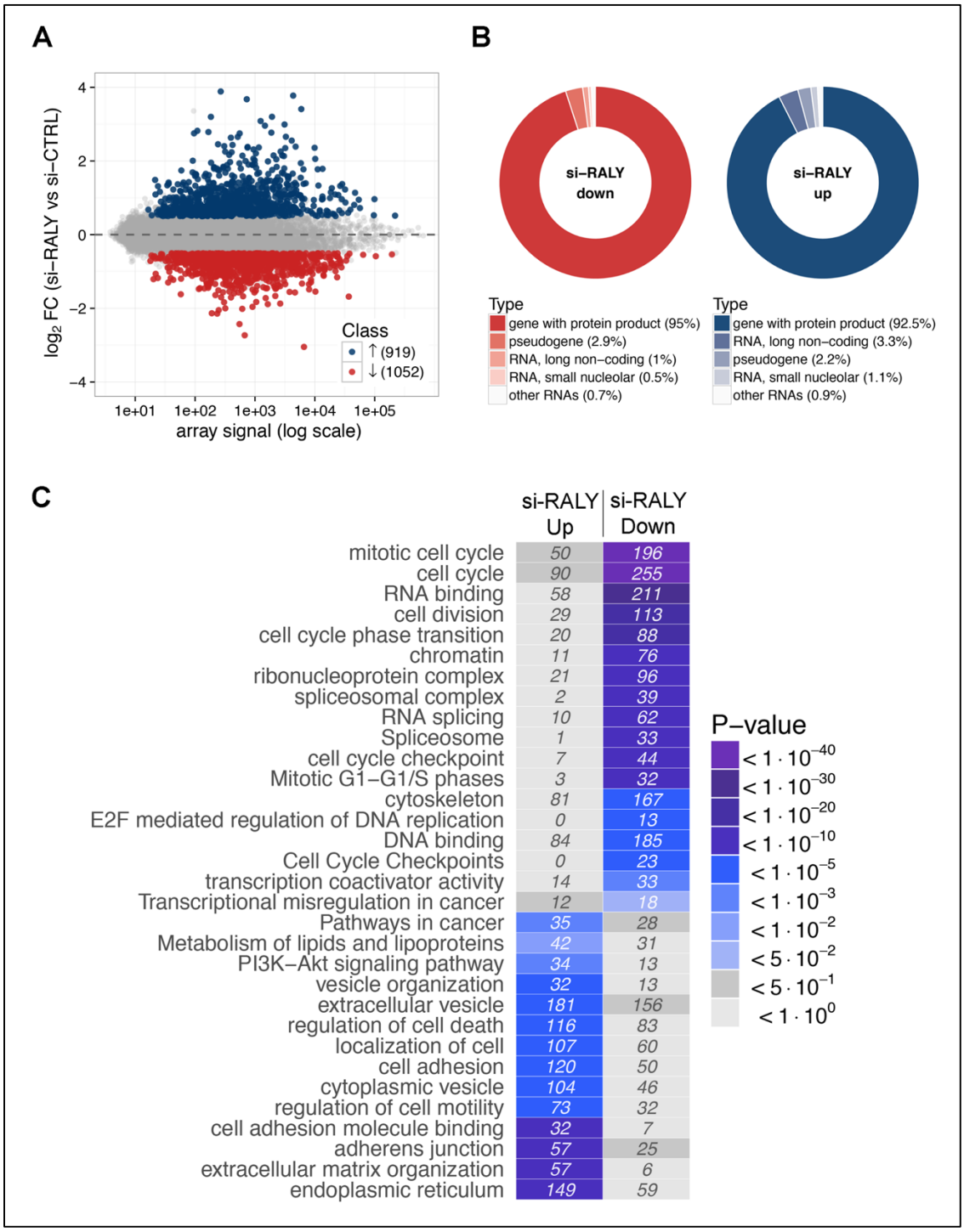

Figure 10. RALY silencing alters the transcriptome of HeLa cells. Three independent microarray experiments were performed using RNA preparations from three independent biological replicates. HeLa cells were transfected for 72 hours with either si-RALY or si-CTRL. (**A**) MA plot of RALY silencing transcriptome profiling. For each gene, the average log10 signal against the RALY silencing log2 Fold Change (si-RALY versus si-CTRL) is plotted. Genes significantly upregulated (blue) or downregulated (red) upon RALY silencing are highlighted. (**B**) Classification of differentially expressed genes (DEGs) upon RALY silencing according to RNA classes. (**C**) Functional annotation enrichment analysis of RALY upregulated and downregulated genes. The heat map, colored by enrichment p-value, displays enriched classes from Gene Ontology terms and KEGG or REACTOME pathways. The number of DEGs falling in each category is displayed inside each tile.

The enrichments of cell cycle- and transcription-related genes suggested to focus on the downregulated genes upon RALY silencing.

To validate the downregulation events detected by microarray, the expression level of factors promoting transcription and/or cell proliferation was measured by qRT-PCR. More specifically, for the cell growth I analyzed Cyclin B1, Cyclin B2 and the Cyclin Dependent Kinase 1 (*CCNB1, CCNB2 and CDK1,* respectively), which together drive the G2/M transition; Cyclin E1 and Cyclin E2 (*CCNE1* and *CCNE2*), two regulators of the entry in S phase; the M-phase inducer phosphatase 1 (*CDC25A*), a phosphatase involved in both the S-phase entry, acting on the CCNE1/Cdk2 complex, and the G2-phase transition, participating to the activation of CCNB1/Cdk1 [137–139]. Regarding the regulation of transcription instead, I analyzed Cyclin T1 (*CCNT1*), member of the positive transcription elongation factor; *GTF2A1* and *GTF2E2*, subunits of general transcription factors important for the assembly of the RNAPII machinery on the promoter and for the initiation of transcription; and the elongation factor ELL2 (*ELL2*), member of the super elongation complex [2,21,140–142]. In addition, I measured also the RNA levels of SPT16 (*SUPT16H*) and SSRP1 (*SSRP1*), subunits of the chromatin remodeling FACT complex, which with its chromatin relaxing activity favors the activity of RNAPII throughout nucleosomes, and of the transcription factor Dp-1 (*TFDP1*), a dimerization partner of E2F1-6 that stimulates E2F dependent transcription [143–146].

All the target mRNAs were found to be downregulated in si-RALY transfected HeLa cells, confirming the microarray analysis (**Fig. 11A**). In addition, the protein levels of CCNE1 and CCNT1 were measured by Western blot analysis. Both the proteins were found to be downregulated in si-RALY compared to si-CTRL transfected cells (**Fig. 11B and C**).

Taken together, these results show that the silencing of RALY induces a broad downregulation of transcription and cell proliferation related genes, possibly explaining the reduced RNAPII-dependent transcription phenotype and the reduction in cell proliferation observed by Rossi *et al.* [109].

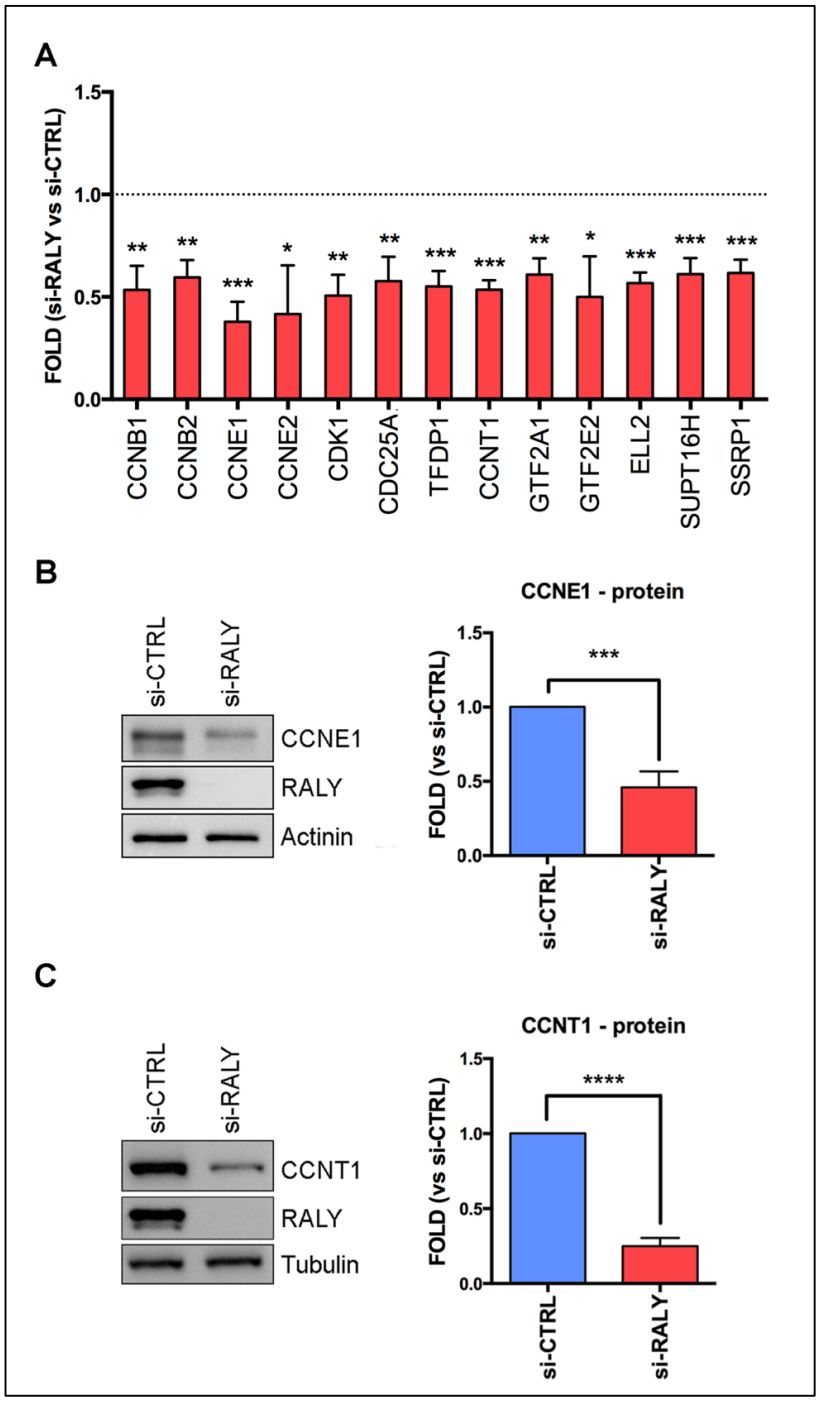

Figure 11. Validation of the microarray results for the downregulated genes upon RALY silencing. HeLa cells were transfected either with si-RALY or si-CTRL for 72 hours and total RNA and proteins were extracted. (**A**) The mRNA level of different factors promoting proliferation (*CCNB1*, *CCNB2*, *CCNE1*, *CDK1*, *CDC25A*, *TFDP1*) and transcription (*CCNT1*, *GTF2A1*, *GTF2E2*, *ELL2*, *SUPT16H*, *SSRP1*) were measured by qRT-PCR and normalized on *GAPDH*. As in the microarray analysis, all the analyzed genes were found to be less expressed in si-RALY compared to si-CTRL cells. (**B-C**) Total protein extracts were subjected to SDS/PAGE and Western blot analysis with the indicated antibodies (on the left). The levels of CCNE1 (**B**) and CCNT1 (**C**) were measured by band densitometry analysis (on the right) and were found to be downregulated in si-RALY compared to si-CTRL transfected cells. All the graphs of the figure show the mean of three independent experiments ± S.D. P-value was calculated using an unpaired two tailed t-test (*P<0.05; **P<0.01; ***P<0.001).

3.7 RALY regulates the expression of E2F1

In order to improve the functional characterization of the transcripts presenting an altered expression upon RALY silencing, we performed a gene set enrichment analysis (GSEA). This process allows to compare the list of misregulated genes from a microarray analysis with the collection of "Hallmark" annotated gene sets provided by the Molecular Signatures Database (MSigDB), detecting enrichments in specific gene families and cellular functions [147]. This analysis identified "E2F targets" as the most significantly enriched class among genes downregulated upon RALY silencing (**Fig. 12A**). This gene set, according to the MSigDB definition, includes "g*enes encoding cell cycle related targets of E2F transcription factors*".

The E2F family comprises eight transcription factors playing key roles in the regulation of cell cycle progression. Its members can be divided into activators (E2F1-3) and repressors (E2F4-5, E2F6, E2F7-8) of cell proliferation. In particular, the E2F1-3 group, regulating the transactivation of target genes necessary for the G1/S phase transition of the cell cycle, was found to be downregulated upon RALY silencing in the microarray results (**Table 2 in Appendix**) [146,148–150]. Interestingly, *E2F1* and *E2F2* mRNAs were detected as RALY mRNA targets in the RIP-seq analysis performed by Rossi *et al*. in MCF7 cells (**Fig. 12B**) [109].

Among the activators of the E2F family I focused on E2F1, a key character of cell cycle regulation and apoptosis induction, also involved in carcinogenesis when misexpressed [136,151–154]. To validate the microarray analysis, the downregulation of E2F1 in response to RALY silencing was confirmed by qRT-PCR and Western blot analyses. Both the mRNA and protein forms of E2F1 resulted to be downregulated in si-RALY transfected HeLa cells (**Fig. 12C-D, respectively**).

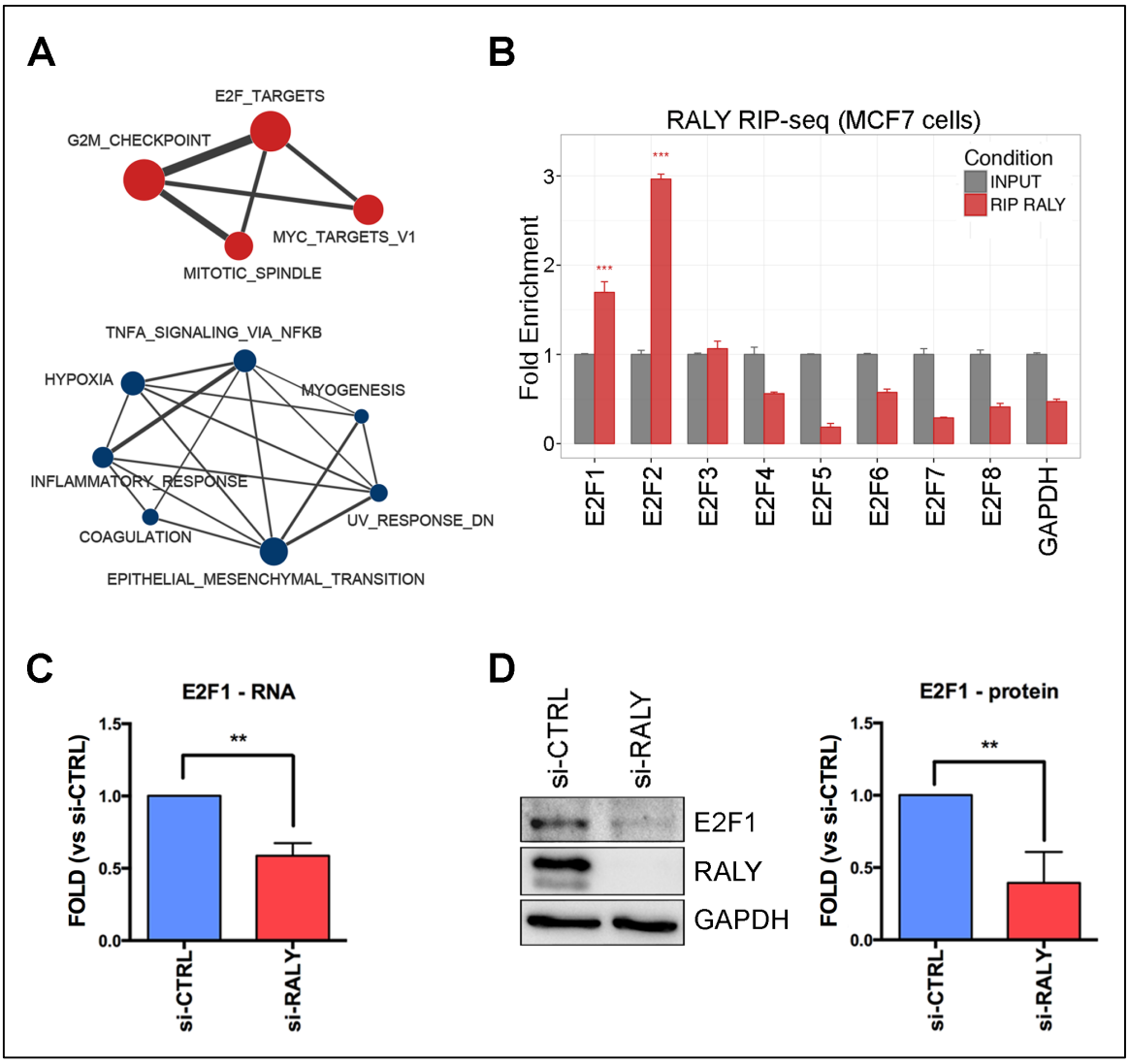

Figure 12. RALY regulates the expression of E2F1. (**A**) Functional map of global changes in gene expression in response to RALY silencing. Enrichment results from GSEA were mapped as a network of gene sets (nodes) related by mutual overlap (edges), where blue identifies upregulated and red downregulated gene sets. The size of each node is proportional to the size of the gene set. The size of each edge is proportional to the mutual overlap between two nodes. The gene set "E2F_targets", containing "genes encoding cell cycle related targets of E2F transcription factors", resulted to be the most enriched in downregulated genes upon RALY silencing. (**B**) The mRNAs coding for *E2F1* and *E2F2* were found to be enriched in RALY-containing immunoprecipitated particles in the RIP-seq analysis performed in MCF7 cells [109]. On the contrary, no enrichments were detected for the other members of the E2F family. (**C-D**) HeLa cells were transfected either with si-CTRL or si-RALY for 72 hours and both total RNA and total protein extracts were collected. The RNA was analyzed by qRT-PCR and the signals normalized on *GAPDH*, while proteins were subjected to SDS/PAGE and Western blot with the indicated antibodies followed by band densitometry analysis. Both the (**C**) *E2F1* mRNA and (**D**) E2F1 protein are downregulated in HeLa cells silenced for RALY. The graphs show the mean of three independent experiments ± S.D. P-value was calculated using an unpaired two tailed t-test (**P<0.01).

Successively, I analyzed the direct interaction between RALY and *E2F1* mRNA through RNA immunoprecipitation in UV-crosslinked HeLa cells. *E2F1* mRNA was observed to be enriched in RALY-containing immunoprecipitated complexes (**Fig. 13A**). *RALY* and *GAPDH* mRNAs were respectively used as positive and negative controls (**Fig. 13A**) [109]. We previously observed RALY to preferentially interact with poly-U stretches of four or more Uridines in the 3'UTR region of transcripts [109]. To confirm the direct interaction and to identify the region bound by RALY on E2F1 mRNA, I selected a suitable poly-U stretch in the 3'UTR of *E2F1* mRNA and performed an RNA pull-down experiment using both a wild-type RNA probe containing the poly-U stretch and a mutated RNA probe where the poly-U stretch had been interrupted. As positive control, I used the probe containing the poly-U stretch of *H1X* 3'UTR, which we previously observed to be bound by RALY [109]. The pulled-down proteins were then analyzed by Western blot. Interestingly, RALY was captured only by the wild-type sequence, while not binding the mutated probe (**Fig. 13B**). On the contrary, neither Tubulin nor Actinin were observed to be captured by either of the probes, excluding the possibility of unspecific bindings (**Fig. 13B**). As expected, the positive control pulled-down RALY from the cell lysate (**Fig. 13B**). These results demonstrate that RALY directly binds *E2F1* mRNA.

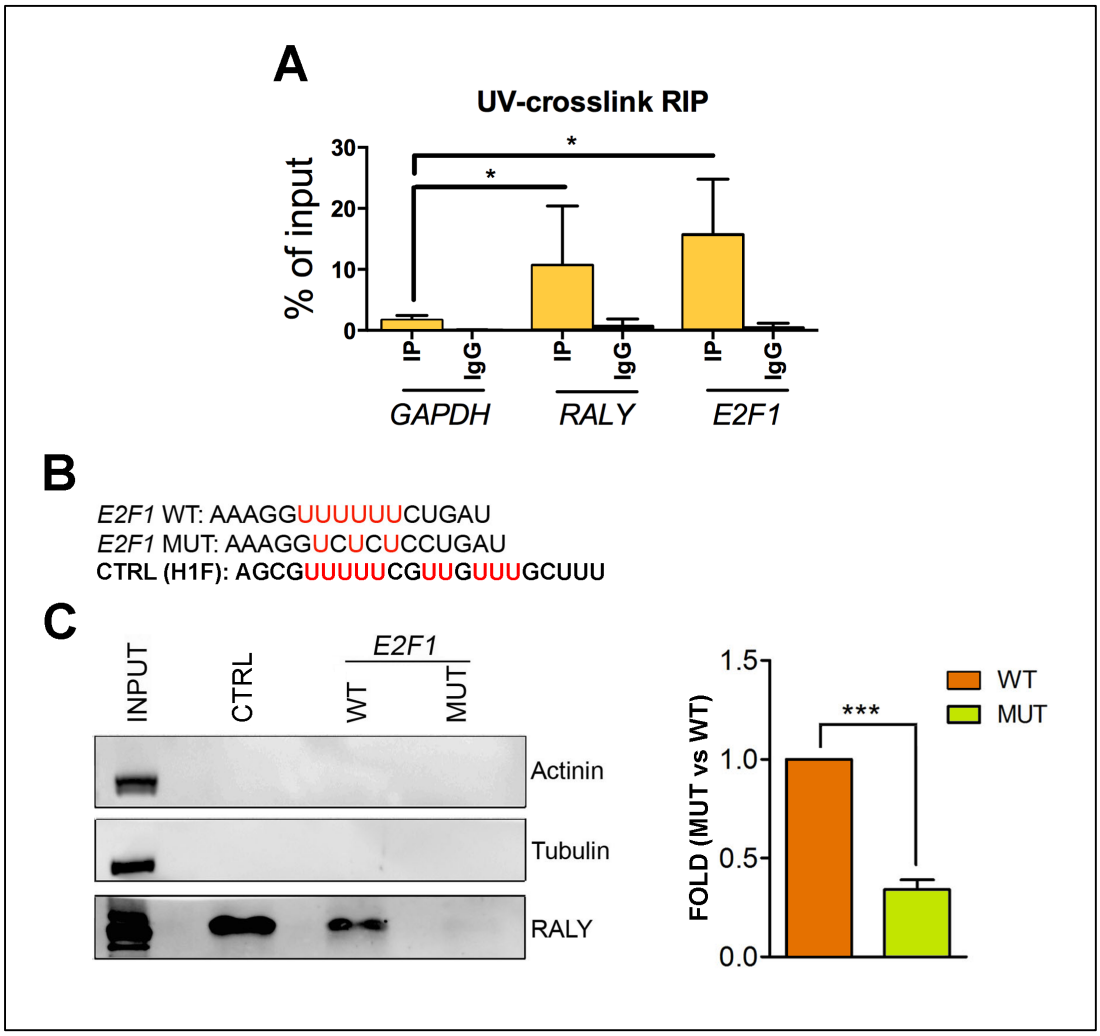

Figure 13. RALY binds E2F1 mRNA on a poly-U stretch in the 3'UTR. (**A**) qRT-PCR was used to compare the indicated mRNAs isolated upon UV-crosslinked RALY immunoprecipitation with RNA recovered after immunoprecipitation with IgG. *E2F1* mRNA is enriched in RALY-containing RNPs. The relative abundance was compared to the 10% of input. The experiments were performed at least three times. Bars represent means ± S.D. P-value was calculated comparing the amount of each mRNA with the amount of GAPDH using an unpaired two tailed t-test (*P<0.05). (**B**) Sequences of the biotinylated wild-type and mutant *E2F1* 3'UTR probes, and of the positive control H1FX 3'UTR probe. (**C**) HeLa cells total protein extract was incubated with either *E2F1* wild-type or mutant biotinylated RNA probes (50 pmol) and captured by streptavidin beads. *H1X* 3'UTR biotinylated probe was used as positive CTRL, as described by Rossi and colleagues [109]. Western blot of RNA-bound proteins shows the positive in vitro interaction of RALY with the wild-type poly-U sequence, but not with the mutant probe. Immunoblotting with anti-Actinin and anti-Tubulin served as negative control. The graph shows the mean values of three independent experiments. Bars represent mean ± S.D. P-value was calculated by unpaired two-tailed t-test (***P < 0.001).

Since the 3'UTR of transcripts is often targeted by *trans*-acting factors to determine the fate of the mRNA, I decided to investigate a possible regulatory activity of RALY on *E2F1* mRNA at the post-transcriptional level. In the range of possible post-transcriptional modifications, I analyzed transcript stability since other members of the hnRNP family, in

particular the hnRNP-C that with RALY shares a high similarity in the RRM, are described to regulate the stability of specific transcripts [56,59,64]. To investigate a possible similar function of RALY, I analyzed the stability of *E2F1* mRNA in absence of RALY. HeLa cells transfected with either si-CTRL or si-RALY for 72 hours were treated with ActD (5 μg/ml) for different periods of time to prevent new RNA synthesis. The stability over time of the remaining RNAs was analyzed by qRT-PCR. As depicted in **Fig. 14A**, the stability of *E2F1* in RALY silenced HeLa cells was measured to be lower over time, compared to si-CTRL cells, with a half-life of 7.64±1.32 hours. In si-CTRL cells, *E2F1* mRNA was found to be stable over the 9 hours analyzed time-frame, and its half-life could not be reliably calculated. As a control, I used the mRNA coding for GAPDH, which is not a target of RALY (**Fig. 13A**) and which was described to be suitable for assessing general mRNA stability [155]. *GAPDH* mRNA was found to be stable over 9 hours in both si-CTRL and si-RALY transfected cells (**Fig. 14B**).

Taken together, these results show that in absence of RALY the stability of *E2F1* mRNA is reduced. As a consequence, both E2F1 mRNA and protein are downregulated in RALY silenced cells.

3.8 RALY has a positive effect on E2F1 mRNA and many of its targets

In principle, only a fraction of the misregulated mRNAs detected by microarray analysis was affected in its expression consequently to the loss of direct contact with RALY. To preliminarily investigate the global effect of the downregulation of RALY on its target mRNAs, I intersected the microarray list of differentially expressed genes upon RALY downregulation with the list of RALY targets identified by RIP-seq in MCF7 cells by Rossi and colleagues [109]. From this merge, considering the provenience of the data from two different cell lines, I obtained 193 upregulated genes and 359 downregulated genes of the microarray that were also detected as direct RALY targets (**Fig. 14C**). Interestingly, there was an increased frequency of RALY RNA targets among genes downregulated upon RALY silencing (34.1% vs 21%, P-value 1.3e-10) suggesting that the loss of a direct RALY-mRNA interaction mediates the downregulation of the target transcript (**Fig. 14C**).

This analysis, even though preliminar, suggests a positive effect of RALY in determining the fate of its target RNAs.

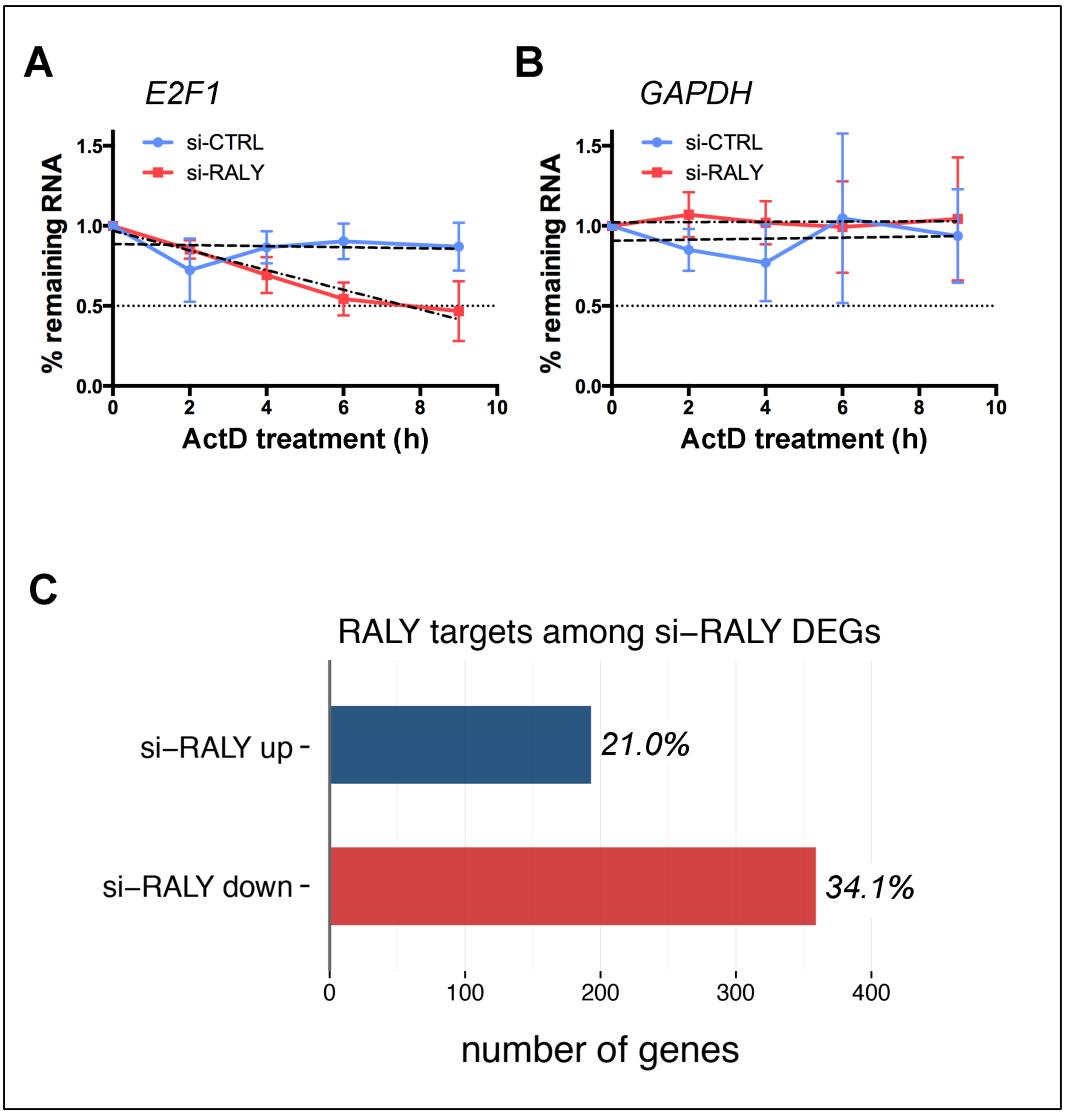

Figure 14. RALY stabilizes *E2F1* mRNA. (**A-B**) HeLa cells were transfected either with si-RALY or si-CTRL for 72 hours and successively treated with ActD (5 µg/ml) for 0, 2, 4, 6 or 9 hours. Total RNA was extracted, analyzed through qRT-PCR and normalized on *ACTB*. (**A**) *E2F1* mRNA was detected to be less stable in si-RALY transfected cells compared to si-CTRL cells. *E2F1* mRNA was measured to have a half-life of 7.64±1.32 hours in si-RALY cells, while being stable for over 9 hours in si-CTRL cells. (**B**) On the contrary, *GAPDH* mRNA showed to be comparably stable between si-RALY and si-CTRL transfected cells. (**C**) Intersection between the lists of si-RALY up (blue) or down (red) regulated genes in HeLa cells and the list of RALY RIP-seq targets previously identified in MCF7 cells (DEG=Differentially Expressed Gene) [109]. The percentage of overlap with respect to the number of DEGs is displayed beside the corresponding bar. Although the intersection could be affected by cell line differences, the increased frequency of RALY RNA targets among genes downregulated upon RALY silencing suggests that the loss of a direct RALY-mRNA interaction is associated with the downregulation of the target. Following the analysis of *E2F1* mRNA (**A**), it is tempting to suggest a positive role of RALY in determining the fate of its target mRNAs.

3.9 The absence of RALY impairs cell proliferation

To confirm the importance of RALY in the regulation of cellular physiology, I measured the proliferation rate and cell cycle distribution of RALY KO cells. I used the CRISPR/Cas9 engineered cells to have a stable RALY negative cell population. In fact, siRNA transfected cells showed a gradual recovery of RALY expression and a slow take-over of not-silenced cells after 5 days from transfection, which would have made the measurement not reliable (**Fig. 15A**). First, I measured E2F1 protein expression in the RALY KO cell model, confirming the downregulation of E2F1 (**Fig. 15B**). Successively, I compared the proliferation of RALY KO and control cells with two different approaches: one based on the count of cells through fluorescence (**Fig. 15C**), another based on the xCELLigence Real Time Cell Analysis platform (**Fig. 15D**). The latter measures the electric impedance of the wells where the cells are plated, which will increase proportionally to the surface occupied by the cells. Interestingly, both the approaches showed a slower cell growth in the RALY KO cell population compared to EV control cells. In the xCELLigence experiment, RALY KO cells did not reach the confluency plateau during the 62 hours analyzed time frame, while EV control cells stopped growing because of confluency around 45 hours after seeding. RALY KO cells do not show an exponential grow phase, presenting instead a linear growth.

To measure the distribution of RALY KO cells through the different phases of the cell cycle, I used an approach based on fluorescence microscopy. I cultured the cells for 24 hours and I incubated them with 5-Ethynyl deoxyuridine (5EdU), a thymidine analogue that is incorporated in DNA during cell replication. Successively, I stained 5EdU with 5FAM as for the RNA synthesis assay and I quantified the DNA content of every cell through the High Content Imaging System. Based on the work done by Massey and colleagues, I analyzed the cells to sort them in the different phases of the cell cycle depending on their DNA content [156]. As shown in **Fig. 15E**, RALY KO cells were detected to be enriched in the G1 phase and in parallel less present in the S and G2 phases of the cell cycle compared to EV control cells. These results are in accordance with the downregulation of numerous proliferation and cell cycle related genes in cells where RALY was silenced and finally confirm RALY as an important player in the regulation of cell proliferation.

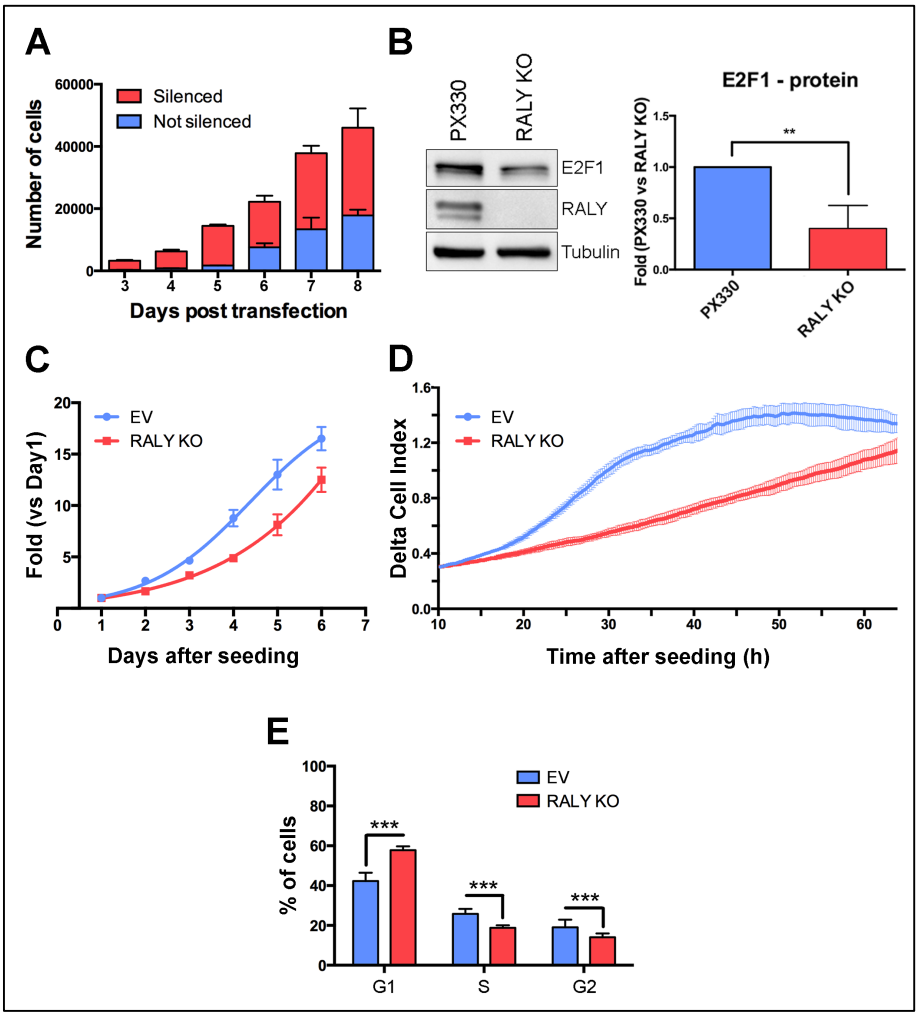

Figure 15. The absence of RALY impairs cell proliferation. (**A**) HeLa cells were transfected with si-RALY and si-CTRL and then fixed and processed for immunofluorescence after different days post transfection. RALY was stained with an anti-RALY antibody and an intensity threshold was set to discriminate the RALY downregulated cells. The population of not RALY-downregulated cells gradually increases from 5 days after transfection. The graph shows the mean values of three independent experiments. Bars represent mean ± S.D. (**B**) Total RALY KO and EV cell lysates were subjected to SDS/PAGE and Western blot with the indicated antibodies. The bands were analyzed by densitometry analysis. E2F1 is downregulated in RALY KO cells compared to controls. The graph shows the mean values of three independent experiments. Bars represent mean ± S.D. P-value was calculated by unpaired two-tailed t-test (**P < 0.01). (**C**) EV and RALY KO cells were fixed and stained with DAPI different days after seeding. The number of cells was measured counting the nuclei through the High Content Imaging System, and every measurement was normalized on the number of cells present in the wells at Day 1. RALY KO cells grow slower compared to EV cells. The graph shows the mean values of three independent experiments. Bars represent mean ± S.D. (**D**) EV and RALY KO cells were seeded in a xCELLigence E-plate and cell proliferation was monitored for 62 hours. All the values were normalized on the impedance detected at 10 hours post seeding, considered as the moment where all the cells were attached to the surface and started to proliferate. RALY KO cells grow slower than EV control cells. The graph shows the mean values of three experimental replicas. The experiment was repeated three times, giving the same output. Bars

represent mean ± S.D. (**E**) EV and RALY KO cells were cultured for 24 hours and then incubated for 1 hour with 5EdU to stain newly synthesized DNA. 5EdU was then stained with 5FAM and the amount of DNA in each cell was measured through the High Content Imaging System, successively sorting the cells in the different phases of the cell cycle. RALY KO cells tend to stall in the G1 phase of the cell cycle, with a consequent lower percentage of cells in the S and G2 phases. The graph shows the mean values of three independent experiments. Bars represent mean ± S.D. P-value was calculated by unpaired two-tailed t-test (***P < 0.001).

4. Discussion

The aim of my thesis was to better characterize the role of the hnRNP RALY in RNA metabolism. In fact, RALY was described to bind numerous mRNAs and to regulate the expression of specific targets [109], but its function was still elusive. I showed RALY to be differently associated to distinct nuclear fractions, possibly to exert multiple functions, and to regulate the expression of several transcription- and proliferation- promoting factors, ultimately affecting global transcription and cell proliferation. Given the results of the microarray analysis, I focused on the role of RALY in the post-transcriptional regulation of the proliferation marker E2F1. I showed RALY to promote the stability of *E2F1* mRNA, eventually determining the amount of E2F1 protein inside the cells.

RALY contains an RRM at the N-terminal domain, two putative NLSs, and an RGG in the C-terminal region [105,114]. My results demonstrate that RALY interacts with transcriptionally active chromatin and that this association is partially abrogated upon RNA degradation (**Fig. 1C**). Based on my analyses, I propose that the association of RALY with transcriptionally active chromatin is mediated by two types of interaction: one RNA-dependent through the RRM domain, hypothetically to bind and process RNAs in synergy with other factors; one RNA-independent mediated by the C-terminal domain (**Fig. 2D**). Interestingly, the fragment of RALY involved in the RNA-independent interaction (amino acids 143-306) contains a region similar to a basic-leucine zipper-like motif (bZLM) presented by hnRNP-C and typical of DNA-binding proteins, more precisely located in between amino acids 146-214 [66,114]. Since the construct containing only the RGG domain was not found on transcriptionally active chromatin, I conclude that this bZLM is necessary for the RNA-independent association of RALY with active chromatin. This interaction might allow RALY to function as a transcriptional co-factor, binding DNA, but could also serve to anchor RALY on chromatin to allow its N-terminal RRM domain to contact newly synthesized transcripts [114]. In fact, different RBPs, some of which interacting with RALY, can associate with chromatin to have access to newly synthesized RNAs [69,75,100,157]. Further experiments will better characterize the binding properties of this predicted bZLM domain and the function of the two different associations of RALY to transcriptionally active chromatin.

To gain more information about the cellular distribution of RALY, I analyzed the localization of the RBP FUS, classified as hnRNP-P2, which was recently studied by Yang and colleagues (**Fig. 1A**) [94,100]. FUS is a well studied RBP because of its association with multiple diseases, in particular affecting the nervous system, playing different biological functions, including transcription regulation, RNA biogenesis and post-transcriptional modification [99,158,159]. FUS is mainly present in the nuclear-soluble and the transcriptionally active chromatin fractions, where it was described to have a transcriptional control function [100]. In particular, RALY and FUS share the association with transcriptionally active chromatin. However, the association of FUS with chromatin was described to be strictly RNA dependent, while a portion of RALY persists on chromatin even in absence of RNA, suggesting different molecular functions of the two RBPs. Considering that FUS was identified as an interactor of RALY through mass spectrometry and that RALY was characterized as a transcriptional co-factor, this shared association opens the possibility for a co-operative function of the two hnRNPs in transcriptional control, but also in RNA maturation [105,114].
The association of RALY with transcriptionally inactive chromatin is not affected by RNase treatment, leaving this interaction uncharacterized. Further experiments are needed to understand the biological function of this relationship.

RALY showed a partial RNA-dependence in the interaction with chromatin when RNA was degraded *in vitro*. *In vivo*, the presence of RNA on chromatin is mainly due to ongoing transcription and the inhibition of RNA synthesis might lead to changes in the association of RBPs with chromatin. The treatment with Actinomycin D induced a progressive delocalization of RALY from the nucleus, more specifically from transcriptionally active chromatin, to the cytoplasm, suggesting that the interaction of RALY with this fraction of chromatin requires the active synthesis of new transcripts (**Fig. 3B and 3C, Fig. 4A**). The block of transcription might abrogate the dynamic recruitment of RALY on chromatin and on newly synthesized RNAs, disrupting the existing interactions and preventing the formation of new associations. If the RNA-dependent and -independent interactions are differentially affected by the block of transcription remains to be elucidated.

Given the strong effect that the block of transcription had on RALY localization, I investigated if the reverse relationship was also standing, namely the absence of RALY affecting transcription. The downregulation of RALY induced a global decrease of RNAPII-dependent transcription (**Fig. 5-7**), without affecting RNAPII dynamics (**Fig. 9**). To explain the decrease of RNA synthesis, I analyzed the transcriptome profile of RALY silenced HeLa cells. In fact, being RALY an RNA-binding protein, it was plausible that the effects on cell physiology upon RALY downregulation were due to a misregulation of RALY mRNA targets. The microarray experiment detected an altered expression of several transcription- and cell cycle-related genes (**Table 2 in Appendix**), in accordance with the decreased RNA transcription described above and the misregulation of cell cycle previously described by Rossi and colleagues [109]. I proceeded validating by qRT-PCR the mRNA expression of factors involved in the two processes.

Regarding the decreased transcriptional activity of RALY silenced cells, I measured the mRNA expression of different transcription promoting factors, such as *CCNT1*, the Cyclin subunit of the Positive Transcription Elongation Factor b (P-TEFb), a complex responsible for the release of RNAPII from the promoter-proximal pausing and for the activation of RNAPII productive elongation; *GTF2A1* and *GTF2E2*, subunits of general transcription factors involved in transcription initiation, and *ELL2*, a member of the transcription-promoting Super Elongation Complex (**Fig. 11A**) [2,21,140,141,160]. In addition, the two subunits of the Facilitating Transcription (FACT) complex, *SUPT16H* and *SSRP1*, were found to be downregulated at the mRNA level in RALY silenced cells (**Fig. 11A**). The FACT complex was described to associate with active RNAPII and to promote nucleosome disassembly in order to facilitate transcription [143,144,161]. The downregulation of these transcription-promoting factors in RALY silenced cells, together with other transcription-related targets detected by microarray analysis, possibly explains the observed global decrease in RNA synthesis.

Regarding cell cycle progression, I validated the mRNA expression of different cell cycle-related factors, such as *CCNB1*, *CCNB2* and *CDK1*, which together drive the G2/M transition; *CCNE1* and *CCNE2*, two regulators of the entry in S phase, and *CDC25A*, a phosphatase involved in both the S-phase entry and the G2-phase transition (**Fig. 2A**) [137–139]. All the targets resulted to be significantly downregulated in RALY silenced cells.

In order to improve the functional characterization of the deregulated transcripts upon RALY silencing, I performed gene set enrichment analysis (GSEA) over the collection of "Hallmark" annotated gene sets provided by the Molecular Signatures Database (MSigDB) [147]. The analysis revealed an enrichment of cell cycle-related targets of the E2F family in the downregulated genes upon RALY silencing (**Fig. 12A**). The family of the E2F transcription factors comprises eight members (E2F1-8) that regulate cell physiology by either promoting or repressing cell cycle progression. A tight equilibrium stands between these two "factions" and the most abundant one will in the end decide the fate of the cell [146]. Interestingly, all the cell cycle promoting E2Fs (E2F1-3) were found to be downregulated by microarray analysis upon RALY silencing (**Table 2 in Appendix**). I focused on E2F1, a well-known marker of cell proliferation and promoter of S-phase, which can also induce apoptosis through different p53-dependent and -independent mechanisms [162–164]. Due to this dichotomous role as proliferation- and apoptosis-promoting factor, E2F1 was described as both oncogene and tumor suppressor gene [165]. The overexpression of E2F1 has been reported in numerous types of cancer, and its ectopic expression was described to drive S-phase entry in quiescent cells and hyperplasia [150,151]. I showed that RALY interacts with and regulates the stability of the mRNA coding for E2F1, determining the expression of E2F1 protein (**Fig. 12C-D, Fig. 13, Fig. 14A**). Inside the cell, E2F1, together with its activating partner transcription factor DP1 (TFDP1), is normally kept in an inactivated state by the Retinoblastoma protein (Rb) and is released at the G1/S phase transition after the phosphorylation of Rb by the Cyclin D-CDK4/6 dimers. The E2F1-TFDP1 complex will induce the expression of S-phase promoting factors, as Cyclin E1 and E2 (CCNE). In a positive feedback loop, CCNE will bind its partner CDK2 to further phosphorylate Rb, consequently increasing the levels of free E2F1-TFDP1 complexes [166–168]. At this point, the cell is strongly directed towards S-phase. Interestingly, an excess of CCNE was observed to promote G1 progression and a CCNE upregulation was found in different cancers, correlating with poor prognosis [162,169,170]. A decreased expression of E2F1 therefore de-potentiates the G1/S phase transition, lowering the expression of proliferation promoting genes, as CCNE [168]. As a consequence, I observed a lower expression of direct targets of E2F1 transcriptional control activity, as in fact *CCNE1* and *CCNE2,* and *CDC25A*, *CDK1* and *CCNB1* (**Fig. 11A**)

[167,168,171,172]. In addition, also *TFDP1* mRNA was found to be downregulated upon RALY silencing (**Fig. 11A**), suggesting a further decreased activity of all the E2F transcription factors.

Except *CDC25A*, the other mRNA targets were detected in the RIP-seq experiment on MCF7 cells performed by Rossi and colleagues [109]. Although these direct interactions were not validated in HeLa cells, no adverse effects were observed in their stability upon RALY downregulation (**data not shown**). Therefore, I cannot exclude that, beyond E2F1 downregulation, also the lost direct interaction with RALY somehow affected their expression. In fact, even though preliminarily, I observed that the loss of direct interaction RALY-mRNA is often associated with the downregulation of the transcript, suggesting a general positive role of RALY in determining the fate of its mRNA targets (**Fig. 14C**).

To confirm the effects of the downregulation of RALY on cell growth, I analyzed both cell proliferation and cell cycle distribution of RALY KO cells. In absence of RALY, cell proliferation proceeds slower and the cells tend to accumulate in the G1 phase of the cell cycle compared to control cells (**Fig. 15C-E**).

The observation of the downregulation of E2F1 upon RALY silencing, together with the lower expression of numerous transcription- and proliferation-promoting factors (**Fig. 11A and Table 2 in Appendix**), is in accordance with the observed diminished RNAPII-dependent transcription, decreased cell proliferation and misregulated cell cycle progression. In this picture, these processes are directly impaired by the downregulation of the respective promoting factors in response to RALY silencing, and also mutually affect each other. Therefore, I propose RALY as a novel indirect regulator of transcription and cell proliferation.

To bring the analysis of the interaction between RALY and E2F1 to a higher level, the ENCODE project found E2F1 to interact with the promoter of RALY, making tempting to consider RALY as a *bona fide* target of E2F1. Furthermore, the cBioPortal for Cancer Genomics reported RALY and E2F1 to be frequently co-amplified in tumours [173]. The colorectal adenocarcinoma was detected as the tumour type presenting the highest

number of co-amplification events of RALY and E2F1 (31 out of 212 analysed patients). In this malignancy, the expression of the two genes was positively correlated (Pearson coefficient = 0.7) [173]. This analysis intriguingly suggests the existence of a positive feedback loop where E2F1 enhances the expression of RALY, which in turn stabilizes *E2F1* mRNA. However, the fact that RALY and E2F1 loci are located close to each other on chromosome 20 has to be taken into account. Further investigations will better clarify the relationship between E2F1 and RALY.

Considering my results, pieces have been added to the understanding of the biological functions of RALY, but different questions regarding the molecular functions of RALY remain unanswered. Nevertheless, an involvement of RALY in the 3'end processing and termination can be hypothesized. In fact, Fasken and colleagues observed function homologies between the human RALY and the yeast protein Nab3, an RNA-binding protein involved in the exosome-mediated processing, termination and degradation of non-coding RNAs (ncRNAs) [174]. Nab3 and RALY share 31% of sequence homology in the RRM and a certain degree of similarity also in the C-terminal region, and the expression of RALY in Nab3 negative cells rescued the thermosensitive yeast phenotype. Interestingly, to be efficient, both the proteins needed a functional RRM [174]. RALY could be involved in the 3'end processing also of coding RNAs, and not only ncRNAs, ultimately affecting RNA stability and biogenesis, but further experiments are needed to better elucidate this possibility.

5. Conclusion and Future Perspectives

Taken together, the results of my PhD project better define the biology of the hnRNP RALY inside the nucleus, in particular highlighting a transcription-dependent association with transcriptionally active chromatin. The downregulation of RALY induced a decrease of RNAPII-dependent transcription inside the nuclei and affected the expression of numerous transcription- and cell proliferation- promoting factors. As a consequence, RALY KO cells show a slower proliferation and a stalling in the G1 phase of the cell cycle compared to control cells. Specifically, I focused on the transcription factor E2F1, marker of proliferation and inducer of apoptosis. I showed that *E2F1* transcript is enriched in RALY-containing ribonucleoparticles and that the levels of E2F1 protein in cells silenced for RALY were reduced due to a diminished stability of *E2F1* mRNA. In addition, I propose a general positive effect of RALY on its mRNA targets.

As future perspectives, the specific functions of the different interactions of RALY with transcriptionally active chromatin will be investigated in more details. The bZLM-dependent association paves the way for an in-depth study of RALY activity as transcriptional co-factor. It would be interesting to further characterize the possible DNA binding ability of RALY and, if successful, to successively identify a consensus sequence recognized by RALY bZLM to classify target DNA sequences and to understand the biological function of these interactions. With this study, it could be possible to attribute a direct function to RALY in regulating the expression of part of its targets.
In parallel, it would be also fascinating to better understand the function of the nuclear-soluble pool of RALY. In fact, RALY localizes in this fraction both in a RNA-dependent and -independent manner. In the nuclear-soluble fraction, the N-terminal RRM-containing fragment of RALY was affected by RNase A treatment, while the C-terminal domain maintained its localization (**Fig. 2**). It is reasonable to hypothesize that the block of RNA synthesis only affects the pool of RALY directly interacting with RNA, while not disturbing the remaining RALY engaged in interactions with other proteins, for example as member of RNPs or different protein complexes.

Regarding the second part of the project, it would be interesting to better investigate the role of RALY in respect to the downregulated genes upon its silencing. I showed how the absence of RALY determined a decreased stability of *E2F1* mRNA, but other direct targets of RALY did not exhibit the same behavior. Being RBP-complexes very heterogeneous and dynamic in their composition, I can hypothesize that the final effect of the direct interaction RALY-mRNA could depend on the partners of RALY present in the complex. Alternatively, since RALY was proven to bind poly-U stretches, preferentially in the 3' UTRs of transcripts, different poly-U stretches patterns could induce RALY to push the mRNA towards alternative fates [109]. It would be also interesting to better understand if the stabilizing effect RALY on *E2F1* mRNA occurs the the level of the nucleus or the cytoplasm. Further experiments are needed to find the solution to this puzzle and better characterize the inferred general positive effect of RALY on its mRNA targets.

6. Additional Results

6.1 RALY and Splicing

During my PhD, I tried to better characterize the involvement of RALY in the process of splicing. Hereby are presented the results I obtained exploiting different approaches of investigation.

RALY was previously shown as a member of the EJC and as an interactor of important splicing factors, such as U2AF2, RbFOX1 and PRP19, but its downregulation did not affect splicing [105,111,112,175]. However, a proper characterization of RALY in relation to splicing was still missing. For this reason, I investigated the involvement of RALY in the process of splicing, particularly focusing on splicing dynamics and efficiency in cells where RALY was downregulated.

The study of the involvement of RALY in pre-mRNA splicing was carried on together with the analysis of RNAPII elongation rate in absence of RALY. Following the "kinetic model" in fact, an influence of RALY on RNAPII elongation speed could have resulted in an alteration of splicing. Furthermore, the linker histone H1X, strongly downregulated in RALY silenced cells, was proposed to participate to the regulation of RNAPII speed to allow specific splicing events [109,133].

As first parameter, the timing of splicing was analyzed. To compare splicing speed between si-CTRL and si-RALY transfected cells, I used an approach similar to the measurement of RNAPII elongation rate (**Fig. 9**). In addition to the exon-intron spanning primers to detect the transcription of specific genomic areas, I used primers to identify splicing events: the forward primer was located in an exon, while the reverse primer was designed in the +1 downstream intron. The qRT-PCR would therefore amplify the sequence only when the exon and the +1-intron are closer, meaning when splicing occurred to extrude the intron immediately downstream to the selected exon. The qRT-PCR C_t value where the spliced transcript becomes detectable can be compared with the time taken by RNAPII to reach that specific exon, eventually obtaining the time took by

the spliceosome to extrude the +1 intron. A representative scheme of the technique is depicted in **Fig. 16**.

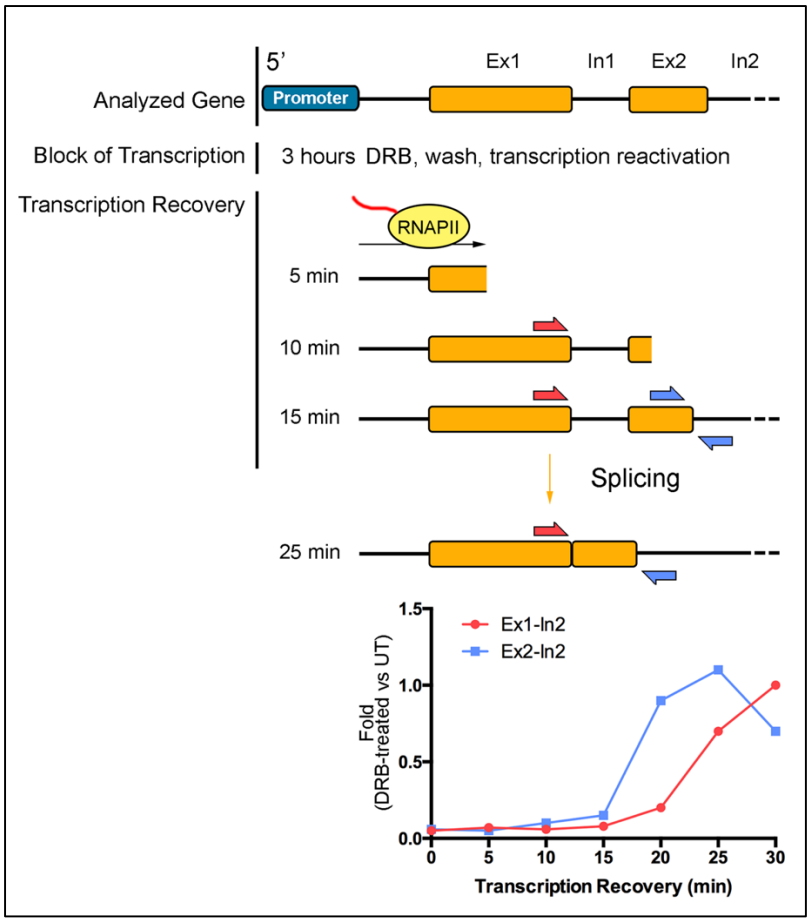

Figure 16. Measurement of splicing speed. RNAPII-dependent transcription is blocked by treatment with DRB (100 µM) for 3 hours and synchronously re-activated by incubation of the cells in fresh medium. After re-activation, RNAPII starts again to transcribe genes. By extracting total RNA at different times after re-activation is possible to detect the occurrence of splicing events through qRT-PCR using exon - +1-intron primers on long genes. The procedure is described in details by Singh and Padgett [134].

In literature, splicing was described to occur in a 10 minutes time window after the transcription of the acceptor splice site of the downstream exon [134]. In both si-CTRL and si-RALY HeLa cells, the splicing of the selected introns in the genes *OPA1* and *ITPR1* occurred in this time frame, suggesting that the timing of splicing was not affected by the silencing of RALY (**Fig. 17A and B**). These results are still preliminary. Further measurements of the splicing speed in other genes and especially in RALY RNA targets will better define any involvement of RALY in splicing dynamics. More sensitive techniques could offer a more precise characterization of splicing dynamics in absence of RALY.

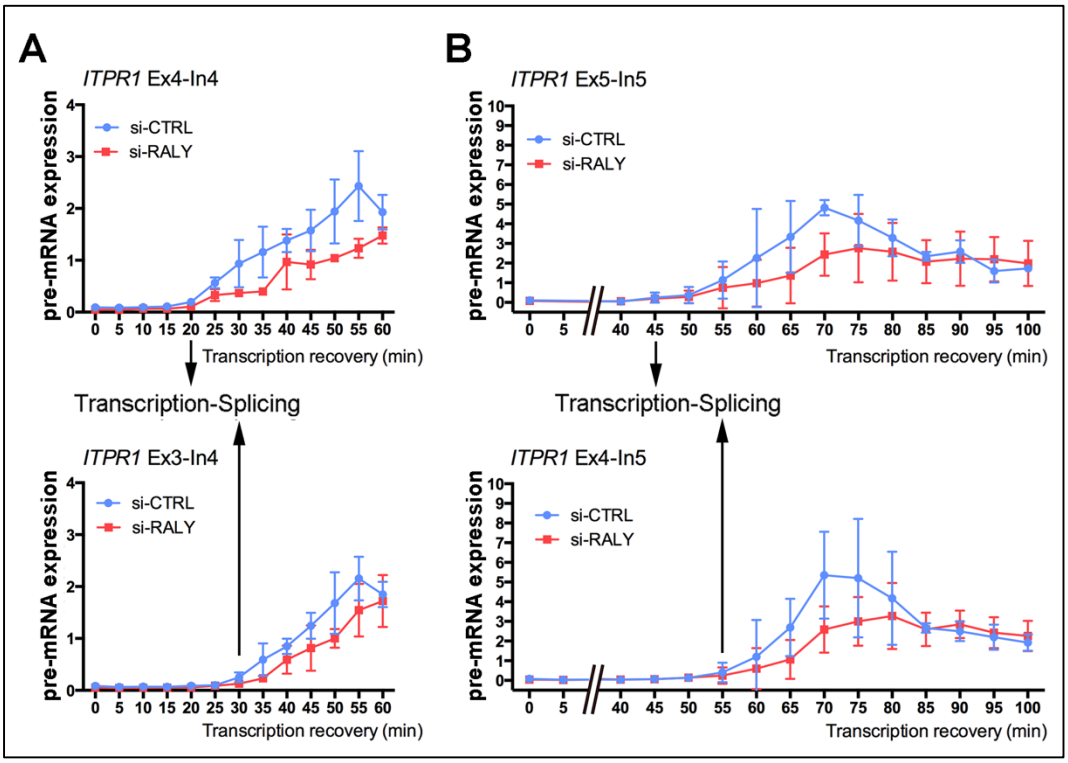

Figure 17. Analysis of the speed of splicing upon RALY silencing. (**A-B**) HeLa cells transfected either with si-RALY or si-CTRL for 72 hours were treated with DRB (100 µM) for 3 hours. After washing with PBS, the cells were incubated in DMEM to recover transcription. Successively, total RNA was collected at 5 minute intervals. qRT-PCR with different exon-intron and exon- +1-intron primers of *ITPR1* allowed the observation of the splicing timing of *ITPR1* intron 3 and 4 (left and right graphs, respectively). The pre-mRNA expression values are plotted relative to the expression level of the no DRB treatment control, which is set to 1 in all experiments. RALY does not impair splicing timing. All the graphs of the figure show the mean of three independent experiments ± S.D.

Next, I analyzed the efficiency of splicing in RALY downregulated HeLa cells. To this aim, I calculated the ratio between the mRNA and the pre-mRNA forms of different transcripts by qRT-PCR (**Fig. 18A**). To perform a more detailed analysis, the analyzed targets comprised both U2 and U12 introns. The minor spliceosome spliced introns were identified using the U12DB and the Ensembles online databases [176,177]. In particular, of the analyzed transcripts, *ASCC2* was detected as a mRNA target of RALY by RIP-seq in MCF7 cells, *Thoc2* was detected to be downregulated and *CDK6* and *EVI5* to be upregulated by microarray analysis in HeLa cells silenced for RALY (**Table 2 in Appendix**). No significant differences were detected between the constitutive splicing efficiencies of si-CTRL and si-RALY cells for all the analyzed splicing events (**Fig. 18B**).

To better investigate the possible relationship between RALY and alternative splicing, I successively analyzed the splicing pattern of *CD44*. This gene presents 10 constant exons and 9 clustered variable exons and, given its numerous isoforms, is considered as a reference gene for alternative splicing (**Fig. 18C**) [178]. The microarray analysis detected an upregulation of *CD44* in si-RALY HeLa cells, and this behavior was validated measuring the amount of three *CD44* constant exons by qRT-PCR and calculating the average expression (**Fig. 18D, upper panel**). Measuring the levels of different variable exons and normalizing them on the average expression of the whole gene, the alternative splicing pattern of *CD44* presented some significant differences in si-RALY compared to si-CTRL cells (**Fig. 18D, lower panel**).

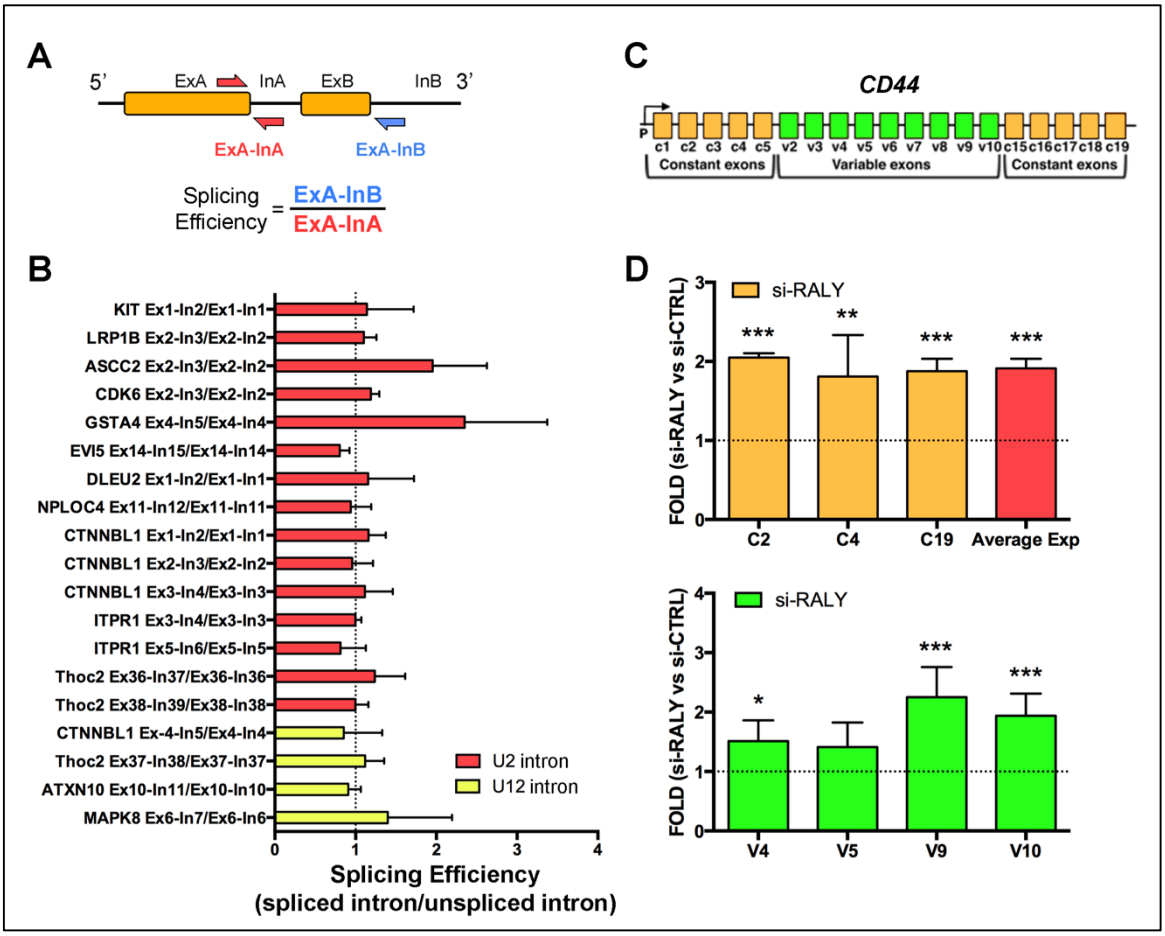

Figure 18. Analysis of the efficiency of splicing upon RALY silencing. (**A**) Representation of the measurement of splicing efficiency through qRT-PCR. (**B**) HeLa cells were transfected either with si-CTRL or si-RALY for 72 hours. Total RNA was extracted and analyzed by qRT-PCR, normalizing the signals on *GAPDH*. The ratio between the spliced intron and the unspliced intron signals in si-RALY cells was compared to si-CTRL cells, which were set to 1. The downregulation of RALY does not impair the splicing efficiency of U2 nor U12 introns. (**C**) Representation of the CD44 gene, taken and modified from Jimeno-Gonzalez *et al.*[178]. (**D**) HeLa cells were transfected either with si-CTRL or si-RALY for 72 hours. Total RNA was extracted and analyzed by qRT-PCR, normalizing the signals on

GAPDH. In the upper panel, the levels of three constant exons of *CD44* was analyzed and the average expression was calculated (Average Exp). The signals of si-CTRL cells were set to 1. In the lower panel, the levels of different variable exons of *CD44* were measured and normalized on the Average Exp of *CD44* in si-RALY cells. This allowed the quantization of alternatively spliced exons regardless of the absolute expression of the whole CD44 gene. The signals of si-CTRL cells were set to 1. The downregulation of RALY increases the levels of some alternatively spliced exons of CD44. All the graphs of the figure show the mean of five independent experiments ± S.D. P-value was calculated using an unpaired two tailed t-test (*P<0.05; **P<0.01; ***P<0.001) between the si-RALY values and the si-CTRL signals, which were set to 1.

Taken together, these results suggest that RALY does not have a prominent role in splicing, since its downregulation did not affect constitutive neither splicing efficiency nor timing in different targets. However, upon RALY downregulation, some effects were observed in the alternative splicing pattern of *CD44* mRNA. Further experiments with more sensitive techniques, as for example RNA-seq, could better define the involvement of RALY in alternative splicing, giving a broad overlook on global splicing events but also allowing a specific focus on RALY target transcripts.

As a last analysis of the relationship between RALY and splicing, a splicing specific analysis was performed on the microarray experiment previously shown in **Fig. 10**. In fact, the HTA2.0 microarray previously used for the gene expression analysis allowed also to analyze splicing. Briefly, the array used specific probe-sets targeting differently spliced exons to evaluate their abundance, and therefore the level of specific isoforms, inside the cells. The signal of every probe-set was normalized on the cellular abundance of the whole transcript so to avoid biases due to the upregulation or downregulation of genes or specific transcript isoforms upon RALY downregulation (less total gene X and less exon A = no significance, no changes in gene X and less exon A = differences in isoforms upon RALY downregulation). Compared to si-CTRL cells, in si-RALY cells the bioinformatics analysis detected as misregulated 847 cassette exon events, namely exon skipping or retention, 254 events caused by alternative 5' donor sites, 216 events caused by alternative 3' acceptor sites, and 65 intron retention events (**Fig. 19A and Table 4 in Appendix**). Functional annotation enrichment analysis by Gene Ontology and pathway databases identified cellular localization, protein localization to membrane and cell migration as biological processes significantly associated with the mis-spliced transcripts upon RALY downregulation (**Fig. 19B and Table 4 in Appendix**).

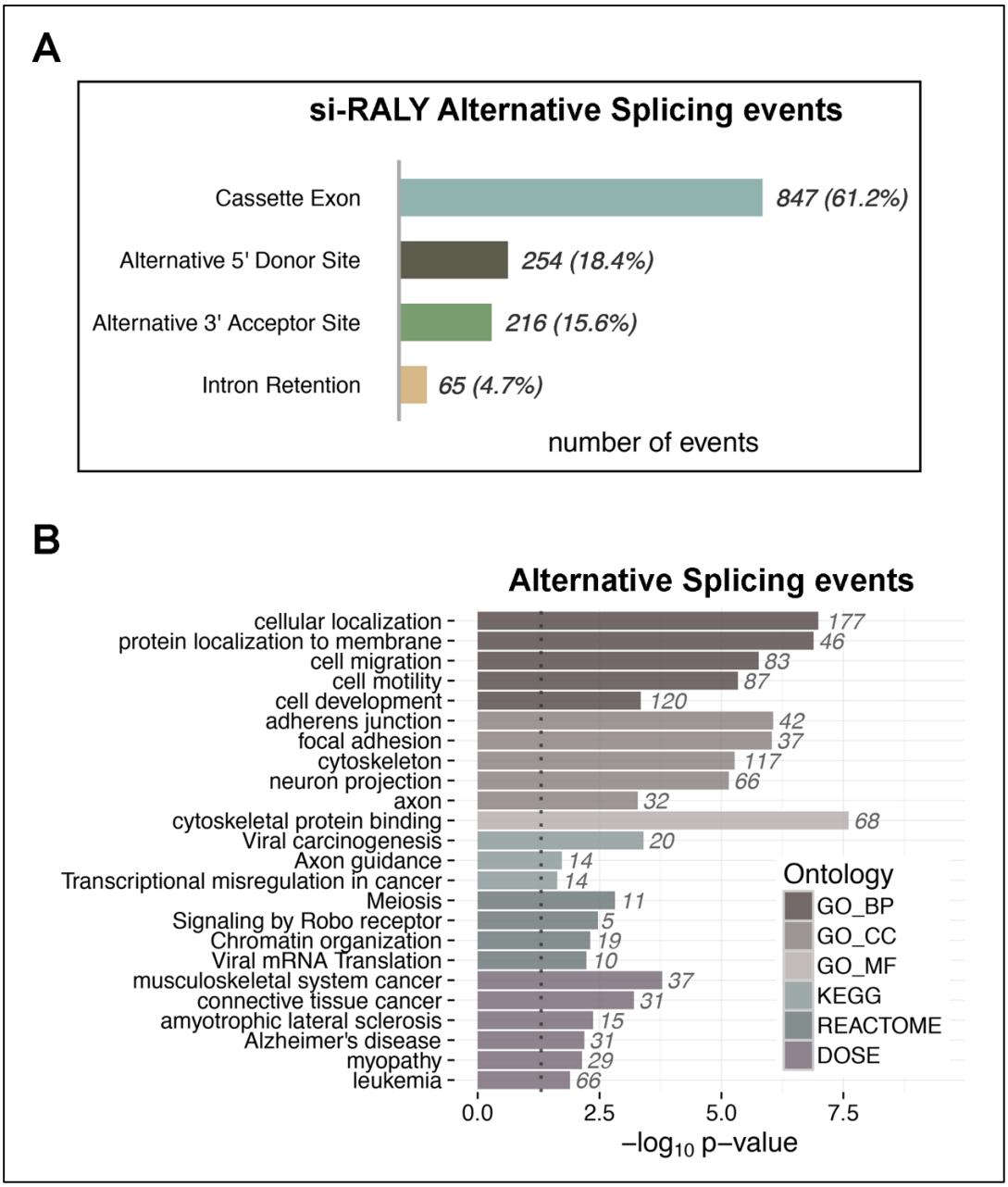

Figure 19. Microarray analysis of splicing upon RALY downregulation. (**A**) Barplot representing the number of altered alternative splicing events detected by microarray analysis upon RALY silencing. (**B**) Functional annotation enrichment analysis of altered alternative splicing events upon RALY downregulation. The bars display enriched classes from Gene Ontology terms and KEGG or REACTOME pathways. The number of DEGs falling in each category is displayed on the right of each bar.

To be noted, none of the represented Gene Ontology classes for the alternative splicing analysis was detected as enriched in the transcriptome analysis. This is due to the fact that the splicing specific signal of the microarray is normalized on the signal of the entire transcript. For this reason, the detection of abnormal alternative splicing events for highly misregulated genes becomes very stringent.

These experiments still need to be validated, but suggest the existence of a relationship between RALY and constitutive/alternative splicing. RALY was shown to interact with different RBPs involved in the co-transcriptional processing of RNAs, such as U2AF2, MATR3, EIF4A3, PRP19 and SR proteins [4,17,122,179–183]. Future experiments with more informative techniques will better understand if RALY is involved in splicing.

6.2 MaDEleNA

During my PhD I contributed to project MaDEleNA (Developing and Studying novel intelligent nanoMaterials and Devices towards Adaptive Electronics and Neuroscience Applications), which aimed to advancements in the field of adaptive electronics.

Neuronal cells can organize and form highly articulated networks, constituting one of the most complex and interesting biological systems in nature. The extraordinary degree of interaction achieved by neuronal nets allows neurons to form "computing units" with a large parallel processing power able to greatly overcome the capacity of modern computing devices [184]. To overtake this limit, new architectures are being investigated. A synapse is essentially a two-terminal device and bears striking functional resemblance to an electrical device termed memristor (memory + resistor) [185]. Neurons modify their conductive and resistive capacity depending on their previous "stimulatory experience", allowing the information to proceed only when the received stimuli exceed the voltage threshold of the axon hillock. The memristor's electrical resistance is not constant but depends on the currents that have previously flowed through the device. After a current input, the memristor retains an electrical charge, establishing a dynamic threshold that will have to be surpassed by the next input to make the output being transmitted [186]. Processors built with memristors will vary their physical characteristics according to their "experience", progressively developing features such as learning and decision making at the hardware level (at present, computer learning is essentially implemented in the software level). As a future application, devices with these characteristics could potentially be able to "communicate" with neurons themselves, sensing their electrical activity and adapting to it. The application of these technology could be endless, from the modelling of artificial neural networks to the construction of novel surgical devices.
Given the dynamicity of memristors being intrinsic in their constitutive materials and given the rigid environmental and culturing conditions required by neurons, our main aim have been to test the bio- and neuro-compatibility of different substrates with memristive electrical properties. In particular, we analyzed the bio-permissive properties of Polyaniline (PANI) and TiO_2, the latter in different Oxygen stoichiometric ratios. The results are presented in the Publications section.

7. EXPERIMENTAL PROCEDURES

Cell cultures and transient transfections.

HeLa cells and MCF7 were grown in DMEM supplemented with 10% FBS as previously described [109]. The transfections of the plasmids were performed using the TransIT transfection reagent (Mirus, Bio LLC) according to the manufacturer's protocols. The transfections of ON-target plus SMART-pool siRNAs and control siRNAs (Thermo Fisher) were performed with the INTERFERin transfection reagent (Polyplus-Transfection), according to the manufacturer's protocols.

Quantitative real-time PCR.

Total RNA was purified from cells using the TRIzol Reagent (Thermo Fischer) and retro-transcribed using the RevertAid First Strand cDNA Synthesis Kit (Thermo Fisher). The KAPA SYBR FAST qPCR Kit (KAPA Biosystems) was used for the qRT-PCR. All primers were purchased by IDT (TEMA Ricerca) and are listed in the **Appendix, Table 5**. The samples were incubated in the BioRad CFX96 Thermo Cycler for 40 cycles and the results were analysed with the Bio-Rad CFX Manager version 2.1. The relative expression was calculated according to the $2^{-\Delta\Delta Ct}$ method.

Constructs.

RT-PCR was performed on total RNA isolated from HeLa cells using the TRIzol Reagent. Human RALY cDNA was amplified with the Phusion High-Fidelity DNA polymerase (New England BioLabs), and then cloned in frame with the c-Myc/DDK tag (pCMV6-Entry Vector, Origene). The sequences of the primers used in this study are listed in the **Appendix, Table 5**.

Preparation of cell extracts and Western blot analysis.

The cells were washed with pre-warmed PBS, lysed in lysis buffer [0.1% Triton X-100, 0.5% NP-40, 150 mM NaCl, 20 mM Tris-HCl pH 8.0, plus protease inhibitor mixture (Roche), including 1 mM phenylmethylsulfonyl fluoride (PMSF)]. Equal amounts of proteins were separated on 12% SDS-PAGE and blotted onto nitrocellulose (GE Healthcare). Western blots were probed with the following primary antibodies: rabbit

polyclonal anti-RALY (A302-070A, Bethyl Laboratories); mouse monoclonal anti-c-Myc (M4439, Sigma Aldrich), rabbit polyclonal anti-H1X (ab31972, Abcam), mouse monoclonal anti-Beta Tubulin (sc-53140, Santa Cruz Biotechnology), rabbit polyclonal anti-H3 (ab1791, Abcam), mouse monoclonal anti-E2F1 (sc-251, Santa Cruz Biotechnology), mouse monoclonal anti-Ubiquitin (3936, Cell Signaling Technology), mouse monoclonal anti-GAPDH (sc-32233, Santa Cruz Biotechnology), rabbit polyclonal anti-FUS/TLS (ab2349, Abcam). Horseradish peroxidase (HRP)-conjugated goat anti-mouse and anti-rabbit antibodies (Santa Cruz Biotechnology) were used as secondary antibodies.

Cell Fractionation in Soluble and Chromatin-bound protein fractions.

HeLa cells were resuspended in radioimmunoprecipitation assay (RIPA) buffer with protease inhibitors and was homogenized with a 23G needle. After 20 min incubation on ice, cell lysates were centrifuged at 1000 × g for 10 min at 4 °C. The supernatants were collected (soluble fraction) and the pellet was resuspended with RIPA buffer supplemented with 0.3% SDS and 250 units/mL benzonase. Pellet suspensions were incubated on ice for 10 min and centrifuged again at 1000 × g for 10 min at 4 °C. The supernatant from the second centrifugation contained most of the chromatin-bound proteins.

Micrococcal Nuclease Cell Fractionation.

Cell fractionation was performed as described by Yang *et al*. [100]. Briefly, HeLa cells were washed with pre-warmed PBS and incubated for 10 minutes on ice in Nucleus Separation Buffer (NSB) [10 mM KCl, 1.5 mM $MgCl_2$, 0.34 M Sucrose, 10% Glycerol, 1 mM DTT, 0.1% Triton X100] supplemented with Protease Inhibitors (Roche) and RNase Inhibitors (NEB) depending on the experiment. For RNase treatment, RNase A (100 μg/ml) was added to NSB and after the 10 minutes on ice, the extract was incubated at 37°C for 10 minutes. The samples were then centrifuged at 1400 x g at 4°C for 10 minutes to pellet the nuclei. The supernatant (cytosolic fraction) was collected and the nuclei pellet was resuspended in NSB supplemented with 1 mM $CaCl_2$ and 2000 U/ml Micrococcal Nuclease (MNase) (NEB), and incubated at 37°C for 10 minutes. EGTA was afterwards added to arrive at a 2 mM concentration to stop MNase reaction. The

samples were centrifuged at 1400 x g at 4°C for 10 minutes to pellet chromatin. The supernatant (Nuclear Soluble fraction) was collected and chromatin was resuspended in NSB supplemented with 150 mM NaCl. The tubes were left in rotation at 4°C for 2 hours. The samples were then centrifuged at 1400 x g at 4°C for 10 minutes to pellet chromatin again. The supernatant was collected (Low-Salt Soluble fraction, transcriptionally active chromatin) and the remaining pellet was dissolved in NSB supplemented with 600 mM NaCl. The samples were left in rotation at 4°C overnight. The tubes were finally centrifuged at 1400 x g at 4°C for 10 minutes to pellet the insoluble fraction of chromatin and the last supernatant was collect (High-Salt Soluble fraction, transcriptionally inactive chromatin). After the supernatant collections, every fraction was clarified with a centrifugation at 16400 x g at 4°C for 2 minutes.

RNA synthesis assay.

HeLa cells were grown for 24 hours to obtain a 70-80% of confluence. The cells were incubated in complete DMEM supplemented with 0.5 mM 5-Ethynyluridine (5EU) (Jena Bioscience) for the desired amount of time. When required, Actinomycin D was added at the concentration of 125 ng/ml at the same moment. The cells were then fixed for 30 minutes at room temperature in fixing solution [125 mM Pipes pH 6.8, 10 mM EGTA, 1 mM $MgCl_2$, 0.2% Triton X-100, 3.7% formaldehyde]. Click reaction with 5FAM-Azide (5FAM) (Jena Bioscience) was performed in staining solution [100 mM Tris pH 8.5, 1 mM $CuSO_4$, 10 µM 5FAM, 100 mM ascorbic acid (in water)]. The cells were washed three times in TBS supplemented with 0.5% Triton X-100. Staining of proteins was achieved incubating the cells in blocking solution and following the immunocytochemistry protocol as described below. Further details in Jao *et al*. [124].

Immunocytochemistry and fluorescence microscopy.

Cells were washed in pre-warmed PBS and then fixed in 4% PFA for 15 minutes at room temperature. Immunocytochemistry was carried out as previously described [109]. The following primary antibodies were used: rabbit polyclonal anti-RALY (A302-070A, Bethyl Laboratories); mouse monoclonal anti-c-Myc (M4439, Sigma Aldrich). Alexa 594- and Alexa 488-coupled goat anti-mouse and anti-rabbit IgG were used as secondary antibodies (Santa Cruz Biotechnology). Microscopy analysis was performed using the

Zeiss Observer Z.1 microscope implemented with the Zeiss ApoTome module. Pictures were acquired using Zen Blue imaging software package (Zeiss) and assembled with Adobe Photoshop CS6. Images were not modified other than adjustments of levels, brightness and magnification.

RNA immunoprecipitation (RIP).

RIP was performed as reported by Keene and colleagues with some modifications [187]. In brief, $4x10^7$ cells were grown in 15 cm-culture dishes and were irradiated once with 150 mJ/cm^2 at 254 nm using an UVLink UV-crosslinker (Uvitec Cambridge). The cells were lysed in lysis buffer [10 mM HEPES pH 7.4, 100 mM KCl, 5 mM $MgCl_2$, 0.5% NP40, 1mM DTT, 100 U/ml RNase Out and Protease inhibitor cocktail (Roche)] for 3 hours at – 80°C and centrifuged at 10,000 x g for 20 min at 4° C. Then, the supernatants were pre-cleared with 20 µl of A/G magnetic beads for 1hour at 4°C. Successively, the supernatants were incubated for 4 hours at 4°C with protein A/G magnetic beads coated either with antibody anti-RALY (A302-069A, Bethyl Laboratories) (3 µg) or with normal rabbit IgG polyclonal antibody (Millipore). The beads were then washed three times with NT2 buffer [50 mM Tris-HCl pH 7.5, 300 mM NaCl, 1 mM $MgCl_2$, 0.05% NP-40, 1 % Urea]. Each wash was performed on a rotating wheel, for 10 minutes, at 4°C. The RNA was isolated with TRIzol and processed for qRT-PCR analysis as described below. The primers used for the qRT-PCR analysis are shown in the **Appendix, Table 5**.

RNA pull-down.

RNA pull-down was performed as previously described [109]. Briefly, wild-type or mutant probes (50 pmol) were incubated with 30 µl of streptavidin-coupled Dynabeads (ThermoFisher) for 20 min at room temperature in RNA Capture Buffer [20 mM Tris (pH 7.5), 1 M NaCl, 1 mM EDTA 20 mM Tris (pH 7.5)]. HeLa cells were washed with PBS and then lysed with the lysis buffer [20 mM Tris (pH 7.5), 50 mM NaCl, 2 mM MgCl2, 0.1% Tween-20]. The lysate (300 µg) was incubated with biotinylated RNA probes coupled to streptavidin Dynabeads for 1 hour at 4°C under rotation. Dynabeads were then washed three times with washing solution [20 mM Tris (pH 7.5), 15 mM NaCl, 0.1% Tween-20], solubilized in Laemmli reducing buffer and boiled for Western blot analysis.

RNA Polymerase II Elongation Rate and Splicing Speed Analyses.

RNA Polymerase II elongation was monitored according to Singh and Padgett [134]. Briefly, HeLa cells were grown in 6-well plates till 70% of confluence and were successively transfected with either si-CTRL or si-RALY siRNAs. After 72 hours, the cells were incubated for 3 hours in complete DMEM + 100 µM DRB to block RNAPII-dependent transcription. The cells were successively washed twice with ice-cold PBS and incubated with DMEM at 37 °C to re-activate transcription. Cells were afterwards lysed at 5 minute intervals using the lysis buffer of the High Pure Isolation Kit (Roche) and total RNA was extracted. Reverse transcription was performed using the ImProm-II Reverse Transcription System (Promega). The analysis of pre-mRNA levels was done by quantitative Real-Time PCR using the primers listed in the **Appendix, Table 5**. For the analysis, the qRT-PCR results of transfected cells (si-CTRL or si-RALY) not treated with DRB were used as reference values for the normalization of the transfected cells treated with DRB for 3 hours and allowed afterwards to recover transcription.

Real-Time Cell Analysis.

The cell proliferation assay with the xCELLigence system (Roche) was performed according to the manufacturer. Briefly, $5x10^4$ cells were seeded into each well of a specific xCELLigence E-plate. Cell proliferation was monitored by the instrument through the measurement of the impedance of the bottom of the wells for three days with detections every 15 minutes. The proliferation signals were normalized on the signals acquired at 10 hours post seeding, so to normalize on the number of effectively proliferating cells in each well.

Cell Cycle Distribution Analysis.

Cell Cycle distribution was analyzed as described by Massey and colleagues [156]. Briefly, 10^4 cells were seeded into a 96 well plate and grown for 24 hours. Cells were successively incubated in DMEM plus 10 µM 5-Ethynyl-deoxiuridine (5EdU) and successively treated for Click-it reaction as described above. Processed cells were successively analyzed with the High Content Imaging System Operetta as described by Massey and colleagues [156].

High content analysis.

HeLa cells were plated ($1x10^4$ cells/well) in 96-well plates (Corning) and transfected the day after with either si-CTRL or si-RALY siRNAs for 72 hours. Then the cells were incubated for different amounts of time with 5EU 0.5 mM in warm DMEM to mark newly synthesized RNA and stained with 5FAM through Click reaction (see the RNA synthesis assay paragraph); successively, they were processed for immunocytochemistry, starting from the blocking step (see above). Plates were imaged on the High Content Imaging System Operetta™ (PerkinElmer). In each of the 4 wells for every condition, images were acquired in 5 preselected fields with LWD 20x objective over three channels, with λ= 380 nm excitation/λ= 445 nm emission for DAPI, λ= 495 nm excitation/λ= 519 nm emission for Alexa Fluor 488 or 5FAM, with λ= 535 nm excitation/λ= 615 nm emission for Alexa Fluor 594. For feature extraction, the images were analyzed by Harmony software version 4.1 (PerkinElmer). Briefly, individual cell nuclei were segmented based on DAPI staining. For the Click-it experiment (Fig. 4) the Select Population algorithm allowed to identify the sub-population of silenced cells for RALY by setting a fluorescence intensity threshold on the Alexa Fluor 594 signal. The 5FAM signal was then measured only in the cells evaluated as silenced for RALY.

Microarray Analysis.

HeLa cells were transfected for 72 hours with either si-CTRL or si-RALY siRNAs as described above, and total RNA was extracted using the High Pure Isolation Kit (Roche). Samples were prepared for hybridization following the Affymetrix WT PLUS Reagent Kit protocol (http://www.affymetrix.com/estore/catalog/prod770005/AFFY/WT+PLUS+Reagent+Kit#1_1). The amplified and labeled samples were hybridized to the Affymetrix GeneChip Human Transcriptome Array 2.0 (Affymetrix, Santa Clara, CA, USA). Arrays were washed and stained using the Affymetrix Fluidics Station 450, and scanned in the Affymetrix GeneArray 3000 7G scanner. The experiment was performed in biological triplicate.

The CEL files resulting from the GeneChip analysis were analysed with affymetrix Transcriptome Analysis Console (TAC) 3.1 and the Bioconductor library of biostatistical packages (http://www.bioconductor.org/). Raw data were preprocessed and normalized with the Robust Multichip Analysis (RMA) method. Gene average intensities were

determined with the Tukey's Bi-weight Average Algorithm, using Affymetrix default analysis settings. Differentially expressed genes (DEGs) upon RALY silencing were determined adopting a double threshold based on statistical significance (unpaired One-Way ANOVA FDR < 0.05) and log2 fold change (> 0.5 for up-regulated, < -0.5 for down-regulated genes).

Frequencies of RALY RIP-seq targets among DEGs were tested with the "Test of Equal or Given Proportions" implemented in R. The ClusterProfiler package was used for enrichment analysis of DEGs lists, using annotations from Gene Ontology (http://www.thegeneontology.org), KEGG (http://www.genome.jp/kegg/), REACTOME (http://www.reactome.org/) databases. The significance of overrepresentation was determined using a FDR threshold of 0.05. Gene Set Enrichment Analysis (GSEA) was performed on ranked DEGs against the the Molecular Signatures Database (MSigDB) "Hallmark" collection of annotated gene sets (v 5.2) [188].

Alternative Splicing analysis was performed with Affymetrix Transcriptome Analysis Console (TAC) 3.1, using default parameters. Significant alternative splicing events upon RALY silencing were filtered using the following criteria: absolute Splicing Index > 2, absolute Splicing Index > 1.75 * absolute gene fold change, splicing One-Way ANOVA p-value < 0.05, Splicing Event Score > 0.1.

8. REFERENCES

1. Sainsbury, S., Bernecky, C. & Cramer, P. Structural basis of transcription initiation by RNA polymerase II. *Nat. Rev. Mol. Cell Biol.* **16,** 129–143 (2015).
2. Thomas, M. C. & Chiang, C.-M. The General Transcription Machinery and General Cofactors. *Crit. Rev. Biochem. Mol. Biol.* **41,** 105–178 (2006).
3. Fredericks, A. M., Cygan, K. J., Brown, B. A. & Fairbrother, W. G. RNA-Binding Proteins: Splicing Factors and Disease. *Biomolecules* **5,** 893–909 (2015).
4. Das, R. *et al.* SR Proteins Function in Coupling RNAP II Transcription to Pre-mRNA Splicing. *Mol. Cell* **26,** 867–881 (2007).
5. Bird, G., Zorio, D. A. R. & Bentley, D. L. RNA polymerase II carboxy-terminal domain phosphorylation is required for cotranscriptional pre-mRNA splicing and 3'-end formation. *Mol. Cell. Biol.* **24,** 8963–9 (2004).
6. David, C. J., Boyne, A. R., Millhouse, S. R. & Manley, J. L. The RNA polymerase II C-terminal domain promotes splicing activation through recruitment of a U2AF65-Prp19 complex. *Genes Dev.* **25,** 972–983 (2011).
7. Gu, B., Eick, D. & Bensaude, O. CTD serine-2 plays a critical role in splicing and termination factor recruitment to RNA polymerase II in vivo. *Nucleic Acids Res.* **41,** 1591–603 (2013).
8. Brody, Y. *et al.* The in vivo kinetics of RNA polymerase II elongation during co-transcriptional splicing. *PLoS Biol.* **9,** (2011).
9. Ramanathan, A., Robb, G. B. & Chan, S.-H. mRNA capping: biological functions and applications. *Nucleic Acids Res.* (2016). doi:10.1093/nar/gkw551
10. Licht, K., Kapoor, U., Mayrhofer, E. & Jantsch, M. F. Adenosine to Inosine editing frequency controlled by splicing efficiency. *Nucleic Acids Res.* (2016). doi:10.1093/nar/gkw325
11. Gilbert, W. V *et al.* Messenger RNA modifications: Form, distribution, and function. *Science* **352,** 1408–12 (2016).
12. Gray, N. K. *et al.* Poly(A)-binding proteins and mRNA localization: who rules the roost? *Biochem. Soc. Trans.* **43,** 1277–84 (2015).
13. Licht, K. & Jantsch, M. F. Rapid and dynamic transcriptome regulation by RNA

editing and RNA modifications. *J. Cell Biol.* **213,** 15–22 (2016).

14. Bowman, E. A. & Kelly, W. G. RNA polymerase II transcription elongation and Pol II CTD Ser2 phosphorylation: A tail of two kinases. *Nucleus* **5,** 224–36
15. Yamaguchi, Y., Inukai, N., Narita, T., Wada, T. & Handa, H. Evidence that negative elongation factor represses transcription elongation through binding to a DRB sensitivity-inducing factor/RNA polymerase II complex and RNA. *Mol. Cell. Biol.* **22,** 2918–27 (2002).
16. Gilchrist, D. A. D. A. *et al.* Pausing of RNA polymerase II disrupts DNA-specified nucleosome organization to enable precise gene regulation. *Cell* **143,** 540–51 (2010).
17. Jonkers, I. & Lis, J. T. Getting up to speed with transcription elongation by RNA polymerase II. *Nat. Rev. Mol. Cell Biol.* **16,** 167–177 (2015).
18. Gilchrist, D. A. *et al.* NELF-mediated stalling of Pol II can enhance gene expression by blocking promoter-proximal nucleosome assembly. *Genes Dev.* **22,** 1921–1933 (2008).
19. Lu, X. *et al.* Multiple P-TEFbs cooperatively regulate the release of promoter-proximally paused RNA polymerase II. *Nucleic Acids Res.* gkw571 (2016). doi:10.1093/nar/gkw571
20. Hsin, JHsin, J.-P., & Manley, J. L. (2012). The RNA polymerase II CTD coordinates transcription and RNA processing. Genes & Development, 26(19), 2119–37. http://doi.org/10.1101/gad.200303.112ing-Ping & Manley, J. L. The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes Dev.* **26,** 2119–37 (2012).
21. Price, D. H. P-TEFb, a cyclin-dependent kinase controlling elongation by RNA polymerase II. *Mol. Cell. Biol.* **20,** 2629–2634 (2000).
22. Ramakrishnan, R., Yu, W. & Rice, A. P. Limited redundancy in genes regulated by Cyclin T2 and Cyclin T1 Limited redundancy in genes regulated by Cyclin T2 and Cyclin T1. *BMC Res. Notes* **4,** 260 (2011).
23. Bhatt, D. M. *et al.* Transcript dynamics of proinflammatory genes revealed by sequence analysis of subcellular RNA fractions. *Cell* **150,** 279–90 (2012).
24. Dujardin, G. *et al.* How Slow RNA Polymerase II Elongation Favors Alternative Exon Skipping. *Mol. Cell* **54,** 683–690 (2014).

25. Fong, N. *et al.* Pre-mRNA splicing is facilitated by an optimal RNA polymerase II elongation rate. *Genes Dev.* **28,** 2663–2676 (2014).
26. Shukla, S. & Oberdoerffer, S. Co-transcriptional regulation of alternative pre-mRNA splicing. *Biochim. Biophys. Acta* **1819,** 673–83 (2012).
27. Veloso, A. *et al.* Rate of elongation by RNA polymerase II is associated with specific gene features and epigenetic modifications. *Genome Res.* **24,** 896–905 (2014).
28. Close, P. *et al.* DBIRD complex integrates alternative mRNA splicing with RNA polymerase II transcript elongation. *Nature* **484,** 386–389 (2012).
29. Wahl, M. C., Will, C. L. & Lührmann, R. The Spliceosome: Design Principles of a Dynamic RNP Machine. *Cell* **136,** 701–718 (2009).
30. Saldi, T., Cortazar, M. A., Sheridan, R. M. & Bentley, D. L. Coupling of RNA Polymerase II Transcription Elongation with Pre-mRNA Splicing. *J. Mol. Biol.* **428,** 2623–2635 (2016).
31. Turunen, J. J., Niemelä, E. H., Verma, B. & Frilander, M. J. The significant other: Splicing by the minor spliceosome. *Wiley Interdisciplinary Reviews: RNA* **4,** 61–76 (2013).
32. Proudfoot, N. J. How RNA polymerase II terminates transcription in higher eukaryotes. *Trends Biochem. Sci.* **14,** 105–110 (1989).
33. Proudfoot, N. New perspectives on connecting messenger RNA 3' end formation to transcription. *Curr. Opin. Cell Biol.* **16,** 272–278 (2004).
34. Connelly, S. & Manley, J. L. A functional mRNA polyadenylation signal is required for transcription termination by RNA polymerase II. *Genes Dev.* **2,** 440–52 (1988).
35. Richard, P. & Manley, J. L. Transcription termination by nuclear RNA polymerases. *Genes Dev.* **23,** 1247–1269 (2009).
36. Gerstberger, S., Hafner, M. & Tuschl, T. A census of human RNA-binding proteins. *Nat. Rev. Genet.* **15,** 829–845 (2014).
37. Dreyfuss, G., Kim, V. N., Kataoka, N. & Medical, H. H. Messenger-Rna-Binding Proteins and the Messages They Carry. *Nat. Rev. Mol. Cell Biol.* **3,** 195–205 (2002).
38. Müller-McNicoll, M. & Neugebauer, K. M. How cells get the message: dynamic assembly and function of mRNA-protein complexes. *Nat. Rev. Genet.* **14 VN-r,** 275–287 (2013).

39. Lunde, B. M., Moore, C. & Varani, G. RNA-binding proteins: modular design for efficient function. *Nat Rev Mol Cell Biol* **8,** 479–490 (2007).

40. van Der Houven Van Oordt, W., Newton, K., Screaton, G. R. & Cáceres, J. F. Role of SR protein modular domains in alternative splicing specificity in vivo. *Nucleic Acids Res.* **28,** 4822–31 (2000).

41. Masliah, G., Barraud, P. & Allain, F. H. T. RNA recognition by double-stranded RNA binding domains: A matter of shape and sequence. *Cellular and Molecular Life Sciences* **70,** 1875–1895 (2013).

42. Glisovic, T., Bachorik, J. L., Yong, J. & Dreyfuss, G. RNA-binding proteins and post-transcriptional gene regulation. *FEBS Letters* **582,** 1977–1986 (2008).

43. Oubridge, C., Ito, N., Evans, P. R., Teo, C. H. & Nagai, K. Crystal structure at 1.92 A resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin. *Nature* **372,** 432–8 (1994).

44. Finn, R. D. *et al.* Pfam: clans, web tools and services. *Nucleic Acids Res.* **34,** D247–D251 (2006).

45. Auweter, S. D., Oberstrass, F. C. & Allain, F. H. T. Sequence-specific binding of single-stranded RNA: Is there a code for recognition? *Nucleic Acids Res.* **34,** 4943–4959 (2006).

46. Maris, C., Dominguez, C. & Allain, F. H. T. The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS Journal* **272,** 2118–2131 (2005).

47. Cléry, A. & Frédéric, H.-T. A. From Structure to Function of RNA Binding Domains. *Landes Biosci. Austin* 137–58 (2012).

48. Dreyfuss, G., Matunis, M. J., Pinol-Roma, S. & Burd, C. G. hnRNP Proteins and the Biogenesis of mRNA. *Annu. Rev. Biochem.* **62,** 289–321 (1993).

49. Martinez-Contreras, R. *et al.* in *Advances in experimental medicine and biology* **623,** 123–147 (2007).

50. Beyer, A. L., Christensen, M. E., Walker, B. W. & LeStourgeon, W. M. Identification and characterization of the packaging proteins of core 40S hnRNP particles. *Cell* **11,** 127–138 (1977).

51. Krecic, A. M. & Swanson, M. S. hnRNP complexes: Composition, structure, and function. *Curr. Opin. Cell Biol.* **11,** 363–371 (1999).

52. Ford, L. P., Wright, W. E. & Shay, J. W. A model for heterogeneous nuclear ribonucleoproteins in telomere and telomerase regulation. *Oncogene* **21,** 580–3 (2002).

53. Huang, Y.-S. & Richter, J. D. Regulation of local mRNA translation. *Curr. Opin. Cell Biol.* **16,** 308–313 (2004).

54. Norton, P. A. Alternative pre-mRNA splicing: factors involved in splice site selection. *J. Cell Sci.* **107,** 1–7 (1994).

55. Swinburne, I. A., Meyer, C. A., Liu, X. S., Silver, P. A. & Brodsky, A. S. Genomic localization of RNA binding proteins reveals links between pre-mRNA processing and transcription. *Genome Res.* **16,** 912–21 (2006).

56. Rajagopalan, L. E., Westmark, C. J., Jarzembowski, J. A. & Malter, J. S. hnRNP C increases amyloid precursor protein (APP) production by stabilizing APP mRNA. *Nucleic Acids Res.* **26,** 3418–3423 (1998).

57. Cirillo, D. *et al.* Constitutive patterns of gene expression regulated by RNA-binding proteins. *Genome Biol.* **15,** R13 (2014).

58. Han, S. P., Tang, Y. H. & Smith, R. Functional diversity of the hnRNPs: past, present and perspectives. *Biochem. J.* **430,** (2010).

59. Chaudhury, A., Chander, P. & Howe, P. H. Heterogeneous nuclear ribonucleoproteins (hnRNPs) in cellular processes: Focus on hnRNP E1's multifunctional regulatory roles. *RNA* **16,** 1449–62 (2010).

60. Singh, R. & Valcárcel, J. Building specificity with nonspecific RNA-binding proteins. *Nat. Struct. Mol. Biol.* **12,** 645–653 (2005).

61. Yang, X. *et al.* The A1 and A1B proteins of heterogeneous nuclear ribonucleoparticles modulate 5' splice site selection in vivo. *Proc. Natl. Acad. Sci. U. S. A.* **91,** 6924–6928 (1994).

62. Mayeda, A. & Krainer, A. R. Regulation of alternative pre-mRNA splicing by hnRNP A1 and splicing factor SF2. *Cell* **68,** 365–375 (1992).

63. Mayeda, A., Helfman, D. M. & Krainer, A. R. Modulation of exon skipping and inclusion by heterogeneous nuclear ribonucleoprotein A1 and pre-mRNA splicing factor SF2/ASF. *Mol. Cell. Biol.* **13,** 2993–3001 (1993).

64. Lemieux, B. *et al.* A Function for the hnRNP A1/A2 Proteins in Transcription Elongation. *PLoS One* **10,** e0126654 (2015).

65. Jean-Philippe, J., Paz, S. & Caputi, M. hnRNP A1: The Swiss Army Knife of gene expression. *International Journal of Molecular Sciences* **14,** 18999–19024 (2013).

66. McAfee, J. G., Shahied-Milam, L., Soltaninassab, S. R. & LeStourgeon, W. M. A major determinant of hnRNP C protein binding to RNA is a novel bZIP-like RNA binding domain. *RNA* **2,** 1139–52 (1996).

67. Tan, J. H. *et al.* The bZIP-like motif of hnRNP C directs the nuclear accumulation of pre-mRNA and lethality in yeast. *J. Mol. Biol.* **305,** 829–38 (2001).

68. Zarnack, K. *et al.* Direct competition between hnRNP C and U2AF65 protects the transcriptome from the exonization of Alu elements. *Cell* **152,** 453–466 (2013).

69. FACKELMAYER, F. O., DAHM, K., RENZ, A., RAMSPERGER, U. & RICHTER, A. Nucleic acid binding properties of hnRNP-U/SAF-A, a nuclear matrix protein which binds DNA and RNA in vivo and in vitro. *Eur. J. Biochem.* **221,** 749–757 (1994).

70. Ye, J. *et al.* hnRNP U protein is required for normal pre-mRNA splicing and postnatal heart development and function. *Proc. Natl. Acad. Sci.* **112,** 201508461 (2015).

71. Kim, M. K. & Nikodem, V. M. hnRNP U Inhibits Carboxy-Terminal Domain Phosphorylation by TFIIH and Represses RNA Polymerase II Elongation. *Mol. Cell. Biol.* **19,** 6833–6844 (1999).

72. Hasegawa, Y. *et al.* The matrix protein hnRNP U is required for chromosomal localization of Xist RNA. *Dev. Cell* **19,** 469–76 (2010).

73. Dejgaard, K. & Leffers, H. Characterisation of the nucleic-acid-binding activity of KH domains - Different properties of different domains. *Eur. J. Biochem.* **241,** 425–431 (1996).

74. Bomsztyk, K., Denisenko, O. & Ostrowski, J. hnRNP K: One protein multiple processes. *BioEssays* **26,** 629–638 (2004).

75. Mikula, M., Bomsztyk, K., Goryca, K., Chojnowski, K. & Ostrowski, J. Heterogeneous nuclear ribonucleoprotein (HnRNP) K genome-wide binding survey reveals its role in regulating 3'-end RNA processing and transcription termination at the early growth response 1 (EGR1) gene through XRN2 exonuclease. *J. Biol. Chem.* **288,** 24788–98 (2013).

76. Castello, A., Fischer, B., Hentze, M. W. & Preiss, T. RNA-binding proteins in Mendelian disease. *Trends Genet.* **29,** 318–327 (2013).

77. Shankarling, G. & Lynch, K. W. Living or dying by RNA processing: Caspase expression in NSCLC. *Journal of Clinical Investigation* **120,** 3798–3801 (2010).

78. Geuens, T., Bouhy, D. & Timmerman, V. The hnRNP family: insights into their role in health and disease. *Human Genetics* **135,** 851–867 (2016).

79. Patry, C. *et al.* Small Interfering RNA-Mediated Reduction in Heterogeneous Nuclear Ribonucleoparticule A1/A2 Proteins Induces Apoptosis in Human Cancer Cells but not in Normal Mortal Cell Lines. *Cancer Res.* **63,** 7679–7688 (2003).

80. Liu, X., Zhou, Y., Lou, Y. & Zhong, H. Knockdown of HNRNPA1 inhibits lung adenocarcinoma cell proliferation through cell cycle arrest at G0/G1 phase. *Gene* **576,** 791–797 (2015).

81. Loh, T. J. *et al.* CD44 alternative splicing and hnRNP A1 expression are associated with the metastasis of breast cancer. *Oncol. Rep.* **34,** 1231–1238 (2015).

82. Anantha, R. W. *et al.* Requirement of Heterogeneous Nuclear Ribonucleoprotein C for BRCA Gene Expression and Homologous Recombination. *PLoS One* **8,** e61368 (2013).

83. Lu, J. & Gao, F.-H. Role and molecular mechanism of heterogeneous nuclear ribonucleoprotein K in tumor development and progression. *Biomed. reports* **4,** 657–663 (2016).

84. Gallardo, M. *et al.* HnRNP K Is a Haploinsufficient Tumor Suppressor that Regulates Proliferation and Differentiation Programs in Hematologic Malignancies. *Cancer Cell* **28,** 486–499 (2015).

85. Yeo, G., Holste, D., Kreiman, G. & Burge, C. B. Variation in alternative splicing across human tissues. *Genome Biol* **5,** R74 (2004).

86. Liu-Yesucevitz, L. *et al.* Local RNA Translation at the Synapse and in Disease. *J. Neurosci.* **31,** 16086–16093 (2011).

87. Holt, C. E. & Bullock, S. L. Subcellular mRNA Localization in Animal Cells and Why It Matters. *Science (80-.).* **1212,** 1212–6 (2010).

88. Tollervey, J. R. *et al.* Analysis of alternative splicing associated with aging and neurodegeneration in the human brain. *Genome Res.* **21,** 1572–1582 (2011).

89. Mills, J. D. & Janitz, M. Alternative splicing of mRNA in the molecular pathology of neurodegenerative diseases. *Neurobiology of Aging* **33,** 1012.e11-1012.e24 (2012).

90. Bekenstein, U. & Soreq, H. Heterogeneous nuclear ribonucleoprotein A1 in health and neurodegenerative disease: From structural insights to post-transcriptional regulatory roles. *Mol. Cell. Neurosci.* **56,** 436–446 (2013).

91. Berson, A. *et al.* Cholinergic-associated loss of hnRNP-A/B in Alzheimer's disease impairs cortical splicing and cognitive function in mice. *EMBO Mol. Med.* **4,** 730–742 (2012).

92. Lee, E. K. *et al.* hnRNP C promotes APP translation by competing with FMRP for APP mRNA recruitment to P bodies. *Nat. Struct. Mol. Biol.* **17,** 732–739 (2010).

93. Borreca, A., Gironi, K., Amadoro, G. & Ammassari-Teule, M. Opposite Dysregulation of Fragile-X Mental Retardation Protein and Heteronuclear Ribonucleoprotein C Protein Associates with Enhanced APP Translation in Alzheimer Disease. *Mol. Neurobiol.* **53,** 3227–3234 (2016).

94. Calvio, C., Neubauer, G., Mann, M. & Lamond, A. I. Identification of hnRNP P2 as TLS/FUS using electrospray mass spectrometry. *RNA* **1,** 724–33 (1995).

95. Sama, R. R. K., Ward, C. L. & Bosco, D. A. Functions of FUS/TLS From DNA Repair to Stress Response: Implications for ALS. *ASN Neuro* **6,** 175909141454447 (2014).

96. Yu, Y. & Reed, R. FUS functions in coupling transcription to splicing by mediating an interaction between RNAP II and U1 snRNP. *Proc. Natl. Acad. Sci.* **112,** 8608–8613 (2015).

97. Dhar, S. K. *et al.* FUsed in Sarcoma Is a Novel Regulator of Manganese Superoxide Dismutase Gene Transcription. *Antioxid. Redox Signal.* **0,** 1550–1566 (2013).

98. Schwartz, J. C. *et al.* FUS binds the CTD of RNA polymerase II and regulates its phosphorylation at Ser2. *Genes Dev.* **26,** 2690–5 (2012).

99. Masuda, A. *et al.* Position-specific binding of FUS to nascent RNA regulates mRNA length. *Genes Dev.* **29,** 1045–1057 (2015).

100. Yang, L., Gal, J., Chen, J. & Zhu, H. Self-assembled FUS binds active chromatin and regulates gene transcription. *Proc. Natl. Acad. Sci. U. S. A.* **111,** 17809–14 (2014).

101. Deng, H., Gao, K. & Jankovic, J. The role of FUS gene variants in neurodegenerative diseases. *Nat. Rev. Neurol.* **10,** 337–348 (2014).

102. Shang, Y. & Huang, E. J. Mechanisms of FUS mutations in familial amyotrophic lateral sclerosis. *Brain Res.* **1647,** 65–78 (2016).

103. Nadeau, N. J. *et al.* Characterization of Japanese quail yellow as a genomic

deletion upstream of the avian homolog of the mammalian ASIP (agouti) gene. *Genetics* **178,** 777–86 (2008).

104. Michaud, E. J., Bultman, S. J., Stubbs, L. J. & Woychik, R. P. The embryonic lethality of homozygous lethal yellow mice (Ay/Ay) is associated with the disruption of a novel RNA-binding protein. *Genes Dev.* **7,** 1203–13 (1993).
105. Tenzer, S. *et al.* Proteome-wide characterization of the RNA-binding protein RALY-interactome using the in vivo-biotinylation-pulldown-quant (iBioPQ) approach. *J. Proteome Res.* **12,** 2869–84 (2013).
106. Jiang, W., Guo, X. & Bhavanandan, V. P. Four distinct regions in the auxiliary domain of heterogeneous nuclear ribonucleoprotein C-related proteins. *Biochim. Biophys. Acta - Gene Struct. Expr.* **1399,** 229–233 (1998).
107. Busch, A. & Hertel, K. J. Evolution of SR protein and hnRNP splicing regulatory factors. *Wiley Interdisciplinary Reviews: RNA* **3,** 1–12 (2012).
108. Khrebtukova, I., Kuklin, A., Woychik, R. P. & Michaud, E. J. Alternative processing of the human and mouse raly genes(1). *Biochim. Biophys. Acta* **1447,** 107–12 (1999).
109. Rossi, A. *et al.* Identification and dynamic changes of RNAs isolated from RALY-containing ribonucleoprotein complexes. *Nucleic Acids Res.* (2017). doi:10.1093/nar/gkx235
110. Rhodes, G. H., Valbracht, J. R., Nguyen, M. D. & Vaughan, J. H. The p542 gene encodes an autoantigen that cross-reacts with EBNA-1 of the Epstein Barr virus and which may be a heterogeneous nuclear ribonucleoprotein. *J. Autoimmun.* **10,** 447–54 (1997).
111. Singh, G. *et al.* The cellular EJC interactome reveals higher-order mRNP structure and an EJC-SR protein nexus. *Cell* **151,** 750–64 (2012).
112. Jurica, M. S., Licklider, L. J., Gygi, S. R., Grigorieff, N. & Moore, M. J. Purification and characterization of native spliceosomes suitable for three-dimensional structural analysis. *RNA* **8,** 426–39 (2002).
113. Pineda, G. *et al.* Proteomics studies of the interactome of RNA polymerase II C-terminal repeated domain. *BMC Res. Notes* **8,** 616 (2015).
114. Sallam, T. *et al.* Feedback modulation of cholesterol metabolism by the lipid-responsive non-coding RNA LeXis. *Nature* **534,** 1–20 (2016).

115. Tsofack, S. P. *et al.* NONO and RALY proteins are required for YB-1 oxaliplatin induced resistance in colon adenocarcinoma cell lines. *Mol. Cancer* **10,** 145 (2011).

116. Raffetseder, U. *et al.* Splicing factor SRp30c interaction with Y-box protein-1 confers nuclear YB-1 shuttling and alternative splice site selection. *J. Biol. Chem.* **278,** 18241–8 (2003).

117. Gaudreault, I., Guay, D. & Lebel, M. YB-1 promotes strand separation in vitro of duplex DNA containing either mispaired bases or cisplatin modifications, exhibits endonucleolytic activities and binds several DNA repair proteins. *Nucleic Acids Res.* **32,** 316–27 (2004).

118. Chen, C. Y. *et al.* Nucleolin and YB-1 are required for JNK-mediated interleukin-2 mRNA stabilization during T-cell activation. *Genes Dev.* **14,** 1236–48 (2000).

119. McCloskey, a., Taniguchi, I., Shinmyozu, K. & Ohno, M. hnRNP C Tetramer Measures RNA Length to Classify RNA Polymerase II Transcripts for Export. *Science (80-.).* **335,** 1643–1646 (2012).

120. Yamaguchi, A. & Takanashi, K. FUS interacts with nuclear matrix-associated protein SAFB1 as well as Matrin3 to regulate splicing and ligand-mediated transcription. *Sci. Rep.* **6,** 35195 (2016).

121. Henikoff, S., Henikoff, J. G., Sakai, A., Loeb, G. B. & Ahmad, K. Genome-wide profiling of salt fractions maps physical properties of chromatin. *Genome Res.* **19,** 460–9 (2009).

122. Perales, R. & Bentley, D. "Cotranscriptionality": The Transcription Elongation Complex as a Nexus for Nuclear Transactions. *Mol. Cell* **36,** 178–191 (2009).

123. Sobell, H. M. Actinomycin and DNA transcription. *Proc. Natl. Acad. Sci. U. S. A.* **82,** 5328–31 (1985).

124. Jao, C. Y. & Salic, A. Exploring RNA transcription and turnover in vivo by using click chemistry. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 15779–84 (2008).

125. Tóth, K. F. *et al.* Trichostatin A-induced histone acetylation causes decondensation of interphase chromatin. *J. Cell Sci.* **117,** 4277–87 (2004).

126. Robinson, P. J. & Rhodes, D. Structure of the "30nm" chromatin fibre: A key role for the linker histone. *Curr. Opin. Struct. Biol.* **16,** 336–343 (2006).

127. Kalashnikova, A. A., Rogge, R. A., Hansen, J. C., Peper, J. S. & Dahl, R. E. Linker

histone H1 and protein-protein interactions. *Biochim. Biophys. Acta* **22,** 134–139 (2015).

128. Lee, H., Habas, R. & Abate-Shen, C. Msx1 Cooperates with Histone H1b for Inhibition of Transcription and Myogenesis. *Science (80-.).* **304,** 1675–1678 (2004).
129. Shen, X. & Gorovsky, M. A. Linker histone H1 regulates specific gene expression but not global transcription in vivo. *Cell* **86,** 475–83 (1996).
130. Hergeth, S. P. & Schneider, R. The H1 linker histones: multifunctional proteins beyond the nucleosomal core particle. *EMBO Rep.* **16,** 1439–1453 (2015).
131. Stoldt, S., Wenzel, D., Schulze, E., Doenecke, D. & Happel, N. G1 phase-dependent nucleolar accumulation of human histone H1x. *Biol. cell* **99,** 541–52 (2007).
132. Takata, H. *et al.* H1.X with different properties from other linker histones is required for mitotic progression. *FEBS Lett.* **581,** 3783–3788 (2007).
133. Mayor, R. *et al.* Genome Distribution of Replication-independent Histone H1 Variants Shows H1.0 Associated with Nucleolar Domains and H1X Associated with RNA Polymerase II-enriched Regions. *J. Biol. Chem.* **290,** 7474–7491 (2015).
134. Singh, J. & Padgett, R. A. Rates of in situ transcription and splicing in large human genes. *Nat. Struct. Mol. Biol.* **16,** 1128–33 (2009).
135. Yankulov, K., Yamashita, K., Roy, R., Egly, J.-M. & Bentley, D. L. The Transcriptional Elongation Inhibitor 5,6-Dichloro-1- -D-ribofuranosylbenzimidazole Inhibits Transcription Factor IIH-associated Protein Kinase. *J. Biol. Chem.* **270,** 23922–23925 (1995).
136. Meng, P. & Ghosh, R. Transcription addiction: can we garner the Yin and Yang functions of E2F1 for cancer therapy? *Cell Death Dis.* **5,** e1360 (2014).
137. Lindqvist, A. *et al.* Cyclin B1-Cdk1 activation continues after centrosome separation to control mitotic progression. *PLoS Biol.* **5,** 1127–1137 (2007).
138. Nam, H.-J. & van Deursen, J. M. Cyclin B2 and p53 control proper timing of centrosome separation. *Nat. Cell Biol.* **16,** 538–49 (2014).
139. Donzelli, M. & Draetta, G. F. Regulating mammalian checkpoints through Cdc25 inactivation. *EMBO Rep.* **4,** 671–7 (2003).
140. Zhou, H. *et al.* Uncleaved TFIIA is a substrate for taspase 1 and active in transcription. *Mol. Cell. Biol.* **26,** 2728–35 (2006).

141. Peterson, M. G. *et al.* Structure and functional properties of human general transcription factor IIE. *Nature* **354,** 369–373 (1991).
142. Luo, Z. *et al.* The super elongation complex family of RNA polymerase II elongation factors: gene target specificity and transcriptional output. *Mol. Cell. Biol.* **32,** 2608–17 (2012).
143. Orphanides, G. *et al.* FACT, a factor that facilitates transcript elongation through nucleosomes. *Cell* **92,** 105–116 (1998).
144. Safina, A. *et al.* Complex mutual regulation of facilitates chromatin transcription (FACT) subunits on both mRNA and protein levels in human cells. *Cell Cycle* **12,** 2423–2434 (2013).
145. Wells, J., Graveel, C. R., Bartley, S. M., Madore, S. J. & Farnham, P. J. The identification of E2F1-specific target genes. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 3890–3895 (2002).
146. Attwooll, C., Lazzerini Denchi, E. & Helin, K. The E2F family: specific functions and overlapping interests. *EMBO J.* **23,** 4709–16 (2004).
147. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 15545–50 (2005).
148. Chen, H.-Z., Tsai, S.-Y. & Leone, G. Emerging roles of E2Fs in cancer: an exit from cell cycle control. *Nat. Rev. Cancer* **9,** 785–797 (2009).
149. Zhu, W., Giangrande, P. H. & Nevins, J. R. E2Fs link the control of G1/S and G2/M transcription. *EMBO J.* **23,** 4615–26 (2004).
150. Müller, H. & Helin, K. The E2F transcription factors: key regulators of cell proliferation. *Biochim. Biophys. Acta* **1470,** M1–M12 (2000).
151. Pützer, B. M. & Engelmann, D. E2F1 apoptosis counterattacked: Evil strikes back. *Trends in Molecular Medicine* **19,** 89–98 (2013).
152. Hallstrom, T. C., Mori, S. & Nevins, J. R. An E2F1-dependent gene expression program that determines the balance between proliferation and cell death. *Cancer Cell* **13,** 11–22 (2008).
153. Wu, M. *et al.* E2F1 enhances glycolysis through suppressing Sirt6 transcription in cancer cells. *Oncotarget* **6,** 11252–11263 (2015).
154. Poppy Roworth, A., Ghari, F. & La Thangue, N. B. To live or let die - complexity

within the E2F1 pathway. *Mol. Cell. Oncol.* **2,** e970480 (2015).

155. Kang, M.-J. *et al.* HuD regulates coding and noncoding RNA to induce APP→Aβ processing. *Cell Rep.* **7,** 1401–9 (2014).
156. Massey, A. J. *et al.* Multiparametric Cell Cycle Analysis Using the Operetta High-Content Imager and Harmony Software with PhenoLOGIC. *PLoS One* **10,** e0134306 (2015).
157. Hnilicová, J. & Staněk, D. Where splicing joins chromatin. *Nucleus* **2,** 182–8 (2011).
158. Yasuda, K. *et al.* The RNA-binding protein Fus directs translation of localized mrnas in APC-RNP granules. *J. Cell Biol.* **203,** 737–746 (2013).
159. Daigle, J. G. *et al.* RNA-binding ability of FUS regulates neurodegeneration, cytoplasmic mislocalization and incorporation into stress granules associated with FUS carrying ALS-linked mutations. *Hum. Mol. Genet.* **22,** 1193–1205 (2013).
160. Luo, Z., Lin, C. & Shilatifard, A. The super elongation complex (SEC) family in transcriptional control. *Nat. Rev. Mol. Cell Biol.* **13,** 543–547 (2012).
161. Mason, P. B. & Struhl, K. The FACT complex travels with elongating RNA polymerase II and is important for the fidelity of transcriptional initiation in vivo. *Mol. Cell. Biol.* **23,** 8323–8333 (2003).
162. Whitfield, M. L., George, L. K., Grant, G. D. & Perou, C. M. Common markers of proliferation. *Nat. Rev. Cancer* **6,** 99–106 (2006).
163. Korotayev, K. & Ginsberg, D. Many pathways to apoptosis: E2F1 regulates splicing of apoptotic genes. *Cell Death Differ.* **15,** 1813–1814 (2008).
164. Ginsberg, D. E2F1 pathways to apoptosis. *FEBS Letters* **529,** 122–125 (2002).
165. Johnson, D. G. The paradox of E2F1: oncogene and tumor suppressor gene. *Mol. Carcinog.* **27,** 151–7 (2000).
166. Pellicelli, M., Picard, C., Wang, D., Lavigne, P. & Moreau, A. E2F1 and TFDP1 Regulate PITX1 Expression in Normal and Osteoarthritic Articular Chondrocytes. *PLoS One* **11,** e0165951 (2016).
167. Ohtani, K., DeGregori, J. & Nevins, J. R. Regulation of the cyclin E gene by transcription factor E2F1. *Proc. Natl. Acad. Sci. U. S. A.* **92,** 12146–12150 (1995).
168. Farra, R. *et al.* Effects of E2F1-cyclin E1-E2 circuit down regulation in hepatocellular carcinoma cells. *Dig. Liver Dis.* **43,** 1006–1014 (2011).
169. Hwang, H. C. & Clurman, B. E. Cyclin E in normal and neoplastic cell cycles.

Oncogene **24,** 2776–2786 (2005).

170. Keyomarsi, K. *et al.* Cyclin E and survival in patients with breast cancer. *N. Engl. J. Med.* **347,** 1566–75 (2002).
171. Russo, A. J. *et al.* E2F-1 overexpression in U2OS cells increases cyclin B1 levels and cdc2 kinase activity and sensitizes cells to antimitotic agents. *Cancer Res.* **66,** 7253–7260 (2006).
172. Vigo, E. *et al.* CDC25A phosphatase is a target of E2F and is required for efficient E2F-induced S phase. *Mol. Cell. Biol.* **19,** 6379–6395 (1999).
173. Cerami, E. *et al.* The cBio Cancer Genomics Portal: An open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* **2,** 401–404 (2012).
174. Fasken, M. B., Laribee, R. N. & Corbett, A. H. Nab3 facilitates the function of the TRAMP complex in RNA processing via recruitment of Rrp6 independent of Nrd1. *PLoS Genet.* **11,** e1005044 (2015).
175. Sun, S., Zhang, Z., Fregoso, O. & Krainer, A. R. Mechanisms of activation and repression by the alternative splicing factors RBFOX1/2. *RNA* **18,** 274–83 (2012).
176. Alioto, T. S. U12DB: a database of orthologous U12-type spliceosomal introns. *Nucleic Acids Res.* **35,** D110–D115 (2007).
177. Yates, A. *et al.* Ensembl 2016. *Nucleic Acids Res.* **44,** D710–D716 (2016).
178. Jimeno-González, S. *et al.* Defective histone supply causes changes in RNA polymerase II elongation rate and cotranscriptional pre-mRNA splicing. *Proc. Natl. Acad. Sci. U. S. A.* **112,** 14840–14845 (2015).
179. Wu, T. & Fu, X.-D. Genomic functions of U2AF in constitutive and regulated splicing. *RNA Biol.* **12,** 479–485 (2015).
180. Coelho, M. B. *et al.* Nuclear matrix protein Matrin3 regulates alternative splicing and forms overlapping regulatory networks with PTB. *EMBO J.* **34,** 653–668 (2015).
181. Chan, C. C. *et al.* eIF4A3 is a novel component of the exon junction complex. *RNA* **10,** 200–9 (2004).
182. Chanarat, S. & Sträßer, K. Splicing and beyond: The many faces of the Prp19 complex. *Biochim. Biophys. Acta - Mol. Cell Res.* **1833,** 2126–2134 (2013).
183. Jeong, S. SR Proteins: Binders, Regulators, and Connectors of RNA. *Mol. Cells* **40,** 1–9 (2017).

184. Jo, S. H. *et al.* Nanoscale Memristor Device as Synapse in Neuromorphic Systems. *Nano Lett.* **10,** 1297–1301 (2010).

185. Strukov, D. B., Snider, G. S., Stewart, D. R. & Williams, R. S. The missing memristor found. *Nature* **453,** 80–83 (2008).

186. Pickett, M. D., Medeiros-Ribeiro, G. & Williams, R. S. A scalable neuristor built with Mott memristors. *Nat Mater* **12,** 114–117 (2013).

187. Keene, J. D., Komisarow, J. M. & Friedersdorf, M. B. RIP-Chip: the isolation and identification of mRNAs, microRNAs and protein components of ribonucleoprotein complexes from cell extracts. *Nat. Protoc.* **1,** 302–7 (2006).

188. Liberzon, A. *et al.* The Molecular Signatures Database Hallmark Gene Set Collection. *Cell Syst.* **1,** 417–425 (2015).

9. Appendix

Table 2, containing the results of the Gene Expression analysis performed by microarray in HeLa cells where RALY was downregulated, can be consulted **HERE**[1].

Table 3, containing the results of the Gene Ontology and pathway analysis of the gene expression part of the microarray experiment, can be consulted **HERE**[2].

Table 4, containing the results of the Gene Ontology and pathway analysis of the splicing part of the microarray experiment, can be consulted **HERE**[3].

[1]: https://drive.google.com/file/d/0Bwr0nDDQwOP2dWtTNHQxYWh5Nk0/view?usp=sharing

[2]: https://drive.google.com/file/d/0Bwr0nDDQwOP2S0h3bjVva3N5UW8/view?usp=sharing

[3]: https://drive.google.com/file/d/0Bwr0nDDQwOP2OExSSFBSVWthLXc/view?usp=sharing

TABLE 5. List of oligonucleotides used in this study.

List of primers used to clone RALY fragments into the pCMV6-Entry vector (Origene)

PRIMER	**SEQUENCE**
RALY-c-Myc 1-306 Forward	5'-GCTAGCATGTCCTTGAAGCTTCAGGC
RALY-c-Myc 1-306 Reverse	5'-GCGGCCGCGTCTGCAAGG
RALY-c-Myc 1-225 Forward	5'-GCTAGCATGTCCTTGAAGCTTCAGGC
RALY-c-Myc 1-225 Reverse	5'-GCGGCCGCGTACCCTTCTTC
RALY-c-Myc 143-306 Forward	5'-GCTAGCATGCCTGTGAAGCGACCCCG
RALY-c-Myc 143-306 Reverse	5'-GCGGCCGCGTCTGCAAGGCC
RALY-c-Myc 226-306 Forward	5'-GCTAGCATGGATGGAGGTGGCGCCG
RALY-c-Myc 226-306 Reverse	5'-GCGGCCGCGTCTGCAAGGCC

List of the ds-DNA Cas9 guides targeting RALY cloned into the PX330 vector.

Crispr/Cas9 Guide 1 - exon 1	5'- CACCGATCAACTCTCGAGTCTTCAT 5'- AAACATGAAGACTCGAGAGTTGATC
Crispr/Cas9 Guide 2 - exon 1	5'- CACCGTGGAAACCTCAACACAGCTC 5'- AAACGAGCTGTGTTGAGGTTTCCAC
Crispr/Cas9 Guide 3 - exon 1	5'- CACCGGCCGGCTGTTCTGTGCACA 5'- AAACTGTGCACAGAACAGCCGGCC
Crispr/Cas9 Guide 4 - exon 1	5'- CACCGACGGCCATACTTAGAGAAGA 5'- AAACTCTTCTCTAAGTATGGCCGTC
Crispr/Cas9 Guide 5 - exon 2	5'-CACCGCCTAAGCCTGACAGACCCA 5'- AAACTGGGTCTGTCAGGCTTAGGC

List of the primers used for the array validation and for the RNAPII elongation rate measurement.

Gene Name	**Assay Name/Primers**
ACTB	Hs.PT.56a.40703009
B2M	Hs.PT.58v.18759587
GAPDH	Hs.PT.39a.22214836

RALY	Hs.PT.58.26374068
E2F1	FW 5’-GCCAAGAAGTCCAAGAACCA RV 5’-CAGTGTCCTCGGAGAGCAG
OPA1	FW 5’-TGGAGATCAGAGTGCTGGAA RV 5’-AAATAGGGCCACATGGTGAG
ITPR1	FW 5’-GCTACTCTGCCCAAAAGCAG RV 5’-ATTACGGTCCCCAGCAATTT
CTNNBL1	FW 5’-TGCTGTACAGTCGCTTCTCG RV 5’-TGAGCACTTCTGCTCCCTCT
CCNE1	FW 5’-AAGTTGCACCAGTTTGCGTA RV 5’-AGGGGACTTAAACGCCACTT
CCNE2	FW 5’- TCCTTCACCTTTGCCTGATT RV 5’- CCTCATCTGTGGTTCCAAGTC
CCNB1	FW 5’-AGATGGAGCTGATCCAAACC RV 5’-TGGCTCTCATGTTTCCAGTG
CCNB2	FW 5’-GGCTGGTACAAGTCCACTCC RV 5’-ACTTGGAAGCCAAGAGCAGA
CDK1	FW 5’-CCCTCCTGGTCAGTACATGG RV 5’-TCATCAATCAAGAGATTTTGAGG
CDC25A	FW 5’-AGCCTCTGGACAGCTCCTCT RV 5’-CTTCAACACTGACCGAGTGC
CCNT1	FW 5’- TGGGAGATCCCAGTCTCAAC RV 5’- CCTGTTGGGAGTTTTCTCCA
GTF2A1	FW 5’-GCGAACTCGGCAAATACAAA RV 5’-TTCATCCACTCCATCATCCA
GTF2E2	FW 5’-GATTCCATGGACGAGGAGAA RV 5’-ACTCCAGCCAAGTGTTCGTT
ELL2	FW 5’-CCATACAGGGACAGGGTGAT RV 5’-CAGGCCAGTCTCTTTGAAGC
SUPT16H	FW 5’-ATAACGCTGTCATGGACGTG RV 5’-CTAGGGAGCCTTCACGGAAT
SSRP1	FW 5’-TTGTCAACTTTGCTCGTGGT RV 5’-TTTTCGCGTTGACAAAATCA
TFDP1	FW 5’-GGGACCACTTCCTACAACGA

	RV 5’-CTGACATTCCTGAGCCGAGT
ITPR1 Ex5-In5	FW 5’-AGTTCTGGAAAGCCGCTAAGCC RV 5’-ACCAAGGCAGCCACTCACTACT
ITPR1 Ex4-In4	FW 5’-CCTGGTTGATGATCGTTGTGTT RV 5’-TACTCCAGGAAAGCCACCACCTTA
ITPR1 Ex3-In3	FW 5’-TGTACGCGGAGGGATCGACAAA RV 5’-CAACCGCACTTAGAAAGCCAAG
ITPR1 Ex27-In27	FW 5’-AGAGCTTTGGGTGTACAAAGGGCA RV 5’-AGTTTCGAGAAGCCCATCCATCCA
OPA1 In28-Ex29	FW 5’-CCCTGACCTCAGTTTGTTAAATGGG RV 5’-CAGCTGATGGTTAAAGCGCCCGTA
CTNNBL1 Ex5-In5	FW 5’-TGAAGAGGGAGCAGAAGTGCTCAT RV 5’-TATTCACTCCAGCCTCACCACACT

List of the primers used for the splicing measurements.

KIT1 – Ex1-In1	FW 5’- TGCGTTCTGCTCCTACT RV 5’- GTCCTCTCTCCGGATGC
KIT1 – In2	RV 5’- CTCAGTCATCCATATGTCATCC
LRP1B – Ex2-In2	FW 5’- CTGACTGCCCTGATGATTC RV 5’- AACAGCAGTCATCTGTGAAA
LRP1B – In3	RV 5’- ATTTCAGTTGCATTGTTCCAC
ASCC2 – Ex2-In2	FW 5’- ACAAGGACCCGAAGACA RV 5’- GGAACAGATAAAGACGGAATTTAC
ASCC2 – In3	RV 5’-TGGCATGTCATGTGCTTT
CDK6 – Ex2-In2	FW 5’- CTGGAGACCTTCGAGCA RV 5’- TTTCTGGGCCTGAGGATT
CDK6 – In3	RV 5’- CTATAATAGTCGACTGCCTGATAA
GSTA4 – Ex4-In4	FW 5’- CTCTTTGGCAAGAACCTCAA RV 5’- AGTAAATGTGGTGTGCTACAG
GSTA4 – In5	RV 5’- GTTTGTGGCAGCTTCTTTG
EVI5 – Ex14-In14	FW 5’- AGACAGCAAGTCAAGGATTTAG RV 5’- GGTCATCAGGTACTTCAAACTC

EVI5 – In15	RV 5’- CTAGGGATGATACAATCTGTGC
DLEU2 – Ex1-In1	FW 5’- GAGCCCAGAGCTCCGAT RV 5’- GAAAGCCCACAGTGTTGAAAG
DLEU2 – In2	RV 5’- CTAGTCCGATTATATTGCAAAGACA
NPLOC4 – Ex11-In11	FW 5’- CCAGAACAAGCATCCCAA RV 5’- ACGAGGTGGTAATATTTACAGT
NPLOC4 – In12	RV 5’- CCATTCTCCATGTACAGAGTTC
CTNNBL1 – Ex1-In1	FW 5’- GGCGAACTTCTGAGCTACCA RV 5’- GTGAGATGAAAGGGCTCTGG
CTNNBL1 – Ex2-In2	FW 5’- ATGGGGAAGAGGAAGAGGAA RV 5’- GATGGGAAGCTCAGACAAGC
CTNNBL1 – Ex3-In3	FW 5’- GAGCCATTGGATGAAAGCTC RV 5’- AAACGCCACAGAGGAGACAT
CTNNBL1 – In4	RV 5’- GGAGAAAGAGTAACAGCACTTCCC
ITPR1 – Ex3-In3	FW 5’-TGTACGCGGAGGGATCGACAAA RV 5’-CAACCGCACTTAGAAAGCCAAG
ITPR1 – In4	RV 5’-TACTCCAGGAAAGCCACCACCTTA
ITPR1 – Ex5-In5	FW 5’-AGTTCTGGAAAGCCGCTAAGCC RV 5’-ACCAAGGCAGCCACTCACTACT
ITPR1 – In6	RV 5’- TCACGCTGAGAGATGACCAG
Thoc2 – Ex36-In36	FW 5’- GGCAAAGAAAAAGGCAGTGA RV 5’- CAATACCTCTTCTTCCCACTCC
Thoc2 – In37	RV 5’- TCAGCCCTTAAGCATATAAACCA
Thoc2 – Ex38-In38	FW 5’- GTCCTCGGACAAGCACAGAT RV 5’- ACCCGAAGACTTGGTCGTTA
Thoc2 – In39	RV 5’- TACACTGGTAGAATTCGGCAGAT
CTNNBL1 – Ex4-In4	FW 5’- TACCACCTTCTGGTGGAGCTGAAT RV 5’- GGAGAAAGAGTAACAGCACTTCCC
CTNNBL1 – In5	RV 5’- TATTCACTCCAGCCTCACCACACT
Thoc2 – Ex37-In37	FW 5’- GAAACGGGACAGTTCAGGAG RV 5’- TCAGCCCTTAAGCATATAAACCA

Thoc2 – In38	RV 5’- ACCCGAAGACTTGGTCGTTA
ATXN10 – Ex10-In10	FW 5’- TTGATCCTGGACAACTGCAA RV 5’- CATGAGGCAAGGGACTTGTT
ATXN10 – In11	RV 5’- GCATAAAGCCGGAACAAAAA
MAPK8 – Ex6-In6	FW 5’- GGCTACAAGGAAAACGGTCA RV 5’- ATAAGGGCCCCAAGTTTTTC
MAPK8 – In7	RV 5’- ACAATTGAGCAGGGTTCAGA

RNA probes used in the RNA pull-down experiment:

WT E2F1 3’UTR poly-U	AAAGGUUUUUUCUGAU
Mutated E2F1 3’UTR poly-U	AAAGGUCUCUCCUGAU

10. Publications

Hereby are collected the papers Nicola Cornella contributed to:

- **Cornella N**, Tebaldi T, Gasperini L, Singh J, Padgett RA, Rossi A, Macchi P. "The heterogeneous nuclear ribonucleoprotein RALY regulates transcription and the expression of the proliferation-promoting factor E2F1", under revision to the Journal of Biological Chemistry. → *This paper contains all the results described in the first part of this thesis. Here, Nicola Cornella performed all the experiments, except for the ones described in Fig.11, performed together with Annalisa Rossi.*

- Rossi A, Moro A, Tebaldi T, **Cornella N**, Gasperini L, Lunelli L, Quattrone A, Viero G, Macchi P. "Identification and dynamic changes of RNAs isolated from RALY-containing ribonucleoprotein complexes". *Nucleic Acids Res.* (**2017**). Doi:10.1093/nar/gkx235 → *Here, Nicola Cornella engineered the RALY KO HeLa cells through CRISPR/Cas9 technology and performed the relative experiments (**Supplementary Figure S4A-C**). He participated to develop the rationale of the work and performed different experiments on transcript stability in the end not included in the paper.*

- Juarez-Hernandez LJ, **Cornella N**, Pasquardini L, Battistoni S, Vidalino L, Vanzetti L, Caponi S, dalla Serra M, Iannotta S, Pederzolli C, Macchi P, Musio C. "Bio-hybrid interfaces to study neuromorphic functionalities: New multidisciplinary evidences of cell viability on poly(anyline) (PANI), a semiconductor polymer with memristive properties". *Biophys. Chem.* **208,** 40–47 (**2016**). → *Here, Nicola Cornella has been responsible for the biocompatibility experiments of different cell lines on PANI, and for the differentiation of SH-SY5Y cells into neuron-like cells (**Fig. 3, Fig. 4 and Fig. 5**)*.

- Roncador A, Jimenez-Garduño AM, Pasquardini L, Giusti G, **Cornella N**, Lunelli L, Potrich C, Bartali R, Aversa L, Veruchi R, dalla Serra M, Caponi S, Iannotta S, Macchi P, Musio C. "Primary cortical neurons on PMCS TiO 2 films towards bio-hybrid memristive device: A morpho-functional study." *Biophys. Chem.* (**2017**). doi:10.1016/j.bpc.2017.04.010 → *Here, Nicola Cornella performed part of the experiments to evaluate neuron survival and development on the TiO_2 coated slides, refining the use of the High Content Imaging System and the analysis of the samples (**Fig. 5**).*

Declaration of Original Authorship

I, Nicola Cornella, hereby confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Acknowledgements

Vorrei ringraziare il mio tutor, prof. Paolo Macchi, per l'opportunità offertami nello svolgere il dottorato nel suo laboratorio e tutta la struttura organizzativa del programma di dottorato, per l'eccellente supporto.

Un ringraziamento particolare ad Annalisa, per avermi sopportato, supportato ed insegnato molto durante questi 4 anni, sia dentro che fuori dal laboratorio.

Ovviamente, un grazie speciale alla mia famiglia, che in ogni frangente è stata di supporto al mio percorso.

Un grazie a tutte le persone che hanno camminato a fianco me durante questo lungo ed intenso viaggio: gli amici e compagni di ventura e sventura dentro e fuori dal laboratorio, chi è stato di passaggio, chi ho incontrato viaggiando. Non vi nomino, perché siete moltissimi e non vorrei dimenticare nessuno, ma siete stati tutti di fondamentale importanza nel mio raggiungere la meta e nel farmi diventare chi sono. Senza di voi sarebbe stato tutto molto più grigio. Grazie di aver donato colore a tutto quanto, specialmente negli alcuni momenti di frustrazione e sconforto che la ricerca, e in generale la vita, gentilmente hanno offerto. Siete stati parte della mia crescita, soprattutto personale. Ci sarebbero tante belle parole da spendere in modo preciso per ognuno di voi. Il tempo e lo spazio purtroppo non me ne concedono la possibilità, ma sappiate che ve le ho rivolte mentalmente, con un ampio sorriso sulle labbra.

Grazie.