UNIVERSITY OF TRENTO - Italy

International PhD Program in Biomolecular Sciences

Department of Cellular, Computational
and Integrative Biology - CIBIO

31st Cycle

"Beyond transcription: p53-dependent cell death response mediated by DHX30"

Tutor
Prof. Alberto Inga
Laboratory of Transcriptional Networks
Department CIBIO, University of Trento

Ph.D. Thesis of
Dario Rizzotto
Laboratory of Transcriptional Networks

Academic Year 2017-2018

INDEX

1) ABSTRACT

Given its role as “guardian of the genome”, the protein p53 has been extensively studied in the cancer research field, with the aim of exploiting its potential in cancer therapies. Nutlin-3 can be considered the prototype molecule able to activate WT p53 by inhibiting the interaction with its negative regulator MDM2.

Previous evidence coming from literature and experiments performed by other members of the laboratory showed how cell lines undergo different responses when treated with Nutlin-3. Specifically, colon cancer-derived HCT116 cells undergo cell cycle arrest when treated with 10μM of Nutlin-3a for 48 hours, while SJSA1 (osteosarcoma-derived) undergo massive apoptosis. Transcription regulation alone cannot account for the antithetic behaviour of the cell lines since apoptotic factors are transcriptionally modulated in both. Conversely, the analysis of polysome-bound mRNAs revealed a core of differentially regulated genes that share a CG-motif in the 3’UTR whose expression is enhanced in Nutlin-3-treated SJSA1 cells. A luciferase-based assay confirmed that the addition of this motif to a 3’UTR is sufficient to enhance translation of the reporter gene in SJSA1 but not in HCT116 cells treated with Nutlin-3. Analysis of the interactors of the CG-motif by pulldown followed by mass spectrometry revealed a core of interacting RBPs shared between the two cell lines and two specific interactors, DHX30 and MYH9. On the other hand, since the motif was showing a biased composition towards cytosines, we analysed the expression of PCBP family proteins, revealing a differential expression of PCBP2 between HCT116 and SJSA1.

The role of PCBP2 and DHX30 was tested in HCT116 cells, being the cell line expressing higher levels of both RBPs. We confirmed the ability of these proteins to interact with CG-motif containing RNAs and possibly to interact with each other. DHX30 binding affinity to CG-motif RNAs was quantified in vitro by means of ALPHA technology, confirming a strong binding of the protein towards the consensus motif.

The functional role of the proteins was assessed by performing a stable knockdown of either PCBP2 or DHX30 in HCT116 cells via shRNA. In a PCBP2 downregulated context and to a higher extent after DHX30 silencing, the expression of CG-motif RNAs showed enhanced translation, assessed by RNA-sequencing of the polysome-bound mRNAs and confirmed by the luciferase assay, and via qPCR on a selected mRNAs. Analysis of apoptotic cells by annexin V staining in HCT116 shDHX30 cells exposed to Nutlin-3 for 48 hours confirmed the expectations from previous experiments, establishing a role for DHX30 in the translation control of CG-motif-containing pro-apoptotic mRNAs. MYH9 was tested in SJSA1, being the interactor identified in this cell line from mass spectrometry. Its silencing resulted in the enhanced translation of the luciferase reporter and increased propensity towards cell death.

We expanded our model by silencing DHX30 in U2OS cells, which are refractory to apoptosis after Nutlin-3 treatment but share the tissue of origin with SJSA1. Although the silenced clones did not fully recapitulate the phenotype, enhanced expression of CG-motif containing RNAs was observed,

resulting in lower vitality of the cells Nutlin-3 exposure. Instead, overexpression of DHX30 in SJSA1 partially rescued the phenotype, lowering the expression of CG-motif containing targets.

Overall, we demonstrate how post-transcriptional gene regulation can shape p53-dependent responses in cell lines. We identified a combination of cis-acting (CG-motif) and trans-acting factors (RNA binding proteins) that modulate polysome-loading of pro-apoptotic mRNAs resulting in divergent phenotypic outcomes to Nutlin-3 treatment. We have studied the role of the RNA helicase DHX30 in inhibiting translation of CG-motif containing RNAs, resulting in the survival of cells where this protein is highly expressed. We have optimized an in-vitro assay based on ALPHA technology that can be used to screen for small molecule inhibitors of DHX30-RNA interactions in the perspective of developing a combined treatment with non-genotoxic p53 activation with Nutlin-3 or derivatives to promote cell death over cell cycle arrest in cancer cells.

2) INTRODUCTION

2.1 - p53 as a tumor suppressor gene

It is widely known and well accepted that the protein p53 has a pivotal role in preventing tumorigenesis. Discovered in 1979 as an interactor of the T-antigen in cells transformed by SV40 virus[1,2], p53 was at first mistakenly classified as an oncogene, given its higher expression in transformed[3] and tumor-derived cells[4].

Ten years later, p53 was finally recognized as a tumor suppressor factor[5] and it is nowadays considered to be the "guardian of the genome"[6]. Indeed, nearly all the events that can cause harm to the cell (thus being a potential driver for oncogenic transformation) lead to p53 activation, resulting in its stabilization, accumulation and translocation into the nucleus, where it can elicit its function as a sequence-specific transcription factor[7].

Probably the best-known p53-activating phenomenon is DNA damage, either occurring during replication or induced by exogenous agents (such as chemicals, reactive oxygen species and UV or ionizing radiations)[8,9]. In addition, other events like oncogene activation, hypoxia, metabolic and ribosome stress, transposons mobilization and telomere shortening lead to p53 activation[9–11].

Through the action of its target genes, p53 can control different pathways to overcome the stressful situation that originally activated it and prevent cell transformation. The earlier responses to p53 activation are cell cycle arrest and stimulation of DNA repair[12], but extreme situations can induce more drastic responses, such as senescence and apoptosis[13].

The p53-mediated arrest of the cell cycle depends primarily on the transcriptional activation of the CDKN1A gene (also known as p21) and the consequent increase in the protein levels. P21 is a potent inhibitor of the cyclin-dependent kinases (CDK) 2 and 4, that consequently prevents RB phosphorylation that would drive the cell to the S phase of the cell cycle [14,15] (Figure 1). Additionally, p21 can interact with proliferating cell nuclear antigen (PCNA), an essential factor for the DNA replication machinery[16]. Through the action of p21, p53 can halt the cell cycle in the G1/S phase, although this block is reversible and the cell can resume cell cycle once p21 levels drop. Conversely, it has been demonstrated that persistent p21 activation can lead the cell to senescence [17]. Additionally, p53 can influence cell-cycle arrest in G2/M through the increased expression of GADD45 and 14-3-3-σ [18,19].

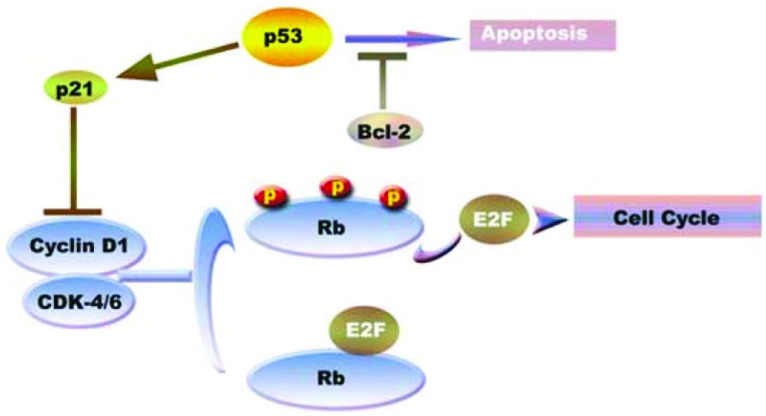


Figure 1: Schematic view of p53-dependent cell cycle arrest in G1/S (adapted from Jayasurya et al, Modern Pathology, 2005 [20])

Another very well-known pathway mediated by p53 activation is apoptosis, both intrinsic (mitochondrial) and extrinsic (death receptor-dependent). As regards the intrinsic pathway, p53 can directly upregulate the expression of pro-apoptotic members of the Bcl2 family, such as BAX, BID, PUMA (BBC3) and NOXA (PMAIP1) [21,22] . BH3-only proteins, like PUMA and NOXA, compete with anti-apoptotic Bcl2 proteins for the activation of BAX and BAK (BH1-3 proteins), whose conformational change allows their homo-oligomerization and insertion in the mitochondrial outer membrane, inducing its permeabilization (MOMP)[23]. The consequence of MOMP is the release of apoptosis-promoting factors, such as DIABLO/SMAC [24] and cytochrome-c. The latter is determinant for apoptosis, since it is complexed to APAF-1 (another p53 target gene), pro-caspase 9 and ATP to form the apoptosome, ultimately starting the caspase activation cascade[23,25].

The extrinsic pathway of apoptosis is controlled by p53 through the expression of death receptors, such as FAS[26], DR4[27] and DR5[28]. The consequence of receptor upregulation is not the immediate induction of apoptosis, but rather a sensitization to apoptotic stimuli provided by FAS-ligand and TRAIL (Tumor Necrosis Factor-related apoptosis-inducing ligand)[29].

In addition, p53 downregulates the expression of BIRC5 (Survivin)[30], a member of the IAP (Inhibition of Apoptosis) family of proteins. Members of this family share a BIR domain (Baculovirus IAP Repeat) and are known to inhibit procaspase activation (blocking both intrinsic and extrinsic pathways of apoptosis), among other functions in cell homeostasis[31].

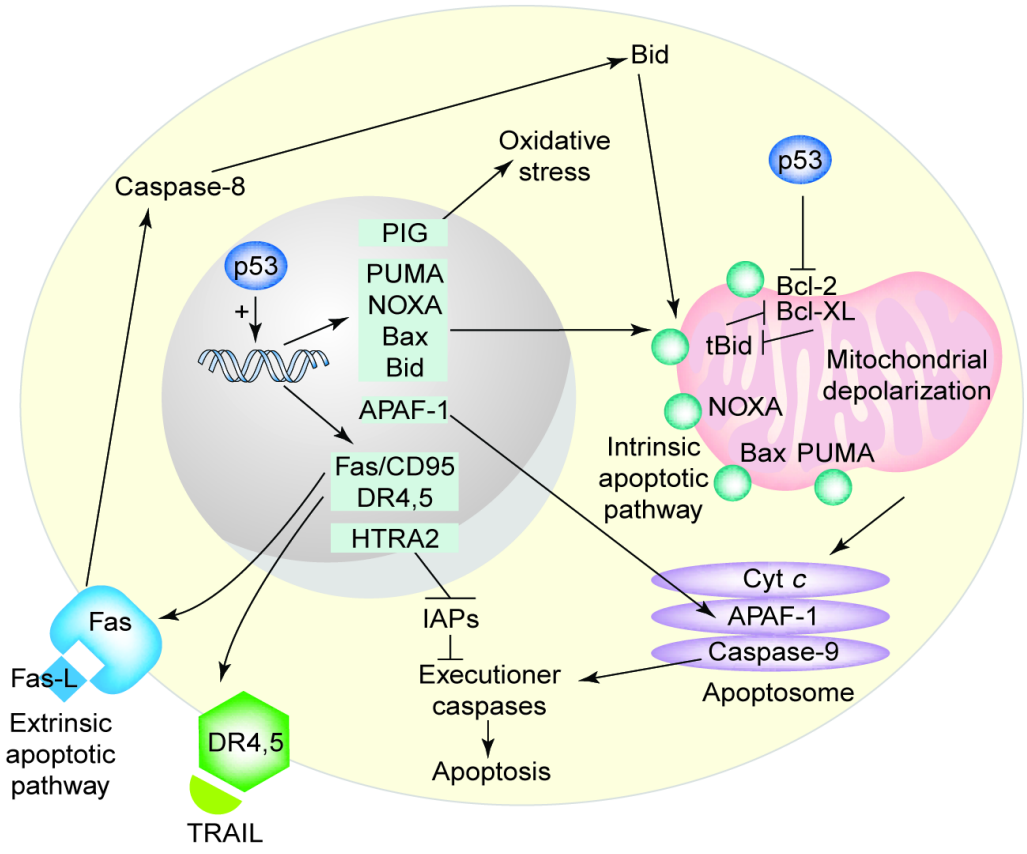


Figure 2: Schematic view of p53 influence on apoptosis pathway (Hofseth et al, Trends in Pharmacological Sciences, 2004 [22])

Besides the transcriptional upregulation of pro-apoptotic and downregulation of anti-apoptotic genes, it was shown that p53 can have direct pro-apoptotic functions in the cytoplasm of the cell by forming complexes with anti-apoptotic proteins like Bcl-xL and Bcl2[32], inhibiting their activity and promoting MOMP. Similarly, p53 can also promote BAX and BAK activation[33,34], resembling the action of BH3-only proteins. It seems that the DNA binding domain is responsible for the interaction of p53 with the members of the Bcl2-family, thus mutations in this domain inhibit both transcriptional and extranuclear p53-dependent promotion of apoptosis[32,35]. Moreover, p53 participates to mitochondrial dynamics, regulating fusion and fission, the latter being a promoter of apoptosis in some circumstances[36]. Direct p53 targets, like mitofusin-2, can directly control the rates of mitochondria fusion-fission[37]. Interestingly, BAK and BAX are involved in the regulation of mitochondria dynamics as well[38], creating a feedback loop between p53, mitochondria and control of apoptosis.

Control of cell cycle arrest and apoptosis are intuitively very prominent tumor suppressive functions of p53, but there is a growing body of evidence suggesting that other p53-controlled activities might be fundamental in preventing oncogenic transformation as well. This concept is supported by the fact that triple negative mouse models *Cdkn1a*$^{-/-}$, *Puma*$^{-/-}$ and *Noxa*$^{-/-}$, retaining WT TP53, do not develop spontaneous cancer compared to *TP53*$^{-/-}$ controls (retaining WT *Cdkn1a*, *Puma* and *Noxa*)[39].

Other important tumor suppressive functions controlled by p53 comprise the control of autophagy, metabolic processes (such as lipid synthesis, fatty acid oxidation and gluconeogenesis), pluripotency and differentiation. All these functions can contribute to maintaining cells' integrity in the context of sub-acute stress[10,40]. Moreover, there is evidence of p53 involvement in the processes of DNA repair, both at the nucleotide level (NER, BER, MMR) and in the repair of DNA double-strand breaks (Homologous Recombination and Non-Homologous End Joining)[12]. Altogether, it is likely the combination of diverse cellular functions controlled by p53 that confers such a strong tumor-suppressive role to this protein and not only the ability to induce cell cycle arrest or apoptosis[7].

In the cancer context, it is not unexpected that inactivation of p53 confers a proliferation advantage to cancer cells. Indeed, more than half of human cancers present a mutation in TP53 gene or have developed other ways to prevent the activity of the WT protein, such as overexpression of negative regulators (in particular the E3 ubiquitin ligase MDM2) or impairment in the upstream activators and/or downstream effectors[13].

On the other side, the consequences of p53 activation and accumulation in healthy cells can be equally deleterious. For this reason, the protein has a very short half-life (and consequently low expression levels) thanks to MDM2, an E3 ubiquitin ligase[9] that binds p53, prevents its transcriptional activity and promotes its proteasomal degradation[41]. This process of degradation is rapidly interrupted when p53 is phosphorylated as a consequence of DNA damage by ATM[42] or by p14^{ARF}-induced MDM2 inhibition[10]. Activation of the p53 transcriptional program leads also to the autoregulation of the pathway, since MDM2 is a transcriptional target of p53 itself [9], causing a

negative feedback loop that self-limits p53 activity. Tight control of the balance between p53 and MDM2 is fundamental to prevent both tumorigenesis and the deleterious effects due to p53 over-activation: *TP53*$^{-/-}$ mouse models are more prone to develop cancer compared to wild-type ones while *Mdm2*$^{-/-}$ mice show embryonic lethality due to constitutive p53 activation, which can be rescued in *TP53*$^{-/-}$ and *Mdm2*$^{-/-}$ double knock-out models [43].

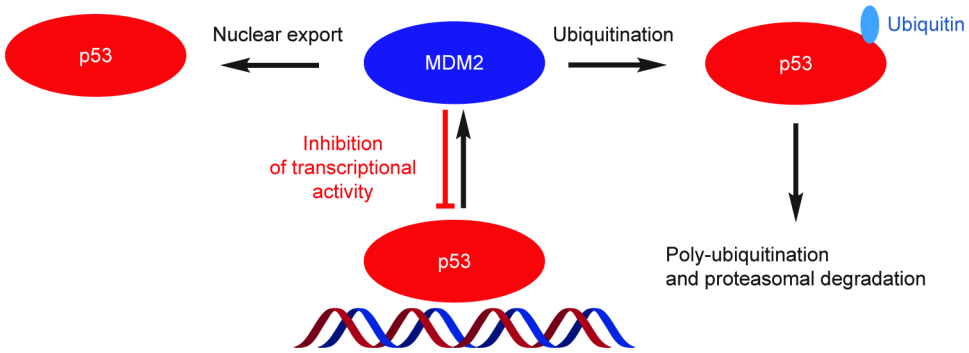


Figure 3: Graphical summary of MDM2-p53 interactions and feedback loop (Wang et al, CSH Perspective in Medicine, 2017 [44])

Figure 4 well summarizes the centrality p53 in the tumor suppressive control of different cellular functions and pathways. It is not a surprise that its high pleiotropy makes p53 a perfect target for inactivation in cells undergoing an oncogenic transformation.

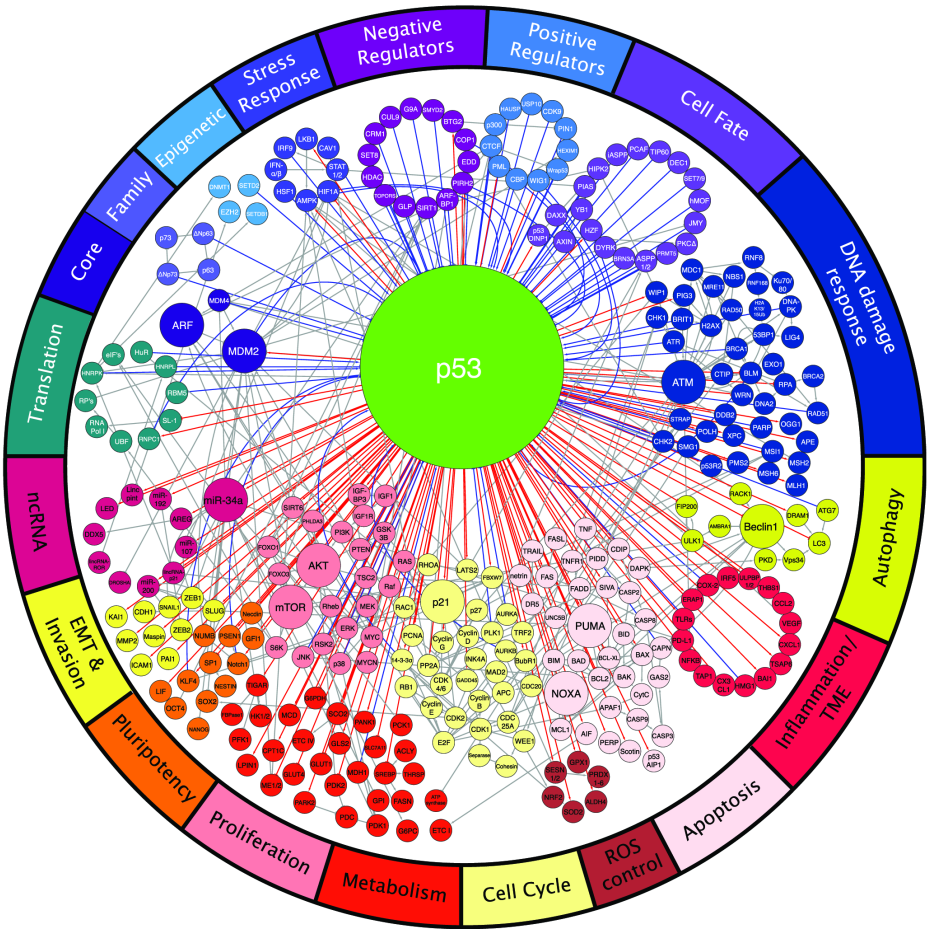


Figure 4: p53 network of target genes divided by pathway and function (Kastenhuber et al, Cell, 2017 [10])

2.2 - p53 as a sequence-specific transcription factor

The p53 protein is 393 amino acids long and presents different domains devoted to different protein functions. Starting from the N-terminus, p53 has (Figure 5):

- a transactivation domain (TAD), containing the binding sites for transcriptional co-activators (in particular p300/CBP) and the negative regulator MDM2[45];
- a proline-rich domain (PRR), that contains five repeats of the SH3-binding motif. The function of this domain is poorly understood, but it might be involved in the interaction with transactivation elements and the overall structure of the protein[45,46];
- a DNA binding domain (DBD) whose function is the recognition of the double-stranded DNA of the response element (RE) on target genes. Most of the p53 mutations found in cancer fall within this domain, with some residues being hotspots for frequent mutations[47](as reported in Figure 5);
- the C-term (CT) region, which contains the nuclear localization signals (NLS), the tetramerization domain (TET) and the nuclear export signal. The TET domain is needed for the oligomerization of the protein, which is functional when combined as a dimer of dimers. Each TET domain is formed by a short β-strand and an α-helix linked by a turn. The C-term region has also a non-specific DNA binding activity, which regulates the activity of p53 [7,45,46].

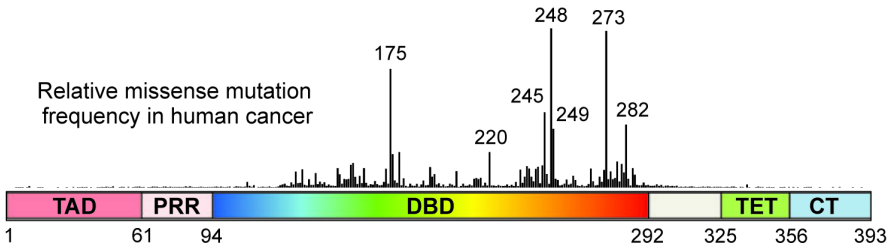


Figure 5: Schematic representation of p53 domains. Bars in the upper part of the picture represent the frequency of mutation and numbers refer to the amino acid most frequently mutated (adapted from Joerger et al, Annual review of Biochemistry, 2008 [45])

Once p53 is freed from MDM2 interaction and its level rises, it translocates into the nucleus where it functions as a sequence-specific transcription factor. The response element (RE) is composed of two repeats of the decamer 5'-RRRCWWGYYY-3' (where R indicates a purine; W an A or a T and Y a pyrimidine) that can be separated by a spacer ranging from 0 to 13 bases. Every p53 monomer in the tetramer interacts with three nucleotides of the RRRCW or WGYYY pentamer[48] (Figure 6).

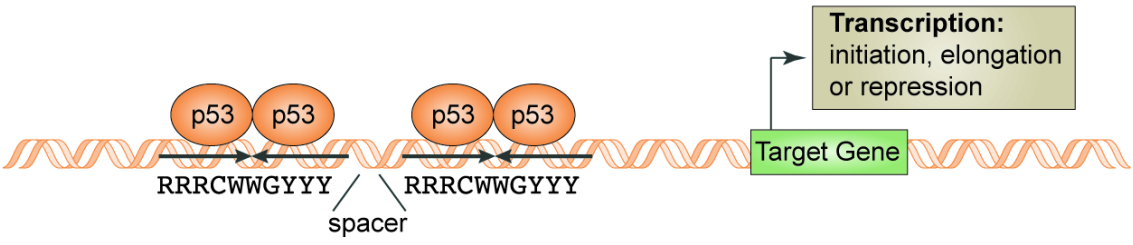


Figure 6: Schematic representation of p53 binding to the response element on target genes (adapted from Menendez et al, Nature Reviews Cancer, 2009 [48])

The relatively high level of degeneration in the RE allows p53 to interact with different strength to the sequences, resulting in different "potency" of the REs. Strong response elements are those containing the CATG sequence in the core region of each decamer and have no or very short spacers (< three nucleotides long) between decamers [49]. Moreover, the number and position of the REs in a gene influence its p53-mediated transcriptional activation. P53 REs are most frequently found in the promoter of the genes, most of the time within 1kb from the transcriptions start site (TSS), but they can be found also in the first intron and, rarely, in the coding sequence[50].

The different strength of the response element sequence combined with their number and position contribute to the regulation of transcription of the target genes, contributing to the outcome of p53 activation. Low levels of the protein allow the transcriptional activation of stronger response elements (such as those present in the p21 promoter) while weak response elements (such as those present in the BAX and NOXA promoters) are activated only when high levels of p53 are reached[48]. Moreover, p53 can recognize and transactivate genes that contain half-sites response elements, although their responsiveness is less than 10% of the p21 RE and thus high p53 levels are needed to induce their transcription[48].

Another layer of complexity in the p53 mediated transcriptional activation is given by the interaction of other cis-acting proteins binding to the promoter of target genes, such as ER[51] and NF-kB[52], and the other member of the p53-family p63 and p73. Both p63 and p73 have distinct and essential roles in development, but they share with p53 some tumor-suppressive functions[53].

The action of p53 as a transcriptional regulator is not limited to activation: it has been demonstrated that p53 can also work as a transcriptional repressor [54]. Most of the genes that are inhibited by p53 encode for cell cycle control[55] and pro-survival functions[56], but the exact mechanism of repression is still debated. It appears that p53 inhibition is not direct, but occurs through p21 and the inhibitory DREAM complex or through the upregulation of some miRNAs [7,57,58].

Despite the effort put in the identification of all p53 transcriptional targets, a comprehensive and unambiguous list of the target genes is not easy to obtain. This is primarily due to the complexity of factors that contribute to the p53-mediated transactivation, such as the strength of the RE, the number and position of the REs in the target gene, p53 post-translational modifications (PTMs) and the interactions with other transcription factors[40]. Moreover, meta-analyses of different ChIP experiments aimed at identifying p53-bound DNA sequences revealed that there is little overlap between different experiments, possibly due to the different cell lines used[59] and the treatment used to activate p53 [60]. It is likely that the flexibility in the p53-mediated activation was evolutionarily selected to allow cells to mount different responses (such as transient cell cycle or apoptosis) depending on the specific situation that initially led to its activation.

Nonetheless, a recent work[59] tried to identify the p53 core transcriptional program by combining different techniques (ChIP-seq, GRO-seq, total RNA-seq and polysome-bound RNA-seq) in different cell lines. An overview of the results is presented in Figure 7. Although some of the genes identified

as part of the p53 core transcription program are well known p53 targets (involved in tumor suppressive functions and identified in previous studies), there are transcripts which were not previously linked to p53 or whose function in tumor suppression is still not clear.

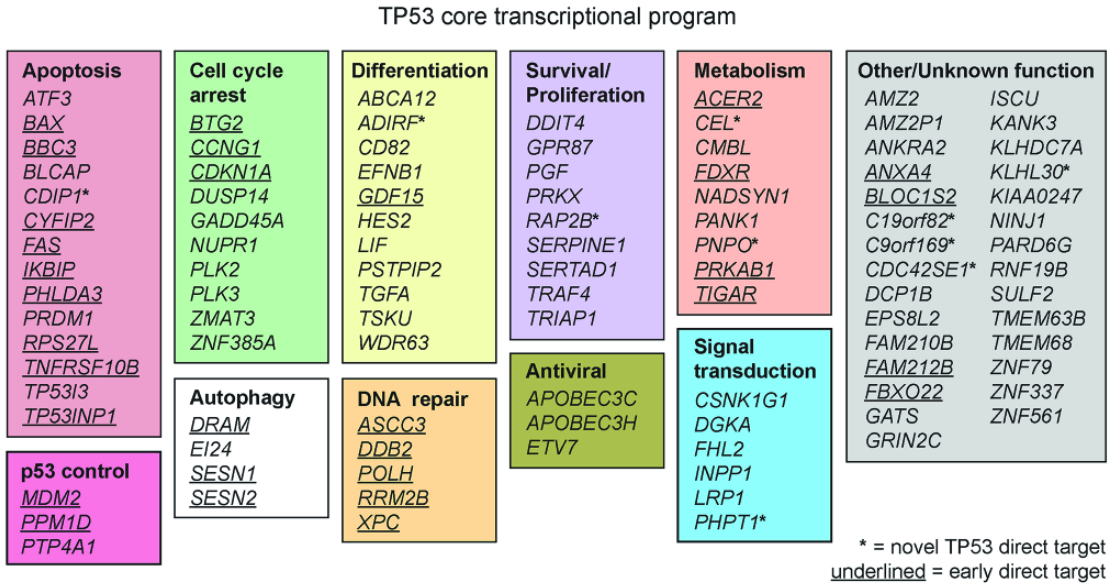


Figure 7: p53 core transcriptional program. Target genes are grouped based on their molecular functions (adapted from Andrysik et al, Genome Research, 2017 [59])

Despite the huge effort in trying to dissect all the features of p53 as a transcription factor, there are still many unsolved questions concerning p53-mediated gene regulation, mainly because the outcome of p53 activation is the result of several layers of control that result in a flexible and diversified response.

2.3 - Mutations of p53

As previously stated, p53 is by far the most mutated gene in human cancers. Most of its mutations are single nucleotides changes, resulting in missense amino acid alterations that eventually cause misfolding of the protein. The type of mutation is strictly cancer dependent: as an example, lung squamous cell carcinomas patients present frequently R158L and V157F substitutions due to the exposure to smoke while ovarian serous carcinoma has often R273H and Y220C mutations[61]. Despite this cancer-specific variability, 95% of all the p53 mutations observed fall in the DBD of the protein[13], with most mutations occurring in only 6 codons (R175, G245, R248, R249, R273, R282) [p53.iarc.fr] (a graphical summary is reported in Figure 8). Different tissues present different susceptibility to p53 mutations, with ovarian, breast, lung and colorectum being the most frequently p53 mutation-bearing cancers and bones, male genital organs and kidney being the less frequently mutated ones [62, p53.iarc.fr].

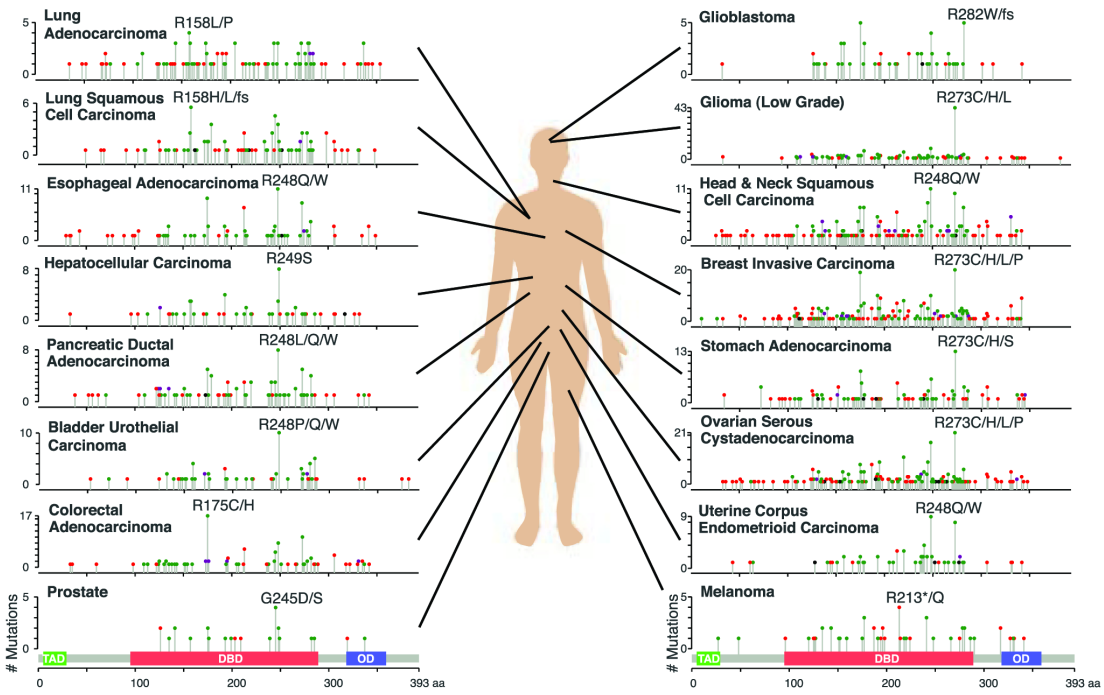


Figure 8: Mutation landscape of p53 in different cancer types divided by tissue of origin (adapted from Kastenhuber et al, Cell, 2017[10])

There are two types of mutant proteins arising from missense alterations: conformational and DNA-contact mutants[10,61]. Although most of the p53 mutations fall within the DNA-binding domain, they can equally have a strong impact on the protein conformation and not only on the DNA-binding capabilities.

Some proteins arising from *TP53* mutations may retain some DNA binding abilities but show an inability to activate certain responses (such as R175P and E180R, defective in inducing apoptosis), while others are termed gain of function (GOF) given their ability to sustain tumorigenesis by promoting drug resistance, angiogenesis and metastasis[10]. Mutant proteins can also have dominant negative functions on the remaining WT allele[13], but most of the time the WT allele is inactivated or lost. Loss of heterozygosity of the p53 locus together with a mutation on the other allele is very often found in cancer and deletion of *TP53* and surrounding genes confer tumors a more aggressive phenotype [63].

Besides missense mutation, nonsense mutations occur in 10% of the cases, with codon 196 and 213 being the most frequently mutated[64]. Those type of mutants are not subject to nonsense-mediated decay and the resulting protein contributes to invasion and metastasis similar to some gain of function mutants[65].

Inherited p53 mutations strongly correlate with increased and highly penetrant susceptibility to cancer, a condition called Li-Fraumeni syndrome. In this condition, 20% on the patients present a mutation in codon R337 (although this is likely due to a founder effect occurred in Brazil[66], increasing the whole percentage of this somatic mutation), followed by other hotspots also found in somatic cancers, such as R175, R248 and R273. R337 mutation causes severe defects in the tetramerization ability of the protein[47], with consequent inability to function.

2.4 - Beyond transcription: p53 as a regulator of post-transcriptional and translational control

The complexity of p53 induced phenotypes is not limited to how this protein functions inside the nucleus to modulate gene transcription: post-transcriptional regulation of mRNA fate and translation control can shape the horizon of p53-mediated phenotypes as well. Among the genes that are transcriptionally regulated by p53, there are also those of micro RNAs (miRNAs) and RNA binding proteins (RBPs), both of which can have a direct contribution in post-transcriptional control.

Micro RNAs are small 20-25 nucleotides RNAs that can base-pair with complementary sequences in the untranslated regions (UTRs, mostly at 3' position) of target mRNAs, lowering their translation. Some p53 target genes, such as *CDKN1A, PUMA, NOXA* and *MDM2*, are targeted by different families of micro RNAs (see Figure 9).

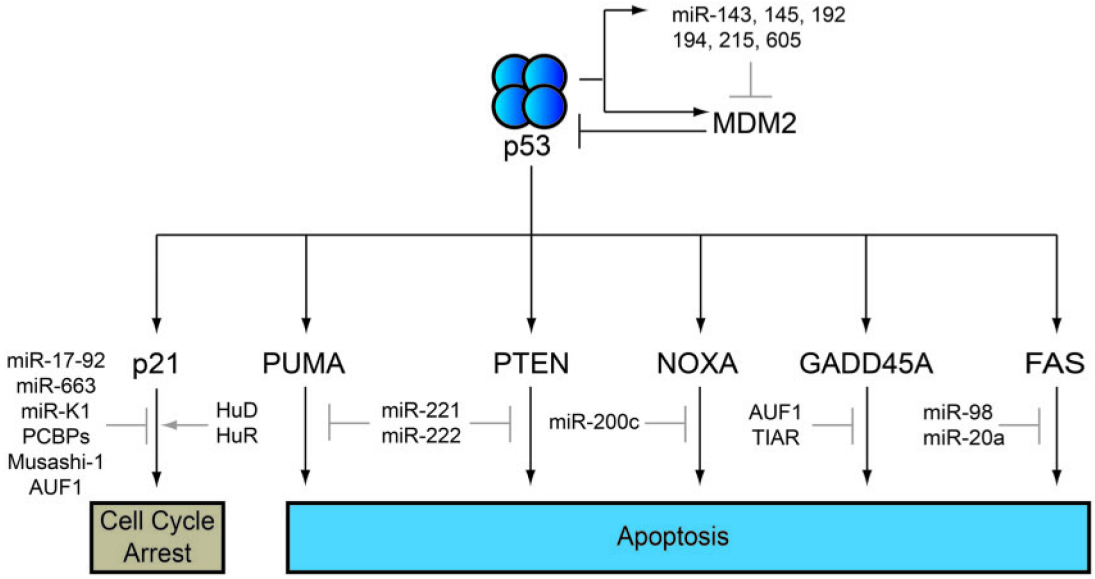


Figure 9: p53 target genes are modulated post-transcriptionally by different RBPs and miRNAs (adapted from Freeman et al, Briefings in Functional Genomics, 2013 [67])

P53 has been directly linked to the enhanced expression of several families of micro RNAs, in particular, the miR-34 one. It was demonstrated that the miR-34 family has tumor suppressive functions by downregulating genes involved in cell cycle progression (*CDK4* and *CDK6*; cyclin D1 an E2), apoptosis (*BCL2*) and epithelial to mesenchymal transition (*SNAIL*)[67]. Loss of or decreased expression of this family of miRNAs has been seen in neuroblastoma, pancreatic and non-small cell lung carcinoma, suggesting a clonal advantage during oncogenesis [68].

Besides micro RNA, RBPs have different roles in regulating mRNA stability and translation initiation rates. The RNA binding protein RBM38 has been found to be a transcriptional target of p53 that participates also in p53 mRNAs translation, leading to its downregulation, through the physical binding with EIF4E[69]. This results in an auto-regulatory loop between p53 and another of its targets, in this case, an RBP. It has also been shown that RBM38 binds p21 mRNA, resulting in increased p21 expression level[69]. Collectively, p53 promotes cell cycle arrest directly by enhancing p21

transcription and indirectly by upregulating RBM38 and miR-34a (which downregulates cyclin and CDKs mRNA) [67].

A confirmation of the effects of p53-mediated post-transcriptional regulation was shown recently by Zaccara et al. In this work the transcriptome and translatome of MCF7 cells were analysed after 16h from a p53-inducing treatment (with doxorubicin or Nutlin-3a). The authors noted how mRNAs presented a certain degree of uncoupling between transcription and translation and identified some new RBPs as p53 transcriptional targets that could account for some of the uncoupling between total (transcription) and polysomal-bound (translation) mRNA levels [70]. In this context, the array of p53 responses and phenotypes could be expanded even further, thanks to the RBPs-mediated translation control of target and even non-target mRNAs.

Ribosome biogenesis is a function often upregulated in cancer cells, that may depend on oncogene (or tumor suppressor loss)-mediated hyperactivation of RNA polymerase I[71]. The gene fibrillarin (*FBL*) codes for a methyltransferase needed in the process of maturation of pre-rRNA (specifically their 2'-O-ribose methylation) and it is often upregulated in cancer cells, resulting in aberrant rRNA methylation patterns. The consequences are decreased translation fidelity and increased IRES-dependent translation, both contributing to tumor growth. P53 binds the promoter of *FBL* gene and downregulates its expression[72]. By regulating the transcription of genes involved in ribosome biogenesis, p53 exerts tumor suppressive functions and, additionally, takes part in translation control by directly managing the production of some ribosome components (summarized in Figure 10).

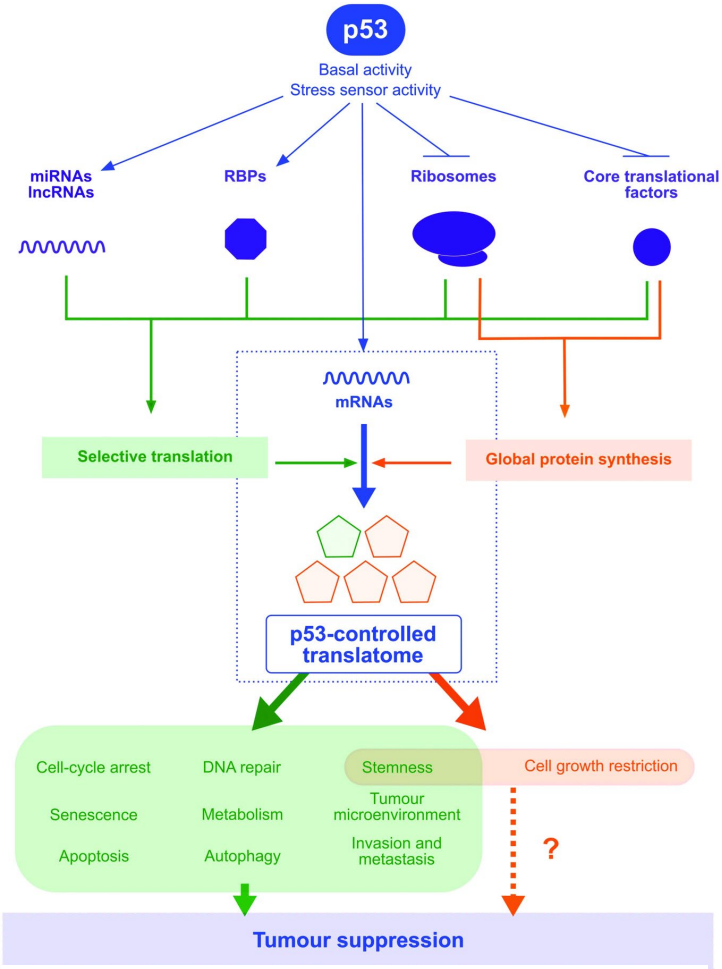


Figure 10: p53 impact on post-transcriptional gene expression regulation (adapted from Marcel et al, Oncogene, 2015 [73])

Additionally, it has been shown by the laboratory of Agami that quiescent and senescent cells (retaining WT p53) downregulate transcriptionally the expression of genes linked to ribosome biogenesis while ribosomal protein production is regulated at the level of translation. The control at translation level is achieved via the modulation of the mTOR pathway by the action of Sestrin 1 and 2, two p53 transcriptional targets[74]. mTOR is a master regulator of cell growth, controlling 5'-CAP-dependent translation mainly by activating EIF4E, which in turn promotes translation initiation[75]. The

same downregulation of the translation apparatus at transcriptional and translational levels was also obtained when p53 was induced using Nutlin-3a, strongly linking p53 and translation regulation through mTOR [74].

Altogether these results expand the complexity of the p53 mediated responses. Besides an intrinsic flexibility of the transcriptional control, p53 can have a significant impact on the translation landscape of the cell, resulting in a complex and often hard to predict the phenotypic output.

Nonetheless, since proteins are ultimately the elements that shape the phenotype, a more comprehensive understanding of how p53 controls translation can have beneficial advantages, particularly in the context of p53 activating therapies.

2.5 - Drugging p53 as a therapeutic strategy

Despite the complexity of p53 dependent responses, its role in preventing tumorigenesis is unquestionable. For this reason, the idea of using p53 as a target for anti-cancer therapies has often been taken into consideration, although p53 is mutated or inactivated in most human cancers[40]. In fact, many molecules that are currently used as standard chemotherapy with DNA-damaging agents (like doxorubicin, carboplatin and camptothecin) or radiotherapy itself rely also on p53 activation to exert their therapeutic benefits. The problem with those types of approaches is the wide spectrum of side effects that they induce in healthy cells. Conversely, p53 targeted therapies could be more beneficial given the non-genotoxic nature of the activation, which, in principle, should not harm healthy cells. Moreover, p53 activation could have a higher impact on cancer cells compared to healthy ones because the first are generally more primed for apoptosis [13,76,77] compared to the latter, which usually respond to p53 activation with cell cycle arrest [44].

Even though different strategies have been developed to reactivate WT p53, I will focus on some approaches targeting the p53-MDM2 interaction using small molecules. When cancer retains WT *TP53* gene, the action of p53 protein is often counterbalanced by its inactivation, mainly achieved via MDM2 overexpression. In these circumstances, the *MDM2* gene is often amplified, but enhanced transcription or translation of its mRNA is also possible. Nonetheless, MDM2 overexpression negatively correlates with p53-inactivating mutations[44]. In this scenario of p53 inactivation, the use of selective inhibitors of the E3-ubiquitin ligase can have beneficial effects: the first compounds discovered to have this selective action are the cis-imidazoline analogues referred to as Nutlins. Nutlin-3a is the prototype molecule belonging to this family that shows the highest capacity to displace MDM2 from p53 by binding the hydrophobic pocket of MDM2 responsible for binding N-term portion of p53 (the structure is reported in Figure 11). This binding causes a dose-dependent WT p53 stabilization and accumulation inside the cell, resulting in early p21 expression, cell cycle arrest (after 24 hours of treatment) and, in some cell lines, also apoptosis (at later time points)[78].

Promising results coming from this molecule boosted compound optimization, which resulted in RG7112 entering in different clinical trials for haematological malignancies[79], liposarcoma, soft tissue

sarcoma, solid tumors and advanced solid tumors[41]. A second generation Nutlin-based MDM2 inhibitor is RG7388 (Idasanutlin), which is more potent than RG7112 and thus expected to induce the same effects at lower concentrations (lowering the chances of unwanted side effects)[41]. Many clinical trials on different types of tumors are still open using this small molecule, often in combination with other drugs (clinicaltrials.gov).

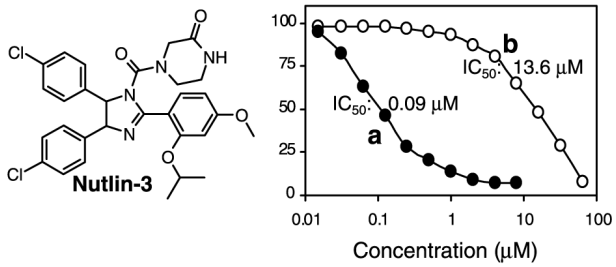


Figure 11: Nutlin-3 structure. The "a" enantiomer shows a more potent binding activity compared to "b" (reported in the graph on the right side). (Adapted from Vassilev et al, Science, 2004 [78])

Another family of compounds able to disrupt p53-MDM2 interaction is the spiro-oxindoles. Contrary to the small molecule screen that led to the discovery of the Nutlins, the development of spiro-oxindoles started from structural data of the p53-MDM2 interaction. These compounds are more potent than Nutlin-3 resulting in lower concentration needed to activate WT p53[80]. The derivative molecule MI-77301 has entered in clinical trials [41]. Many other molecules have been developed and are still in development given the promising results obtained in the preclinical (and sometimes clinical) context of WT p53 activation.

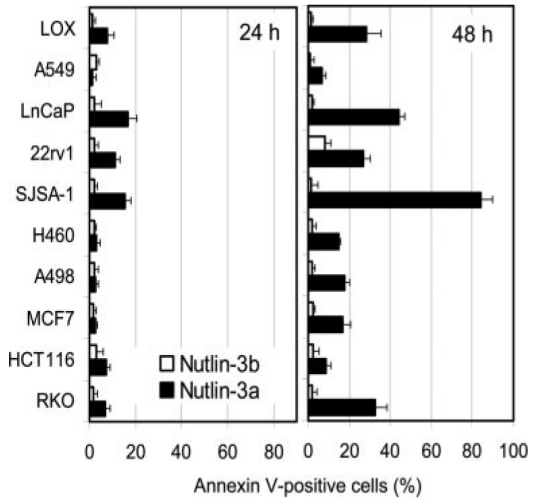
On the other hand, 50% of tumors that bear p53 mutants are, in principle, excluded from this type of therapies. For those cases, small molecules able to restore WT functionalities of p53 mutants have been developed. Besides the biochemical details on the mode of action of these compounds, p53-restoration molecules help mutant p53 to fold properly by acting mainly on the cysteine residues (cysteine-binding compounds), by chelating zinc ions (Zn chelators) or by binding mutated p53 and stabilizing or promoting its refolding (peptides and other small molecules)[61].

An example of a promising p53-reactivating molecule is PRIMA-1. It promotes p53-R273H mutant refolding followed by apoptosis induction of the cells[81]. PRIMA-1 more potent derivative APR-246 is now in different clinical trials, also in combination with other drugs. The advantage of p53-reactivating molecules is given by the fact that most p53 mutant proteins accumulate in the cell. As a consequence of the reactivation, apoptosis-dependent tumor regression was observed in preclinical models[61]. In fact, p53 target genes that control apoptosis are usually transactivated later and/or only when high p53 levels are present. As it was demonstrated by Paek et al, rapid induction of high levels of p53 results preferentially in apoptosis than cell cycle arrest[82]. Small molecule-mediated reactivation of the inactive protein can promptly trigger apoptotic responses, given its high levels already present in the cell. In addition, other p53-independent mechanisms of apoptosis can be modulated by the use of reactivating compounds, which in some cases produce also oxidative stress[61].

3) BACKGROUND AND AIM OF THE PROJECT

In a work published in 2006 by Tovar et al[83] the responses of different cell lines bearing wild type p53 to non-genotoxic activation using Nutlin-3a were evaluated. The authors used a panel of ten cancer cell lines derived from different tissues to model p53 dependent responses. Initially, they focused on p53-dependent cell cycle arrest, modulated by p21 upregulation. As expected, all the cell lines showed a marked reduction in the S-phase of the cell cycle after 24 hours of 10μM Nutlin-3a, suggesting that the treatment triggers a consistent p53-dependent response in all the cells, despite the different tissue of origin.

When they analysed the same panel of cells for p53-dependent activation of apoptosis, the responses were less consistent between the different cell lines. They used annexin V staining to discriminate apoptotic cells (annexin V positive) versus non-apoptotic cells (annexin V negative): at 24 hours all the cells showed a low or very low annexin V positivity (in line with the cell cycle-arrested phenotype that was previously noticed), but at 48 hours cell lines had very different proportions of annexin V positive cells. Some cell lines showed little apoptotic response (i.e.: HCT116, colon cancer-derived cell line), others intermediate (LnCaP, prostate cancer) or high (SJSA1, osteosarcoma-derived cell line) (Figure 12). The authors concluded that this difference was probably due to the different types of p53 inactivation that the cells had developed: SJSA1 have a strong overexpression of MDM2 due to its amplification, while others may have inactivation on some other factors involved in p53-dependent apoptosis. To exclude that the lack of apoptotic responses was not due to inactivation of the cascade, they demonstrated that cells like HCT116, which have no *MDM2* amplification and show low annexin V positivity after 48h of Nutlin-3a, can indeed undergo apoptosis after 48h of 250nM doxorubicin. This suggests that in these cells the apoptotic cascade is functional but cannot be triggered effectively in a strictly p53-dependent fashion. To identify what might be the cause of the inefficient apoptotic commitment in some cell lines, they compared gene expression profiles of apoptotic-prone and resistant cells after 24 hours of Nutlin-3a exposure. They noticed that cell cycle-arresting cells that did not proceed to apoptosis showed expression profiles compatible with defects in the p53-regulated apoptotic signalling, which are preserved in *MDM2* overexpressing cells.



To confirm their observation, they tested the effect of Nutlin-3a treatment in MHM cells, another osteosarcoma-derived cell line which is p53-inactivated via MDM2 amplification. Also in this case, cells respond with initial cell cycle arrest followed by a high apoptotic index (after 48h of treatment).

Figure 12: Apoptotic responses of different cell lines treated with a p53-activating molecule (Nutlin-3a) or its inactive enantiomer (Nutlin-3b). (Adapted from Tovar et al, PNAS, 2006 [83])

The conclusion by Tovar and co-workers is that an intact signalling downstream p53, which is preserved by p53 inactivation via *MDM2* amplification (and/or overexpression), can lead to massive apoptosis. On the other hand, cells which still have WT p53 and no *MDM2* amplification have evolved other ways to inactivate p53-dependent responses resulting in lower susceptibility to cell death. According to their results, this inactivation is, at least partially, transcription-dependent.

Their results can be challenged by the work of Andrysik et al[59], where they tried to identify the core program directed by p53 by combining ChIP-seq, GRO-seq and RNA-seq in different cell lines that have different phenotypic outcomes to Nutlin-3 treatment. They used HCT116, SJSA1 and MCF7 which are apoptotic refractory, apoptotic prone and weakly apoptotic, respectively (as previously noticed by Tovar et al., reported in Figure 12). What they observed is that p53 direct transcriptional targets are not completely conserved between cell lines, but the differences increase much further with time and as gene expression regulation proceeds along the central dogma (Figure 13). Nonetheless, they identified a common core of ≈100 genes (defined as p53 core transcriptional program, reported in Figure 7), which is shared in all the cell lines. This core program comprises 14 pro-apoptotic genes (like BAX, PUMA, NOXA) that are upregulated also in those cell lines that do not undergo apoptosis (HCT116 and in part MCF7), strongly suggesting that phenotype divergence occurs later after transcription.

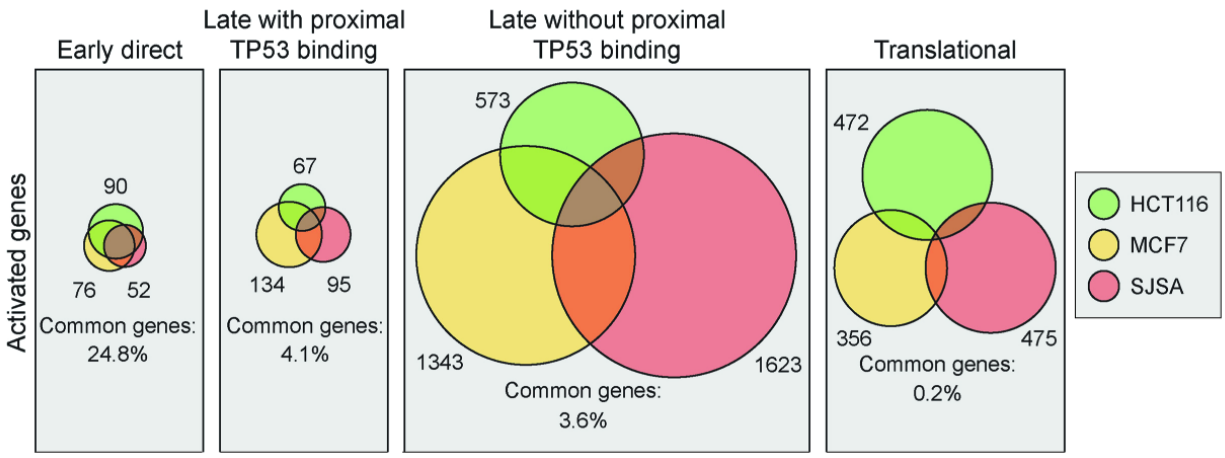


Figure 13: Venn diagram representing the fractions of genes upregulated in response to p53 activation in HCT116, MCF7 and SJSA1. The number of genes that are shared between the cell lines becomes progressively lower along the central dogma. (adapted from Andrysik et al, Genome Research, 2017)

In line with the work of Andrysik et al, Zaccara et al[70] analysed the responses of MCF7 cells to p53 activation using doxorubicin and Nutlin-3. What they discovered is that polysome-associated mRNAs (i.e. translatome) are enriched for apoptosis-related genes, especially after Nutlin-3 treatment, supporting a partial tendency of MCF7 to undergo cell death after exposure to the small molecule. As stated previously in section 2.3, they suggest an indirect p53-role in post-transcriptional gene expression regulation as responsible for the uncoupling between the transcriptome and the translatome.

The aim of my work is the analysis of how post-transcriptional gene regulation impacts on the phenotypic outcome of a cell line after exposure to p53-activating molecules. We used two cell lines to model the antithetic responses that were originally identified by Tovar et al, that is survival (persistent cell cycle arrest, HCT116 cells) or cell death (apoptosis, SJSA1 cells). According to our discoveries, the choice may depend on a cis-acting feature present in the 3'UTR of some mRNAs that can be bound by trans-acting RBPs, to favour or dampen their polysome loading and, consequently, protein production. Since many mRNAs containing the cis-acting feature are involved in apoptosis, the resulting increased protein production influences the phenotypic outcome of the cells to p53 activation by Nutlin-3.

In order to have a more streamlined presentation of the experimental workflow, I included some experiments that were obtained by other people before my arrival in the laboratory in the same sections with other results that were obtained by me. All the people involved in the different experiments are acknowledged at the beginning of each section.

4) RESULTS

4.1 - The CG-motif

4.1.1 - HCT116 and SJSA1 RNA sequencing

Experiments presented in this section were performed by Sara Zaccara and Erik Dassi before my arrival in the laboratory

In order to analyse how translation control can impact on the phenotypic outcome of p53 activation after Nutlin-3 treatment, two different cell lines were taken into consideration: HCT116, colon cancer-derived cell line that shows a persistent cell cycle arrest and SJSA1, osteosarcoma-derived cell line that undergoes initial cell cycle arrest followed by apoptosis [59,83]. These cells were chosen for their divergent responses to an identical, non-genotoxic p53 activation using 10μM Nutlin-3. After 12 hours of treatment, cells were collected and prepared for sucrose gradient polysome fractionation. This time point was chosen because cells show initial p53-mediated responses, such as cell cycle arrest, but SJSA1 are not yet undergoing apoptosis[59]. Polysome profiling is a technique that allows the separation of mRNAs based on their "weight" given by their association with ribosomal subunits (40S,60S), monosome (80S) or poly-ribosomes (polysomes). Lysates are loaded on the top of a 15% to 50% sucrose gradient and ultracentrifuged at 40000 RPM for 100 minutes. Fractionation is achieved using a Teledyne Isco instrument following the absorbance of the gradient at 260nm. Lighter fractions are collected earlier during fractionation while polysomes are present at heavier sucrose densities (thus collected in the last fractions). Fractions containing the free RNA up to the monosome (80S) are considered sub-polysomes (sub) while the others are considered polysome fractions (pol). Fractions belonging to the pol or sub were grouped in two pools and RNAs were extracted using phenol-chloroform before proceeding with library preparation for ION-Torrent sequencing. Total RNAs were extracted from matching cell lines and treatment using commercial RNA extraction kits and processed for library preparation. A schematic view of the experimental setup is presented in Figure 14.

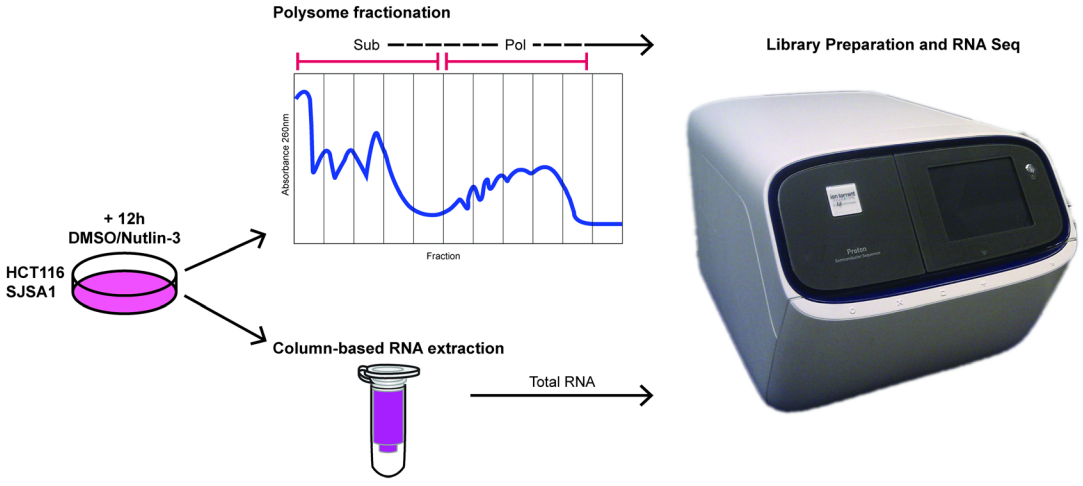


Figure 14: Schematic view of the experimental procedure for RNA sequencing in HCT116 and SJSA1 cell lines. Sub = RNAs present in the sub-polysome fractions; Pol = RNAs present in the polysome fractions

According to the output of the RNA sequencing, genes modulated by the Nutlin-3 treatment that had a $\log_2$ fold change above 1 (upregulated) or below 1 (downregulated) were defined as differentially expressed genes (DEGs) (adjusted p-value <0.1). DEGs can fall into 3 different categories:

- *coupled*: if they have a similar modulation (up- or down-regulation) both at total mRNA and polysome-bound mRNA level;
- *translationally regulated*: if they are modulated at the polysomes level, but not in the total and sub-polysomal mRNAs;
- *unchanged in translation*: genes which show modulation in the subpolysomal fraction that is not followed by enhanced polysome association.

The complete table of DEGs is included in the attachments (section 11.1). Figure 15A reports the relative proportion of coupled, translationally regulated or unchanged genes both up- and down-regulated after Nutlin-3 in SJSA1 and HCT116. The number of up and down-modulated DEGs is similar in the two cell lines, but the ratio of translationally-regulated vs coupled genes changes between them, with HCT116 showing a higher number of coupled upregulated genes compared to translationally-regulated and SJSA1 having a higher number of coupled downregulated genes with respect to translationally downregulated. Figures 15B, 15C, 15D and 15E report the overlap between the different categories of genes in the two cell lines. When looking at the coupled genes (Figure 15B and 15D) HCT116 and SJSA1 show a significant level of overlap both in the up and down-modulated genes.

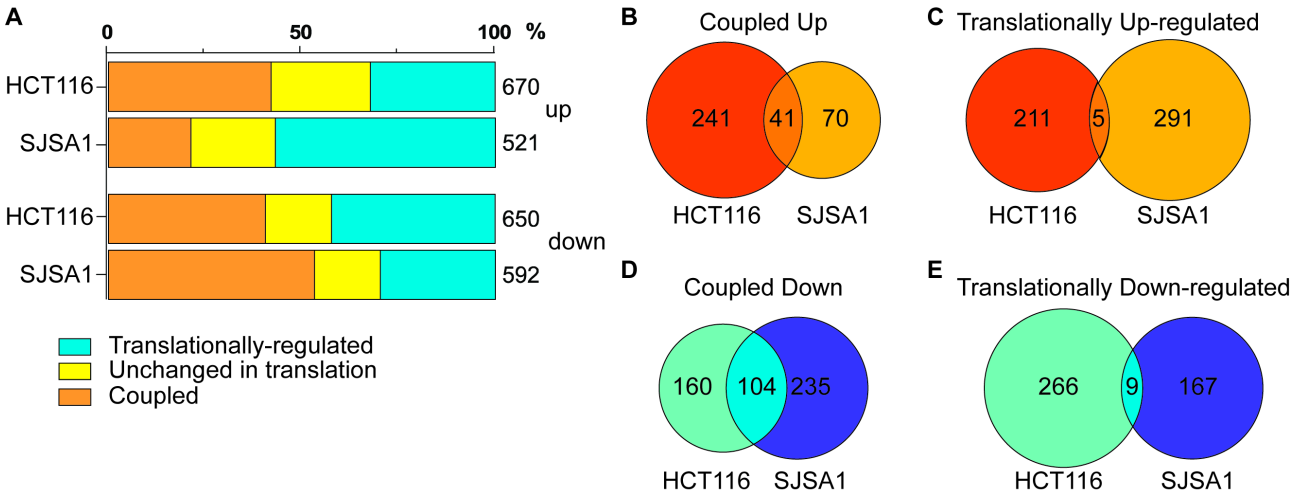


Figure 15: A) DEGs up- and down-regulated in HCT116 and SJSA1 divided by category. Venn diagrams representing the overlap of genes between HCT116 and SJSA1 in the coupled (panels B and D) and translationally regulated (panels C and E) categories

Ontology analysis with Metascape[84] of coupled upregulated genes revealed terms related to p53 activation (Figure 16A) with well-established p53 targets (such as *MDM2*, *CDKN1A*, *BBC3* and *BAX*) upregulated in both cell lines (Figure 17). The relatively high level of overlap between coupled genes in HCT116 and SJSA1 and the presence of genes involved in cell cycle arrest and apoptosis as p53

core targets does not allow a clear understanding of the different phenotype. When comparing the coupled downregulated gene lists, no striking differences where highlighted (Figure 16C) since also in this case, a relatively high proportion of shared mRNAs is present.

When looking at the level of translation regulation (Figure 15C and 15D), cell lines show a more divergent array of regulated genes: only 5 or 9 genes are commonly translationally up- or down-regulated, respectively. By looking at the ontology analysis of these lists (Figure 3B and 3C), apoptosis-related terms appeared only in the translationally upregulated genes in SJSA1, consistent with the phenotype that they will show in the following hours.

These findings suggest that HCT116 and SJSA1 choose their destiny early after p53 activation and translation control may have a very profound impact on the outcome of p53-mediated responses. Indeed, transcriptional programs are very similar between the two cell lines, comprising both cell cycle arrest/pro-survival factors and apoptotic genes. Conversely, translation programs show very little overlap, with terms related to apoptosis particularly represented in the list of SJSA1 translationally upregulated DEGs.

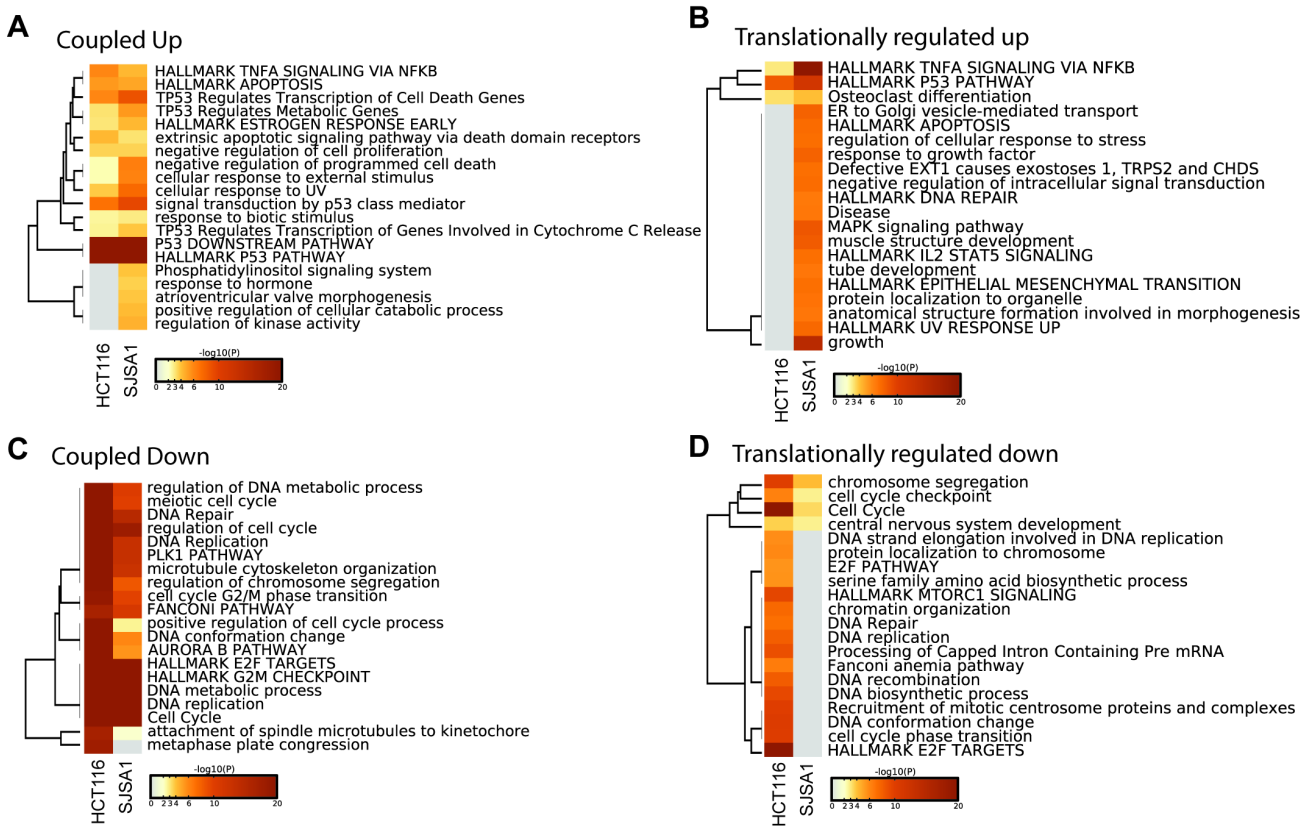


Figure 16: Metascape analysis of the different categories of DEGs in HCT116 and SJSA1 cells

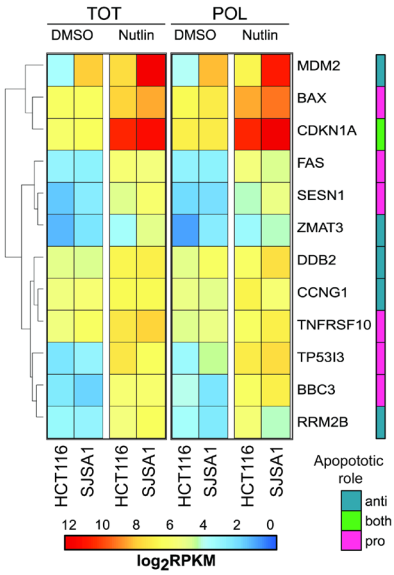


Figure 17: Expression level of some direct p53 target genes in the total (TOT) and polysome-bound (POL) mRNAs in HCT116 and SJSA1 with or without Nutlin-3 treatment. Values are expressed in reads per kilobase million (RPKM); the right bar indicates the role of each gene in the induction of apoptosis.

4.1.2 - The CG-motif discovery

Experiments presented in this section were performed by Sara Zaccara and Erik Dassi before my arrival in the laboratory

Having noticed that the divergence between HCT116 and SJSA1 may stem from a differential translation program, people working on this project before my arrival in the laboratory decided to use different bioinformatics tools to look for common feature shared by mRNAs that could explain why one cell line translates a certain subset of mRNAs and not some others.

There is a complex combination of factors that can account for the differential translation program between HCT116 and SJSA1, but miRNAs and RNA binding proteins may have an immediate and straightforward effect on post-transcriptional gene expression regulation, thus they were promptly investigated.

The search for common miRNA seeds shared on the different gene lists (presented in Figure 15) did not lead to any significant enrichment for specific micro RNA families. For this reason, the attention was placed on the RBPs, firstly by looking for common nucleotide motifs that could work in-cis in the translation regulation after being bound by RBPs. To achieve this goal, the tool Weeder [85] was used. This software was developed to search for conserved motifs in datasets of nucleic acid sequences, which might be instances for binding sites of specific transcription factors or, as in our specific case, RNA binding proteins.

Weeder was employed to analyse the 5'UTRs and 3'UTRs of all the DEGs lists coming from both cell lines and allowed the identification of a putative sequence enriched in the 3'UTR of SJSA1's translationally upregulated genes. The consensus of the motif is 5'-CCCC(A/C)(T/G)GGCCCT-3' and for this reason, it was named "CG-rich" (the logo is reported in Figure 18A). This motif appears to be particularly present in the polysome-bound mRNAs after Nutlin-3 treatment in SJSA1 (Figure 18B), with 65% (194/296) of the DEGs showing at least one instance of it in the 3'UTR.

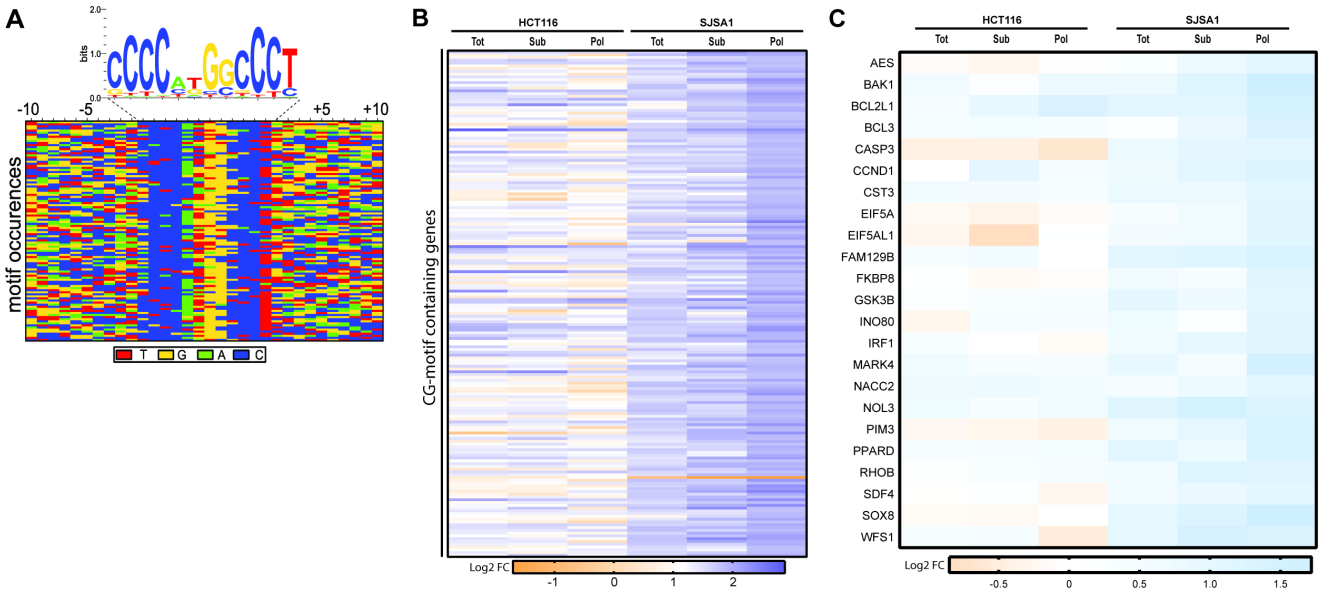


Figure 18: A) The CG-motif consensus; B) Heatmap representing the expression of CG-motif containing mRNAs in the different lists of HCT116 and SJSA1. Values are presented as log2 fold change (FC; Nutlin-3 over DMSO) of the RPKM values. Tot = total RNA; Sub = sub-polysomal RNA; Pol = polysome-bound RNA. C) Heatmap representing the relative expression of apoptosis-related genes in the different lists of RNA coming from the sequencing of HCT116 and SJSA1 samples. Values are presented as fold change (FC) of the RPKM values.

When looking at the subset of genes involved in apoptotic processes (previously identified by Metascape analysis, see Figure 16B) that contain the CG-motif, it is possible to notice a higher level of expression in the polysome fractions of SJSA1, as reported in Figure 18C.

4.1.3 - CG-motif in translation: the luciferase assay

Experiments presented in this section were performed by Sara Zaccara before my arrival in the laboratory.

In order to define if the CG-motif has truly a role in translation modulation, the consensus sequence was added in two copies upstream and downstream the β-Globin 3'UTR, which normally does not contain any instance of the motif. They decided to use two copies of the motif since the vast majority (88,2%) of the 194 genes containing the motif identified in SJSA1's translationally upregulated DEGs contain at least two copies of it. This chimeric 3'UTR was then placed downstream the luciferase 2 coding sequence (*luc2*) inside the pGL 4.13 vector (Promega). A scheme of the plasmid is reported in Figure 19A. As controls, both the vector without the β-Globin 3'UTR or the β-Globin 3'UTR without the addition of the motif were used. The vectors were transfected in HCT116 and SJSA1 that were treated with Nutlin-3 or vehicle DMSO for 12 or 24 hours before proceeding with cells harvesting and dual-luciferase reporter assay (Promega). Results of the luciferase assay are reported in Figure 19B.

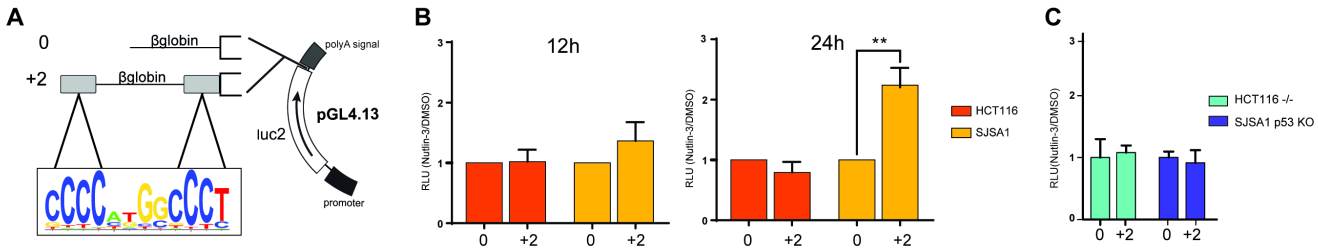


*Figure 19: A) Schematic view of the vectors used to assess the effect of the CG-motif; 0 refers to the β-Globin 3'UTR without motif addition while +2 refers to the addition of 2 copies of the consensus motif. B) Results of the luciferase assay at 12 and 24 hours after Nutlin-3 or DMSO treatment. RLU = relative luciferase units; **= p-value < 0.01 C) p53 dependency of the luciferase assay*

The luciferase assay result suggests that the presence of the motif enhances the luciferase translation in the context of SJSA1 cells treated with Nutlin-3. This effect is already visible at 12 hours after the treatment and becomes statistically significant after 24 hours. We can exclude a transcription-dependent effect since all the relative luciferase units (RLU) presented are normalized on the *luc2* mRNA level quantified by RT-qPCR.

The presence of one copy of the motif alone is sufficient to slightly increase luciferase activity, although not in a statistically significant way (data not shown).

To test the p53 dependency of the phenomenon, p53-deficient clones of HCT116 and SJSA1 were tested after transfection with the same vectors and treated with Nutlin-3 or vehicle (Figure 19C). In this case, no differences in luciferase induction were noticed, not even in the SJSA1 p53 KO clone.

Altogether these data suggest that the presence of the CG-motif in a 3'UTR increases the translation of the genes containing it, as assessed by this luciferase reporter. This effect seems to rely on p53 activation using Nutlin-3 and is limited to SJSA1 cells only.

4.1.4 - Effects of the motif on genes translation

Experiments presented in this section were performed by Sara Zaccara and me.

To assess if the CG-motif on endogenous mRNAs truly enhances their polysome association and translation, confirming the RNA-seq data, we performed a qPCR analysis on a selected panel of targets which have apoptosis-related functions (Figure 20).

As depicted in Figure 20A, all the apoptotic targets considered in qPCR analysis show enhanced polysome association after 12 hours of Nutlin-3 in SJSA1, while this effect is not visible on HCT116 (except for CASP3). To confirm that the enhanced polysome association corresponds to increased mRNA translation and protein production we checked the expression of some targets via western blot (Figure 20B), performed on samples collected 24 and 48 hours after the treatment. As expected, CASP3 and BAK1 expression are increased after Nutlin-3, more strikingly in SJSA1, confirming the RNA seq and qPCR data. Conversely, EIF5A expression detected by western blot seems to disagree with previous data, possibly due to the time offset between qPCR and western blot and/or a different turnover of the mRNA/protein. The slight decrease in CASP3 expression in SJSA1 cells after 48h may be due to the overt apoptotic phenotype that these cells show at that time point, as assessed

also by the presence of PARP cleavage (cPARP). The successful induction of p53 was checked by looking at p21 expression.

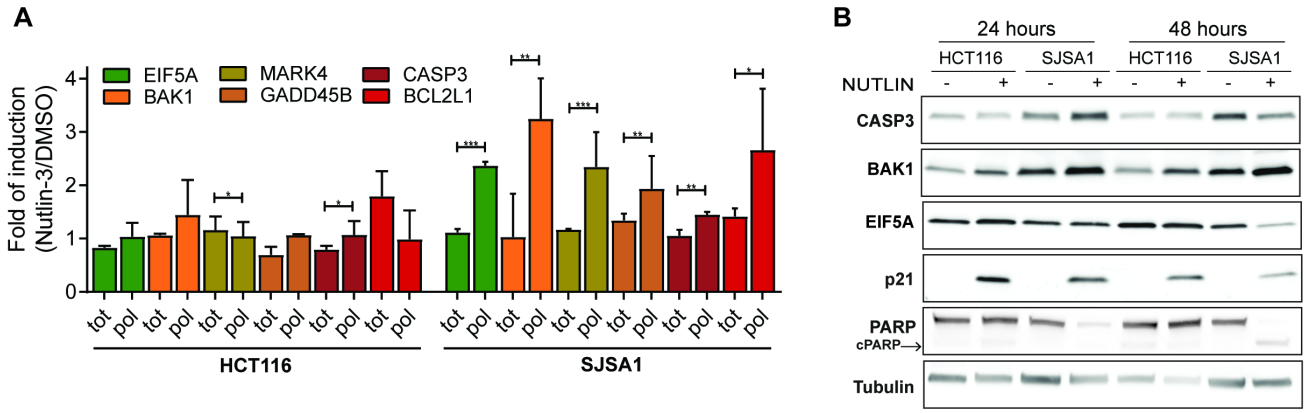


*Figure 20: Panel A) qPCR of selected CG-motif containing, apoptosis-related mRNAs expression in HCT116 and SJSA1. Polysome-bound mRNA levels are normalized over the total mRNA. * = p-value < 0.05; ** = p-value < 0.01. Panel B) Protein expression of some CG-motif containing mRNAs by western blot. P21 was tested to assess successful p53 induction; PARP (and the cleavage product cPARP) were tested to confirm the activation of apoptosis in SJSA1; Tubulin was used as a loading control*

Considering the results from the luciferase reporter assay and the endogenous CG-motif mRNAs expression we can conclude that the presence of the CG-motif enhances translation of the mRNAs in the context of SJSA1 cells, specifically after p53 activation using Nutlin-3.

4.2 - The interactors of the motif

4.2.1 - Knowledge-based approach: PCBP2

Experiments presented in this section were performed by Sara Zaccara and me.

Having determined how the presence of the motif can enhance mRNA translation in SJSA1, the next step was to look for interactors fostering translation of mRNAs containing the CG-motif in this specific cellular context and not in HCT116 cells. These interactors could be induced or repressed by Nutlin-3 or be constitutively differentially expressed in different cell backgrounds.

The first approach used to identify putative interactors was knowledge-based: since many of the nucleotides composing the consensus motif are cytosines (7/12 nucleotides), we looked at the PCBP family of RBPs. Members of this protein family have a high affinity for single-stranded C-rich sequences and share a common structure with two repeats of the KH (hnRNP K homology) domain at the N-term of the protein and a third domain at the C-term. There are 5 loci encoding for the members of this family, that are subdivided in the hnRNP K/J and α-CPs proteins. The only member of the hnRNP K/J subfamily is hnRNP K while the α-CP is composed by αCP-1 (also known as hnRNP-E1 or PCBP1), αCP-2 (aka hnRNP-E2 or PCBP2), αCP-3 (PCBP3), αCP-4 (MCG10)[86]. PCBPs are involved in diverse aspects of RNA biology, from transcription to stabilization and translation control, making them good candidates for further analysis.

By looking at RNA sequencing data, the expression of the members of the PCBP family seems to have little variation after Nutlin-3 treatment: none of these genes passed the threshold (log_2 FC > or

< 1; adj p-value <0.1) to be defined as DEGs neither in HCT116 nor in SJSA1 (Figure 21A). Surprisingly PCBP4, a known p53 target[87], did not show a significant induction in any cell line after the treatment. We also checked the expression of some proteins via western blot (Figure 21B) and noticed that PCBP2 was constitutively less expressed in SJSA1 compared to HCT116, contrary to PCBP1 and PCBP4 which show a similar expression level in both contexts. Moreover, Nutlin-3 treatment had a very modest effect on PCBPs expression levels, confirming RNA sequencing data. Given the constitutive difference in PCBP2 expression between HCT116 and SJSA1, we decided to focus our attention on this protein and investigate in more detail how it may bind the CG-motif and regulate mRNAs containing it.

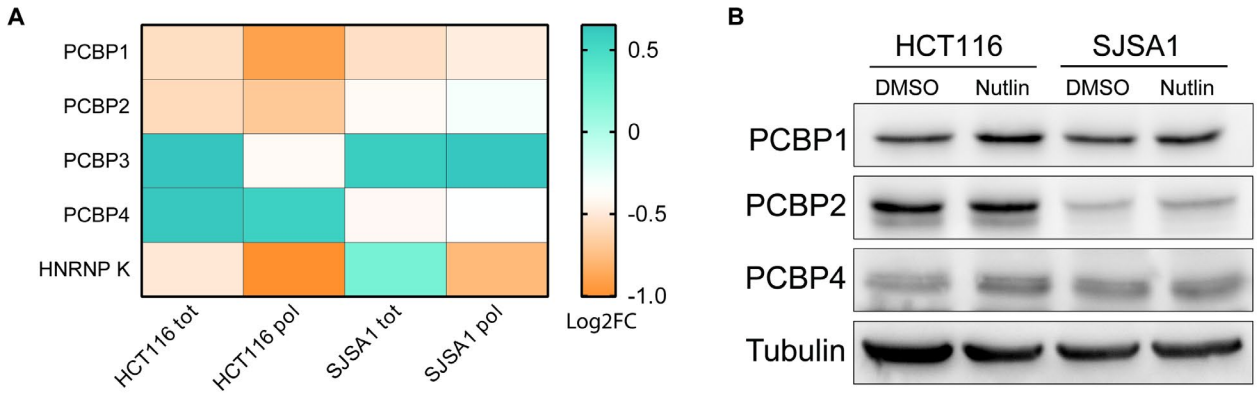


Figure 21: A) Expression level of the members of PCBP RBPs as detected by RNA sequencing on the total and polysome-bound mRNAs. Heatmap shows the log_2 Fold Change (Nutlin-3 over DMSO) of the RPKM. B) western blot representing the expression level of some PCBPs in HCT116 and SJSA1 with or without Nutlin-3 treatment.

4.2.2 - Unbiased approach: mass spectrometry

Sara Zaccara optimized pulldown conditions and prepared samples for mass spectrometry. I worked on the data analysis of the mass spectrometry and subsequent validation experiments.

The other approach aimed at identifying possible interactors of the CG-motif was exploiting a 5' biotin-tagged RNA bait containing the consensus motif sequence to pull-down interacting proteins in vitro for a subsequent identification using mass spectrometry. Protein lysates derived from HCT116 and SJSA1 (either treated with control or Nutlin-3 for 12 hours) were added to Dynabeads coated with the RNA probe for 1 hour, washed and loaded on an SDS-PAGE gel for separation by molecular weight. After staining the gel using Colloidal Coomassie, it was noticed that the upper part of the gel, corresponding to proteins above 70kDa showed a differential band pattern between the cell lines (Figure 22A). For this reason, bands were extracted from the gel and prepared for protein identification using label-free quantification mass spectrometry (LFQ-MS; performed by the IFOM proteomics facility, Milan). Figure 22B shows a core of RBPs that are shared between the two cell lines and two proteins that bind the RNA probe in a cell line-dependent manner: MYH9 for SJSA1 and DHX30 for HCT116.

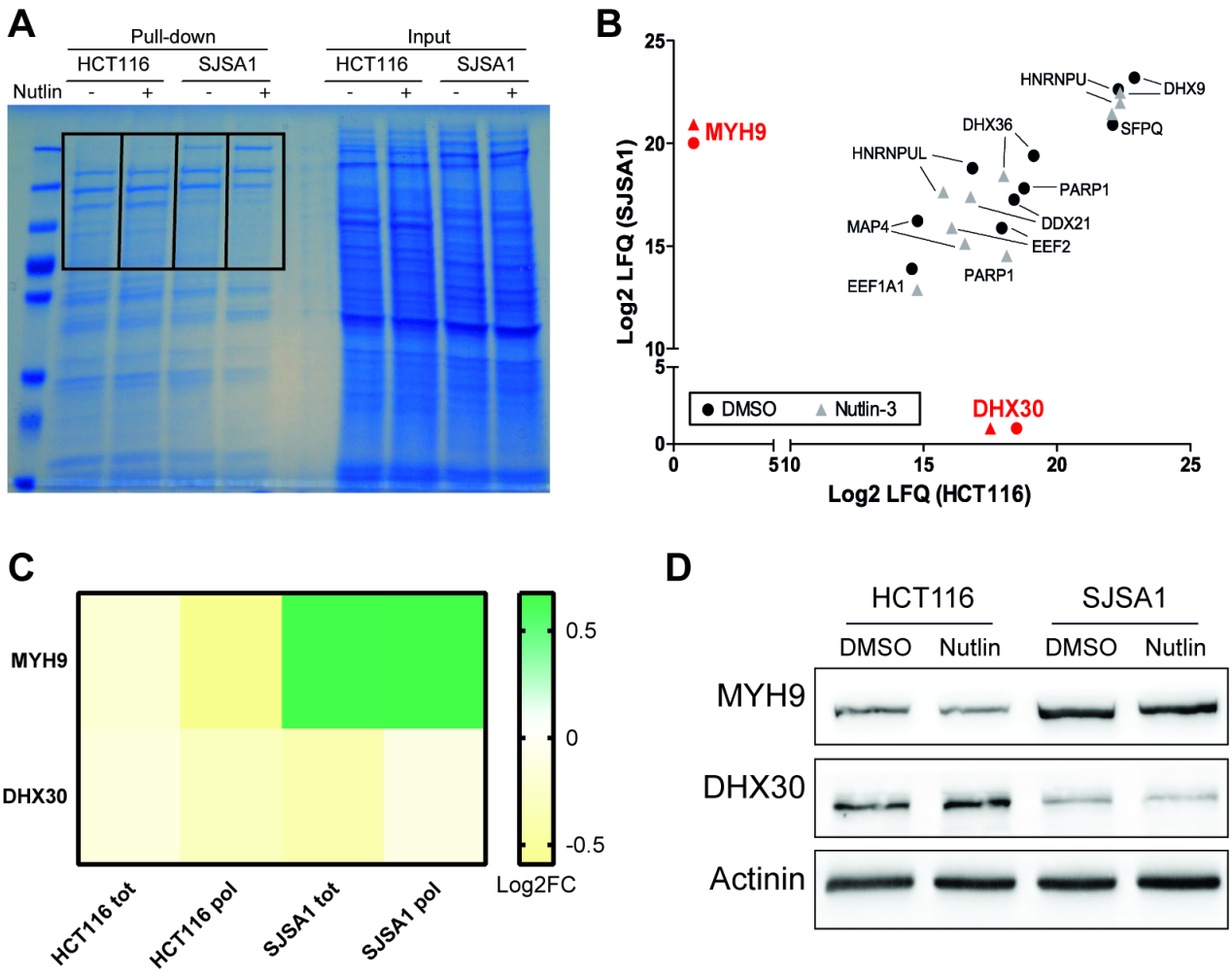


Figure 22: A) Example of an SDS-PAGE gel after Coomassie staining: the four lanes on the left were loaded with samples coming from the protein pull-down experiments while the lanes on the right are the inputs coming from the protein lysates. The areas surrounded by black rectangles were extracted from the gel and prepared for mass spectrometry 2) Proteins identified by mass spectrometry as interactors of the CG motif (LFQ = label-free quantification). Proteins showing a number of unique peptides ≥ 2 and a sequence coverage ≥ 10% were considered putative interactors 3) Fold change of the mRNA expression levels of MYH9 and DHX30 in HCT116 and SJSA1 at total (tot) and polysome (pol) level according to RNA seq data. 4) Expression levels of MYH9 and DHX30 in HCT116 and SJSA1 in control (DMSO) and Nutlin-3 (Nutlin) treated conditions. Actinin was used as a loading control.

The expression level of both *MYH9* and *DHX30* were evaluated from the RPKM coming from the RNA sequencing (Figure 22C) and at protein level via western blot (Figure 22D). Similarly to the PCBPs, also MYH9 and DHX30 do not show a sufficient fold change in mRNA expression to be considered DEGs, with MYH9 showing only a modest induction in SJSA1. Conversely, western blots showed a somehow antithetic expression pattern of the two proteins in HCT116 and SJSA1, with DHX30 being highly expressed in the first cell line and MYH9 in the latter. Consistent with the quantification coming from mass spectrometry, Nutlin-3 does not significantly affect MYH9 or DHX30 expression.

DHX30 is a poorly studied, ATP-dependent RNA helicase, belonging to the DExH/D family and characterized by the presence of the conserved DEAD-box motif (Asp-Glu-Ala-Asp). RNA helicases have been associated with various aspects of RNA biology, from transcription to editing, splicing and translation[88]. Specifically, DHX30 participates in a complex responsible for mtDNA transcription[89], it

is required for mitochondrial ribosome biogenesis[90] and impairs HIV infectivity in concert with other proteins[91,92]. Knock-out of the mouse homolog (*HelG*) leads to early embryonic lethality because of impairments in the central nervous system development [93]. Mutations in the *DHX30* gene have been recently associated with a severe form of developmental retardation[94]. Simsek et al [95] have found this protein as an interactor of the ribosome in mammals and mutants of the protein may impair global translation[94].

Given the relationship between DHX30 and translation, we decided to invest more in the study of this protein in conjunction with PCBP2. The latter was not identified in the original mass spectrometry experiment since the portion of the gel chosen for mass spectrometry sequencing did not match PCBP2's predicted molecular weight (≈ 39kDa). In a following experiment that I performed on the whole gel lane after protein pulldown (molecular weights ranging from approximately 15 to 250 kDa), PCBP2 was indeed identified as an interactor of the RNA probe both from HCT116 and SJSA1 protein extracts.

Considering the data obtained so far, it seems that the constitutive expression profile of the proteins we identified as candidates for binding the CG-motif is mostly responsible for the differential behaviour of SJSA1 and HCT116, suggesting that the specific protein landscape of a cell is responsible for shaping the response. Although we cannot exclude that some RBPs can be induced (or repressed) by p53 activation, contributing to the post-transcriptional control of mRNAs, we could not identify any candidate CG-motif-binding protein to be influenced by treatment.

Some recent reports identified MYH9 as a protein able to bind lncRNAs[96], but the lack of a known RNA binding domain suggests that their interactions are indirect. For this reason, we focused our attention more on elucidating the role of PCBP2 and DHX30, which seemed to be more promising candidates, given also the similar protein expression profile: higher levels in HCT116 and lower levels in SJSA1 (Figure 21B and 22D). In section 4.5.2 is reported an initial study on the possible role of MYH9 in our model of translation control.

4.2.3 - PCBP2 and DHX30: interactions and binding to the CG-motif

Erik Dassi performed the meta-analysis on the 3'UTRs combining eCLIP datasets; Sara Zaccara and I performed the pulldown, Co-IP and RIP.

Taking advantage of the public dataset of RBP target identified via eCLIP[97] in HepG2 and K562 cells as part of the ENCODE project[98], we decided to look at the position of DHX30 and PCBP2 binding sites compared to the CG-motifs in the 3'UTRs of the genes. As depicted in Figure 23A, DHX30 and PCBP2 binding sites are significantly overlapping on the translationally upregulated genes containing the CG-motif identified in SJSA1, supporting the idea that both proteins can bind the CG-motif. To confirm our hypothesis, a pulldown followed by immunoblot of the two proteins was performed starting from the same experimental conditions used for mass spectrometry. An RNA probe containing a mutated version of the consensus motif (mutated consensus: 5'-

CCCCAUGGAGAU-3') was used in these experiments to verify the specificity of the binding of both candidate proteins (Figure 23B). The pulldown confirmed the ability of both proteins to bind the consensus CG-motif in a sequence-specific manner: the introduction of a tri-nucleotide variation strongly compromises the binding of PCBP2 and DHX30, the latter resulting almost undetectable in SJSA1. As regards the wild type consensus, the binding seems to be stronger in HCT116 compared to SJSA1, which may reflect the different protein concentration in the two cell lines (as it is clearly visible in the *INPUT* panel). As expected, Nutlin-3 treatment marginally affects the binding ability of both proteins in each cellular background.

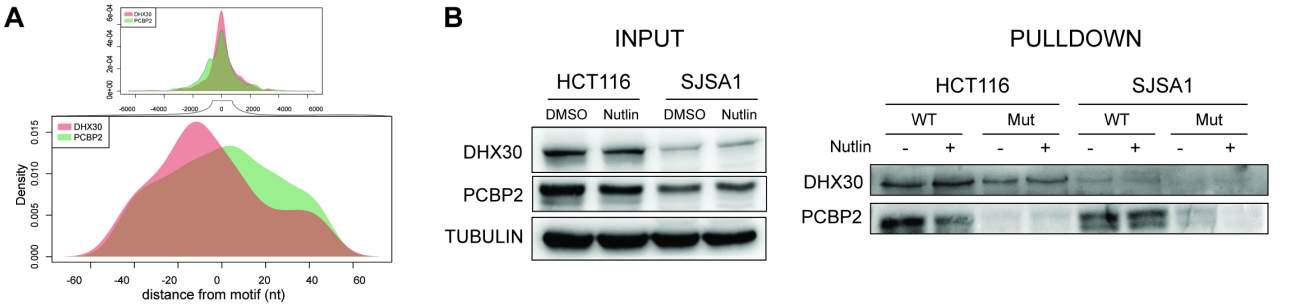


Figure 23: A) Binding sites distribution of PCBP2 and DHX30 identified by eCLIP on the 3'UTR of CG-motif containing mRNAs centred on the CG-motif itself. B) Pulldown experiment followed by immunoblot of DHX30 and PCBP2 using the wild type (WT) or the mutant (Mut) consensus CG-motif.

The data collected so far suggest that PCBP2 and DHX30 may be part of an inhibitory complex that acts in HCT116 by lowering the translation of mRNAs containing the CG-motif, ultimately preventing these cells to commit to apoptosis. This same inhibition is not achieved successfully in SJSA1, most likely because of the low expression levels of at least two components of this inhibitory complex (namely PCBP2 and DHX30), resulting in higher translation rates of CG-containing mRNAs as confirmed by qPCR and luciferase assay (see Figure 20 and 19 respectively).

To confirm the interaction between PCBP2 and DHX30 we performed a co-immunoprecipitation experiment using PCBP2 as bait protein. We chose to perform this experiment on HCT116 cell extracts since they express higher levels of both PCBP2 and DHX30. According to the western blot performed after immunoprecipitation, DHX30 interact with PCBP2 in an RNA-dependent manner (Figure 24). PABPC1 was used as a positive control, being a known PCBP2 interactor [99]. This data supports the idea that PCBP2 and DHX30 can act in concert on the regulation of mRNA translation by binding similar regions in the 3'UTR of target genes and physically interacting with each other in RNA-dependent fashion. Despite our trial to perform the reverse experiment (immunoprecipitation of DHX30 and revelation on PCBP2) we were unable to succeed, most likely because of the poor performance of the DHX30 antibody we had available and the proximity of PCBP2 molecular weight to the IgG chains.

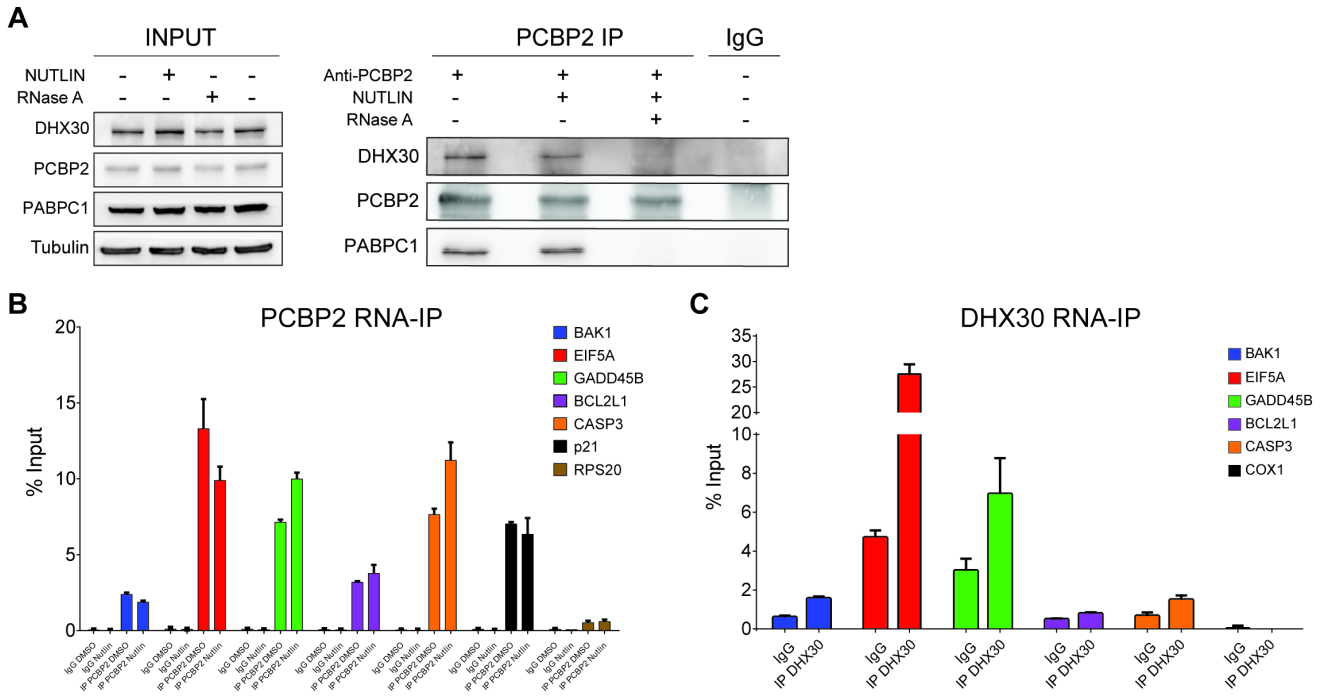


Figure 24: A) Co-immunoprecipitation of PCBP2 and DHX30. Input refers to the initial quantity of protein before immunoprecipitation using anti PCBP2 antibody; IgG refers to the anti-mouse immunoglobulin control; PABPC1 was used as a positive control being a known PCBP2 interactor and Tubulin as a loading control. B) RNA immunoprecipitation using PCBP2 antibody. Both DMSO and Nutlin-3 treated HCT116 were used C) RNA immunoprecipitation using DHX30 antibody in HCT116 cells.

To assess the real interaction between the proteins and mRNAs containing the CG-motif we performed UV-crosslink RNA-immunoprecipitation using antibodies against PCBP2 (Figure 24B) and DHX30 (Figure 24C). Also in these experiments, HCT116 were chosen because of the higher expression level of the proteins considered. The subset of mRNAs containing the CG motif (BAK1, EIF5A, GADD45B, BCL2L1 and CASP3) show a marked enrichment when PCBP2 was immunoprecipitated compared to IgG, both in DMSO and Nutlin-3 cells, similarly to p21 (that was used as a positive control). Conversely, RPS20 showed only a marginal enrichment over IgG and was considered as a negative control.

DHX30 antibody had worse performances in immunoprecipitating RNAs compared to PCBP2 antibody and for this reason, only the untreated condition was taken into consideration. Despite the lower degree of enrichment over IgGs, all the mRNAs containing the motif showed stronger binding compared to the negative control COX1, indicating the ability of DHX30 to bind CG-motif-containing mRNAs.

Taking into consideration all the results obtained so far, we can conclude that PCBP2 and DHX30 are interactors of the CG-motif that show a constitutive different expression level in HCT116 and SJSA1. The binding sites of the proteins are in proximity and partially overlapping, probably because the protein interact in an RNA-dependent manner and are thus part of the same complex. We have also shown that DHX30 and to a higher extent PCBP2 can strongly immunoprecipitate mRNAs containing the CG-motif.

4.3 – Modulating the responses in HCT116

4.3.1 - PCBP2 and DHX30 silencing

Experiments presented in this section were performed by Sara Zaccara and me.

Having demonstrated the ability of PCBP2 and DHX30 to bind the CG-motif consensus in-vitro (using a probe) and in vivo (by binding the 3'UTR of target mRNAs) we decided to modulate the proteins' expression levels to see the effects on mRNA translation. Since the most desirable phenotype in a p53-activating therapy in cancer patients would be cell death and not cell cycle arrest, we decided to downregulate the expression of either one or the other candidate protein in HCT116 by means of short hairpin (sh) RNA.

We used lentiviral vectors produced in HEK293-T cells starting from the pLK0.1 plasmid to transduce HCT116 parental cells. The pools of transduced cells were puromycin-selected before proceeding to single clone selection and testing of the silencing efficiency. To maximize the chances of having a successful protein knock-down, we employed three different shRNA-containing vectors to deplete DHX30 and 2 for PCBP2. As a control, we used a scrambled, non-targeting shRNA sequence. All the plasmids were provided by prof. Joaquin Espinosa and the Functional Genomics Facility (University of Colorado, USA). Plasmids identifiers and sequences are reported in Section 11.4.

All the experiments presented in the following sections were performed on the clones showing the highest level of silencing for either PCBP2 or DHX30, as detected by western blot (data not reported). To confirm that the results could imply a general effect and not a clone-specific phenotype, some selected experiments were performed also on other clones showing similar degrees of silencing.

The first thing we tested was the ability of PCBP2 to bind the CG-motif consensus in a context of DHX30 depletion and vice versa using the protein pull-down strategy previously presented (section 4.2.3). This experiment was done to support the idea that the proteins interact when binding of the RNA probe (and possibly the 3'UTR).

By looking at the inputs of the pulldowns, depletion of each protein does not influence the expression level of the other (Figure 25). However, DHX30 seems to rely on PCBP2 for efficiently binding the probe containing the WT motif consensus (Figure 25A) while PCBP2 binding seems to be independent of DHX30 presence (Figure 25B).

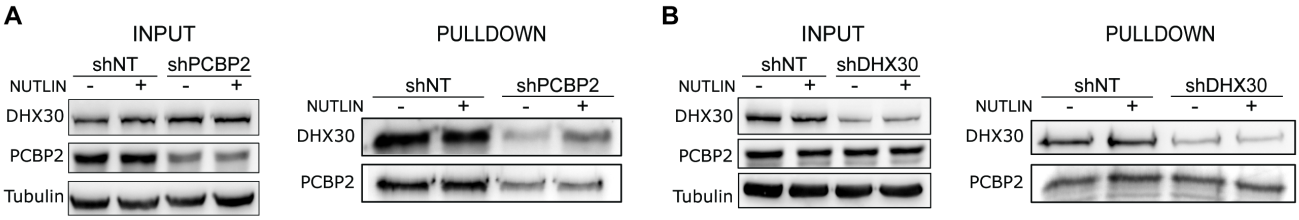


Figure 25: Protein pulldown using the wild type consensus CG motif in PCBP2 (panel A) and DHX30 (panel B) depleted HCT116 cells. shNT refers to the scrambled shRNA-transduced cells, that were used as controls.

We also tested the impact of the silencing using the luciferase reporter assay described previously (see Figure 19, section 4.1.3), after 12 and 24 hours of Nutlin-3 treatment (Figure 26A and 26B

respectively). Cells transduced with scrambled shRNA (shNT) show no induction of the reporter or a modest decrease similarly to parental HCT116 (see Figure 19B). On the other hand, both shPCBP2 and shDHX30 clones show an enhancement in the luciferase reporter, at 12 and 24 hours respectively, somehow resembling SJSA1 behaviour in the same experimental conditions. Surprisingly shPCBP2 reporter decreases its activity at 24 hours. Also in this case, relative luciferase units are normalized over the *luc2* mRNA level to remove any bias given by a transcription-dependent difference of its mRNA level.

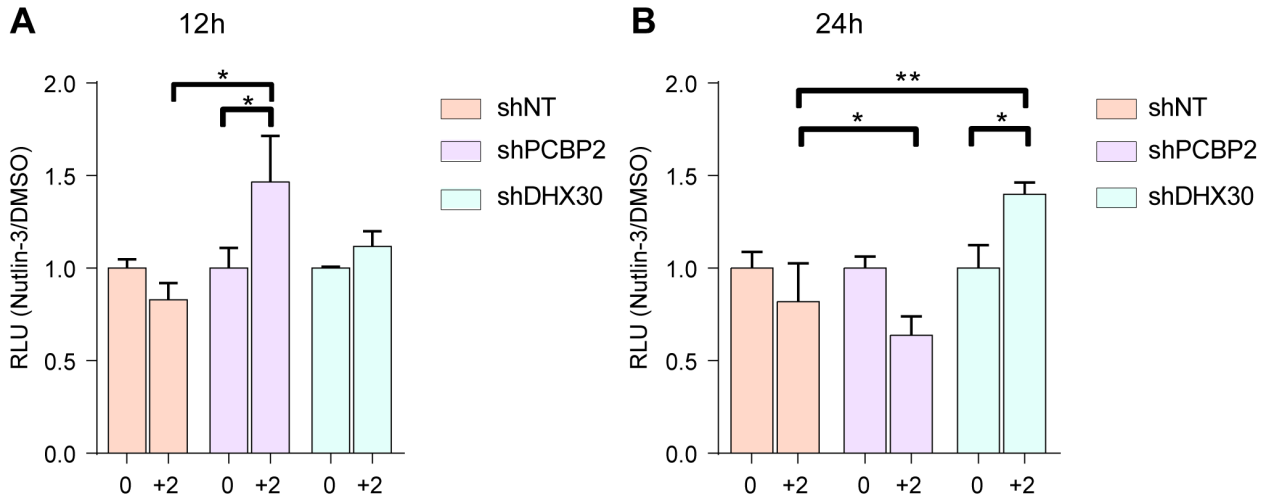


*Figure 26: Luciferase reporter assay results on HCT116 control clone (shNT) or depleted for PCBP2 (shPCBP2) or DHX30 (shDHX30) expression after 12 (panel A) or 24 (panel B) hours after Nutlin-3 treatment. 0 refers to the βGlobin 3'UTR; +2 refers to the βGlobin 3'UTR added with two copies of the CG-motif consensus. * = p-value < 0.05; ** = p-value < 0.01*

These preliminary observations prompted us to test whether the silencing of the proteins could have a significant impact on the translation of CG-motif mRNAs. For this reason, we decided to analyse the translatome of these clones after induction of p53 on a global scale using RNA sequencing.

4.3.2 - HCT116 shPCBP2 and shDHX30: RNA sequencing

Polysome profiling and library preparation were performed by Sara Zaccara and me. NGS facility performed the sequencing. Data analysis was performed by Erik Dassi and Sara Zaccara

In order to have a more comprehensive idea of the impact of PCBP2 and DHX30 silencing on the translation regulation, we performed RNA sequencing on the mRNAs pooled from the polysome fractions of the cells depleted for PCBP2 or DHX30 after 12 hours of 10μM Nutlin-3 treatment. A schematic view of the experimental setup and the validation experiments is reported in Figure 27.

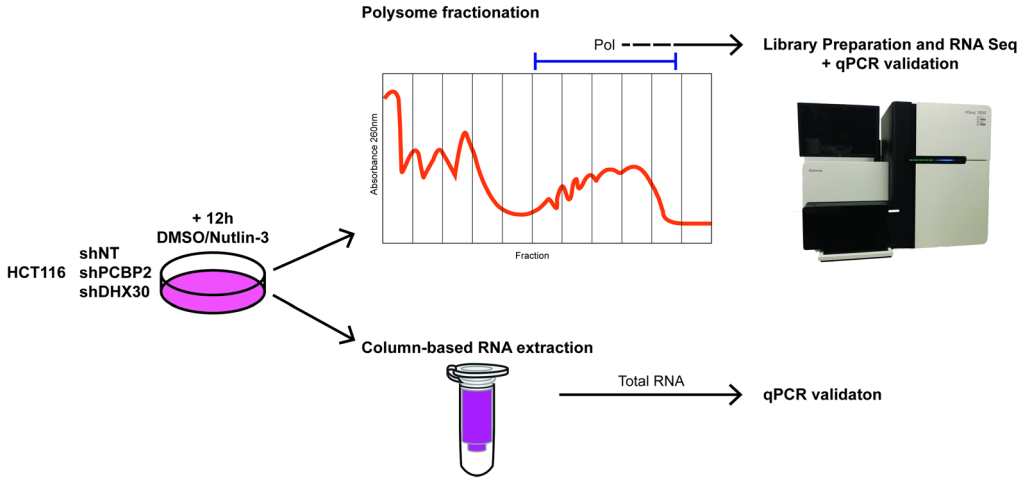


Figure 27: Experimental approach used to characterize HCT116 clones depleted for PCBP2 and DHX30

In this case, differentially expressed genes were defined as those having adjusted p-value < 0.05 and a $\log_2$ Fold Change (Nutlin-3/DMSO) > 0 (upregulated) or < 0 (downregulated). We decided to keep this $\log_2$ FC threshold to account also for silencing-dependent gene expression variations that are poorly or not treatment dependent. Figure 28 reports the Venn diagrams of the polysome-associated mRNAs in HCT116 control cells (shNT), silenced for PCBP2 (shPCBP2) or for DHX30 (shDHX30) as detected by RNA sequencing, either upregulated (panel A) or downregulated (panel B). The results of the RNA sequencing analysis are reported in Section 11.1.

The translatome of HCT116 seems to be heavily affected by the silencing of either PCBP2 or DHX30: silenced clones show a significant number of silencing-specific altered genes both down- (139 for shDHX30 and 157 for shPCBP2) and up-regulated (221 and 157 for shDHX30 and shPCBP2 respectively), while shNT show only a marginal number of exclusive genes (21 up and 5 down respectively). Notably, 189 and 302 genes are commonly upregulated and downregulated in both silencing conditions, but not in the control clone, possibly indicating a common role of the proteins on the regulation of some mRNAs.

Similarly to what was previously done in parental HCT116 and SJSA1, we performed ontology analysis on gene lists using Metascape (Figure 29). Despite the different number of genes present in the lists, no molecular function or ontology was altered in a silencing-dependent manner in the upregulated genes, with terms referring to p53 pathway and apoptosis enriched in all the conditions (panel 29A). By looking at the downregulated genes, ontologies referring to DNA repair and cell cycle arrest were mostly represented in the silenced clones (panel 29B), but also in this case, a clear phenotype difference between control and silenced HCT116 clones could not be spotted.

The higher divergence in the ontology analysis of the translationally downregulated genes between shNT and shDHX30/shPCBP2 clones might be dependent on the high number of genes that are shared between the two silenced clones (302 genes, see Figure 28B), leaving only 5 genes being downregulated only in the shNT clone.

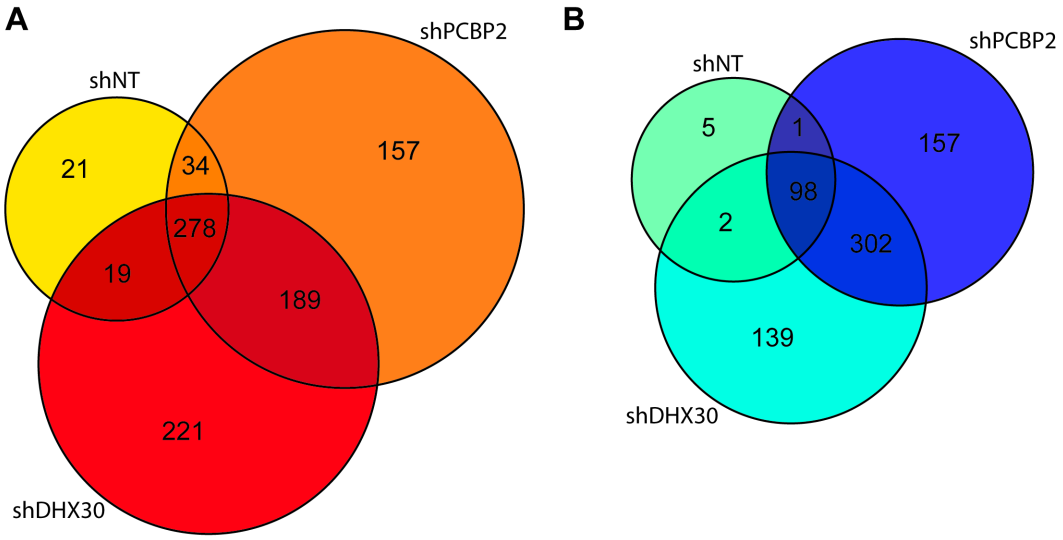


Figure 28: Upregulated (log2FC > 0; panel A) or downregulated (log2FC <0; panel B) genes in the polysome fractions of HCT116 control (shNT) and silenced for PCBP2 (shPCBP2) or DHX30 (shDHX30)

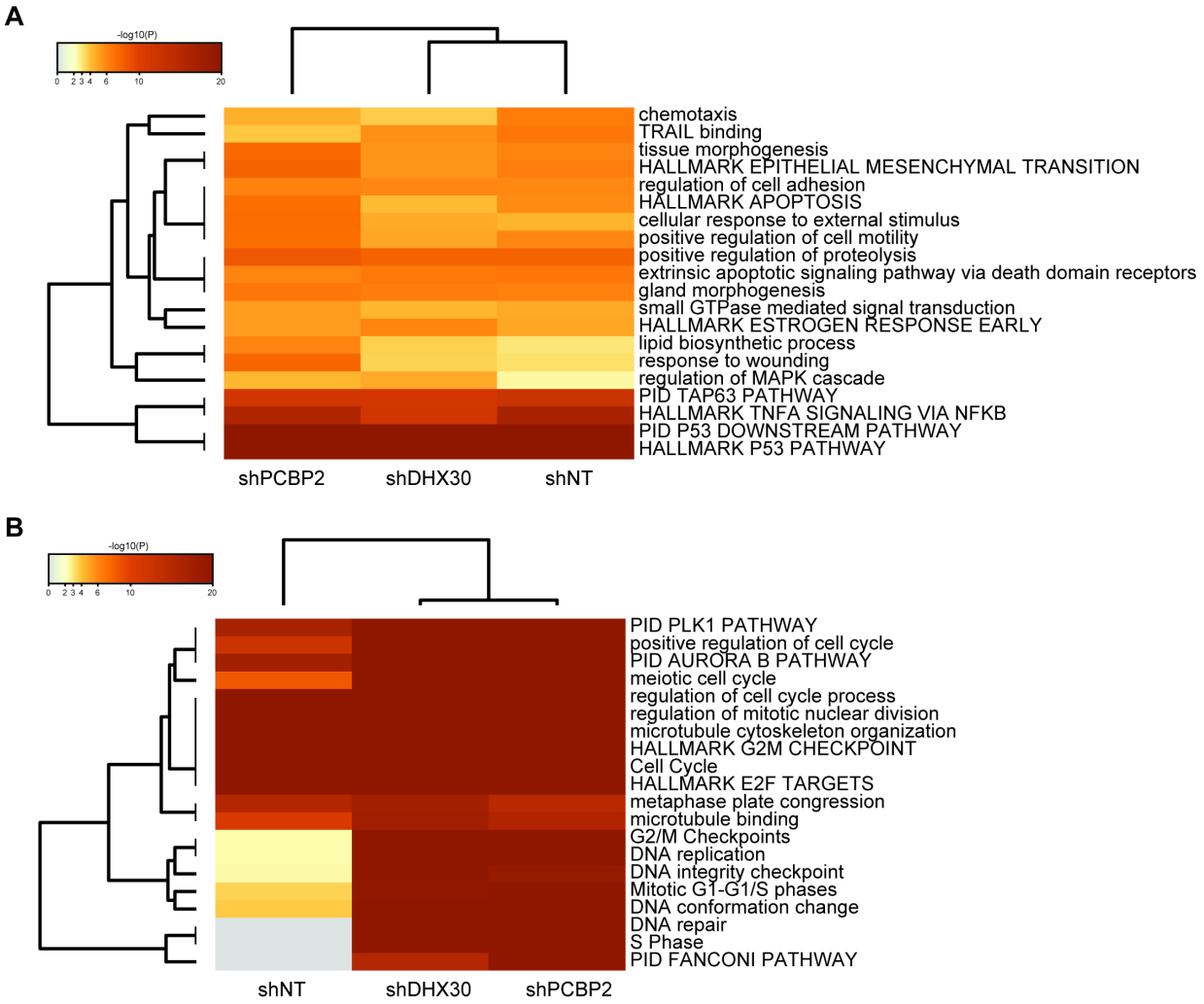


Figure 29: Comparative Metascape analysis of the upregulated (panel A) and downregulated (panel B) polysome-bound genes in HCT116 shNT, shPCBP2 and shDHX30 clones

Since Metascape does not take into consideration the relative expression of the terms in the lists but only the relative enrichment for pathways and molecular functions, we supposed that differences induced by the silencing could not be effectively appreciated, given the identical cellular background of the silenced clones. Indeed, the previous comparison between HCT116 and SJSA1 translatomes (Figure 15) revealed a much lower level of overlap between the gene lists. In this context, the number of genes that are shared between the control and the silenced clones is relatively high (278 genes upregulated and 98 genes downregulated) and the effect of the silencing could impact the levels of expression of the mRNAs more than altering the translational program, despite the relatively high number of terms uniquely present in either one or the other silenced clones. For these reasons we decided to apply Gene Set Enrichment Analysis (GSEA)[100] on the upregulated gene lists of control, PCBP2 and DHX30 silenced clones. Contrary to Metascape analysis, GSEA takes into consideration the level of expression (log_2 FC) of the terms and ranks the results also depending on the expression levels of the genes involved in a specific pathway and not only on the relative enrichment of a molecular signature over the reference. In this way, differences given by the contribution of the subtle alteration of many different genes belonging to a common pathway can be highlighted. The top upregulated and downregulated pathways resulting from GSEA on the HCT116 clones translationally modulated genes are reported in Figure 30.

Looking at the downregulated pathways, both the control and the silenced clones show terms referring to DNA repair, cell cycle and cell division, confirming what was previously obtained in Metascape and meeting the expectation of p53-dependent responses.

shPCBP2 shows pathways related to metabolism and signalling as the most upregulated, and in part also the control clone. Conversely, shDHX30 is the only clone showing a clear p53 pathway upregulation and terms referring to apoptosis, suggesting a more determinant role of this protein in the control of apoptotic mRNAs translation. According to our model, the enhanced polysome association of mRNAs of genes controlling apoptosis could be CG-motif dependent, similarly to what occurs in SJSA1. Indeed, the luciferase assay result (Figure 26) suggests that this might be the case also in HCT116 shDHX30 clones.

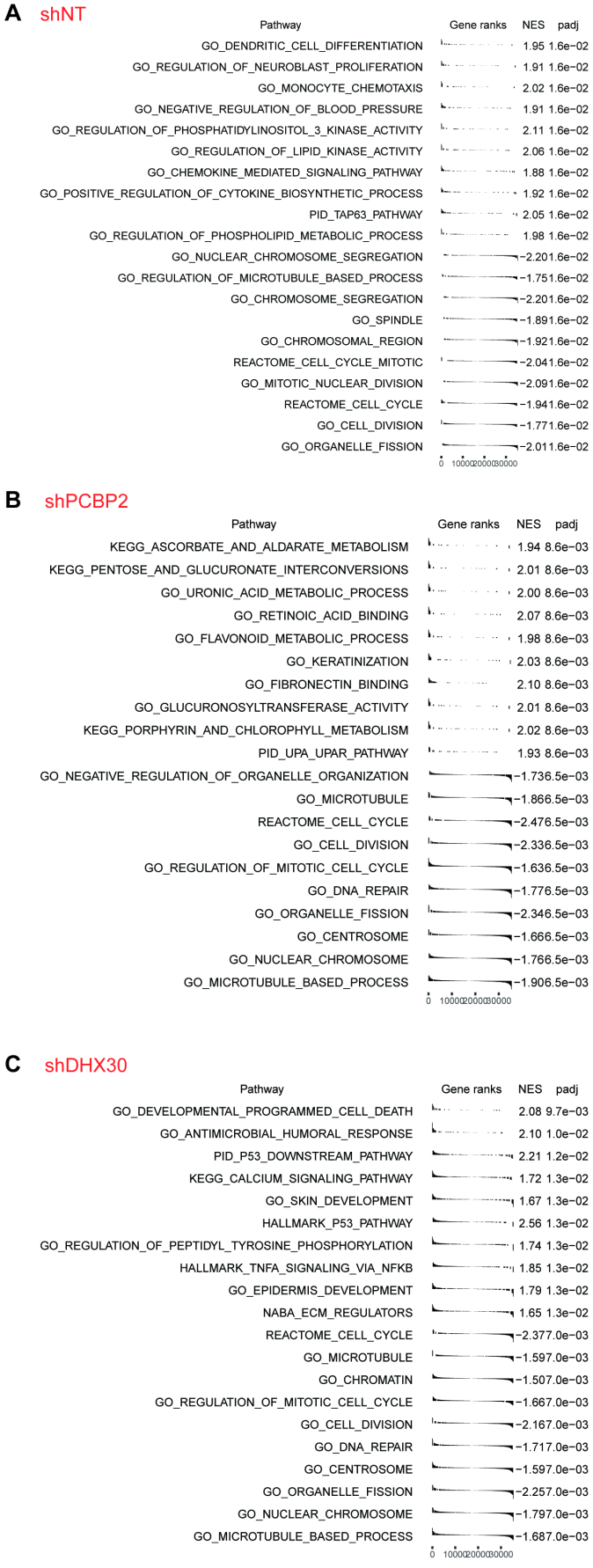


Figure 30: GSEA analysis on the polysome upregulated genes in control shNT(panel A), shPCBP2 (panel B) and shDHX30 (panel C) HCT116 clones. NES = normalized enrichment score; padj = adjusted p-value

In line with this, the number of polysome-bound genes whose function is implicated in apoptosis is higher in shPCBP2 and shDHX30 samples compared to control cells (Figure 31A).

In order to test whether the high representation of apoptosis-related genes was dependent on a motif in the UTR regions of the polysome-bound genes, Weeder was applied to search for enriched motifs in the 5' and 3'UTRs of the genes identified in HCT116 clones by RNA sequencing. Some new motifs were identified in the UTRs of all the polysome-bound upregulated mRNAs for all the clones (Figure 31B, motifs are reported in the lower part of the heatmap), but only those identified in the 3'UTR were somehow resembling the CG-consensus. By comparing the newly discovered 3'UTR motifs with the top three motifs originally identified in SJSA1 (Figure 31B, right part of the heatmap), we noticed a high level of correlation for the new motifs identified in shDHX30 clones (Figure 31B, higher Pearson's correlation coefficients indicate higher similarity between the CG-motif identified in SJSA1 and those enriched in the 3'UTR of polysome-bound mRNAs in HCT116 clones).

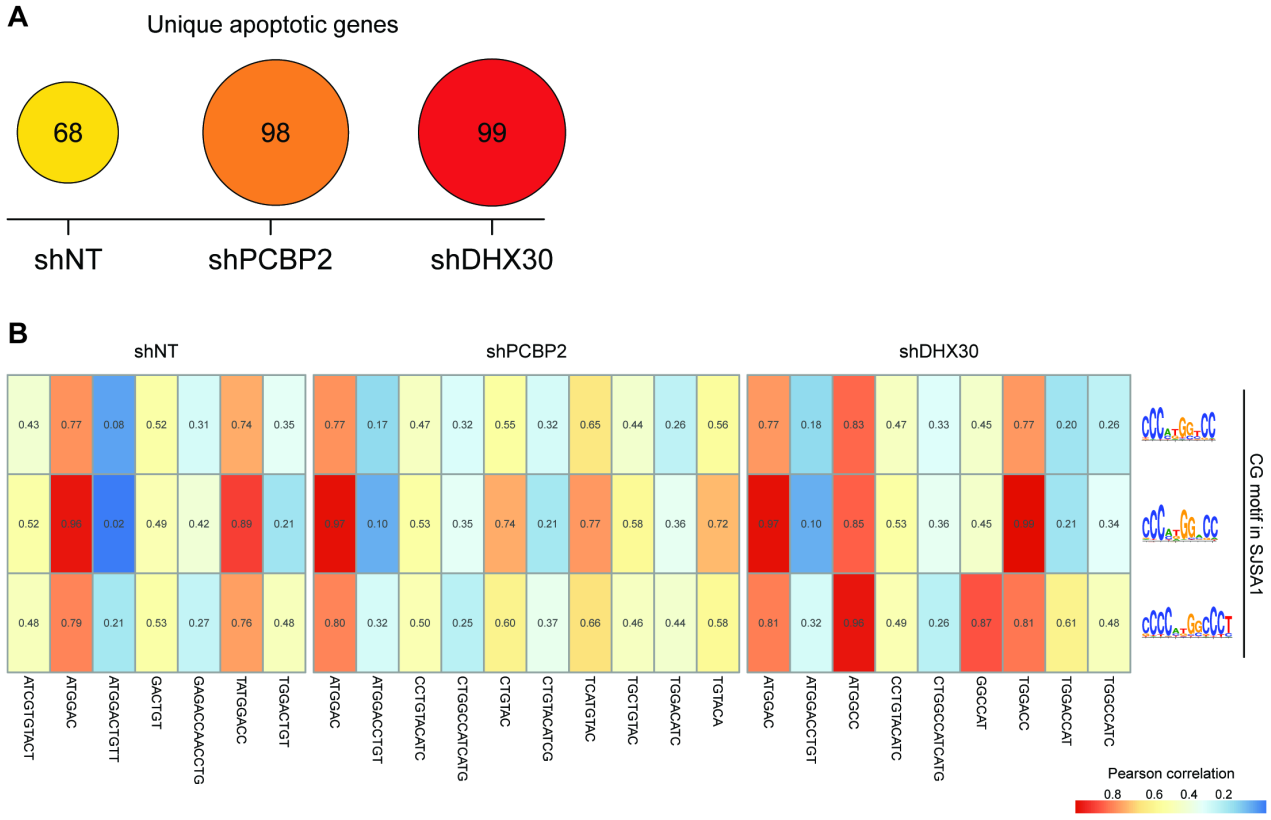


Figure 31: Panel A: Number of unique genes involved in apoptosis in HCT116 control (shNT) or silenced for PCBP2(shPCBP2) and DHX30 (shDHX30). Panel B: Enriched motifs identified in the 3'UTR of the polysome-bound mRNAs in HCT116 clones (lower part of the heatmap) compared to the top 3 motifs originally identified in the 3'UTRs of the translationally enhanced genes in SJSA1. Results are ranked according to the Pearson correlation of the position frequency matrices (PFM) of enriched motifs.

According to these data, it seems that DHX30 silencing has a stronger impact compared to PCBP2 in promoting translational regulation of mRNAs containing a CG-like motif in the 3'UTR. Moreover, mRNAs having apoptotic-related functions are more polysome-associated in DHX30 clones compared to shNT and shPCBP2 clones. Silencing of either DHX30 or PCBP2 has a strong impact on the polysome-bound mRNAs composition, both upregulated and downregulated after the

treatment with Nutlin-3, suggesting that both proteins have a strong influence on cells' translation program.

4.3.3 – RNA sequencing validation: qPCR and western blot

qPCR and western blots were performed by me and Sara Zaccara. RT2 qPCR arrays were designed by me and Alberto Inga

To validate the results obtained from the RNA sequencing we decided to test a panel of genes using RT-qPCR. We opted for a custom-made gene array (RT2 qPCR Primer Assay, QIAGEN) containing 23 different primer sets to amplify an equal number of genes. Targets were selected for being:

- polysome upregulated in shPCBP2 and/or shDHX30 gene lists (Figure 32 and 33);
- genes originally identified in SJSA1 translatome containing the CG-motif, independently from being identified as pol UP on the shPCBP2/shDHX30 RNA-sequencing lists (Figure 34);
- direct p53 targets (Figure 35).

In order to verify if the silencing of either PCBP2 or DHX30 had an impact only in translation, we also analysed the gene expression at the transcriptome level by looking at the total mRNA extracted from matching clones and treatment (DMSO or Nutlin-3 for 12hours, as reported in Figure 27). Besides the RT2 qPCR Primer Assay, the expression of some of the targets belonging to the original SJSA1 CG-motif-containing list was tested by qPCR using the primer sets used previously (see Figure 20 and section 11.6).

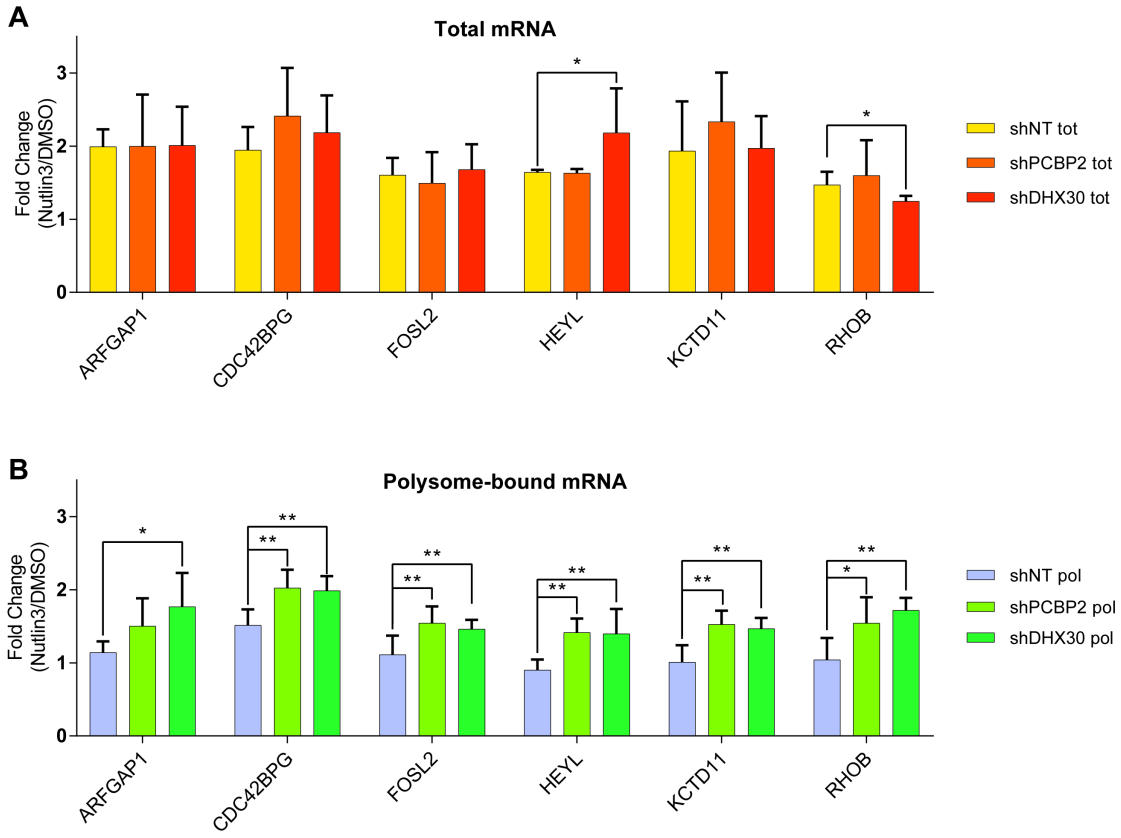


*Figure 32: qPCR results on some selected target identified in the shPCBP2 pol-UP genes. Panel A: expression at total mRNA level; Panel B: expression at polysome level. * = p-value < 0.05; ** = p-value < 0.01*

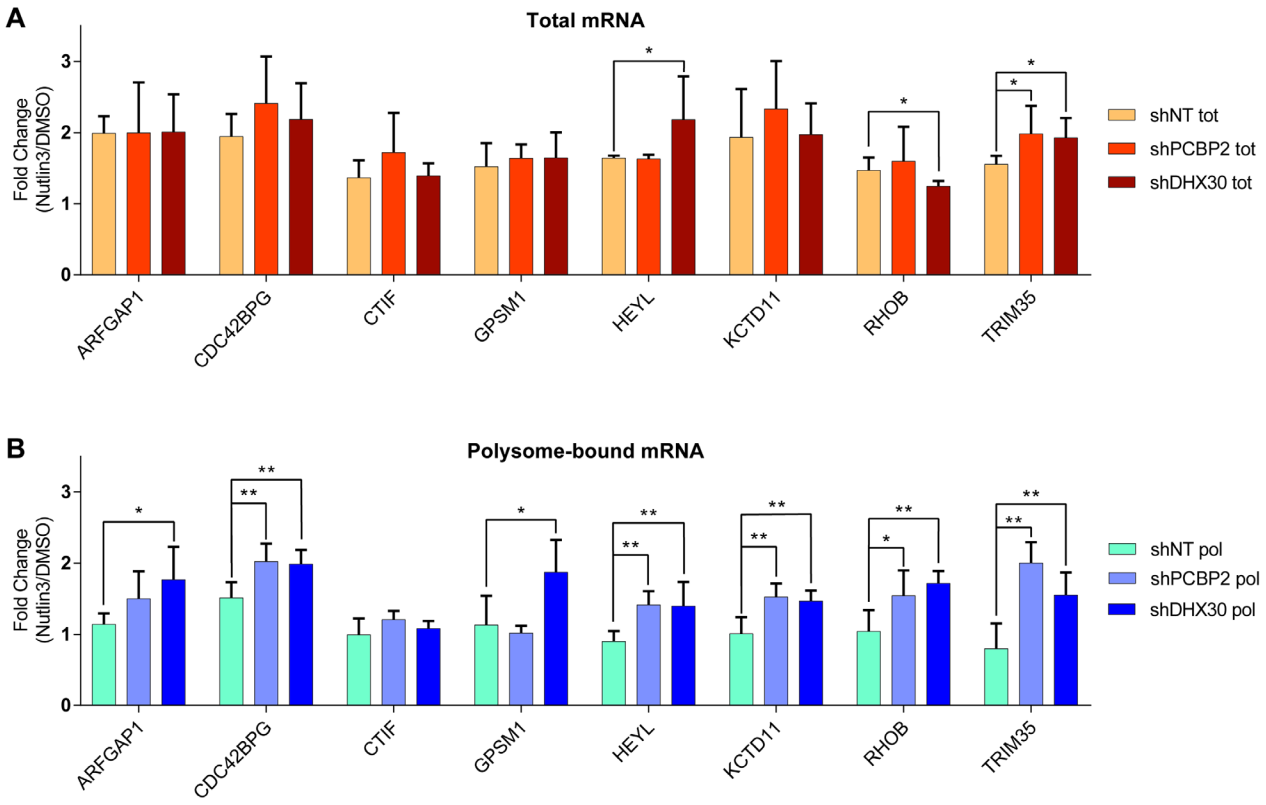


Figure 33: qPCR results on some selected target identified in the shDHX30 pol-UP genes. Panel A: expression at total mRNA level; Panel B: expression at polysome level. * = p-value < 0.05; ** = p-value < 0.01

With few exceptions, RT-qPCR data confirmed the results of the RNA sequencing analysis and the expectations according to the model: while most of the selected genes do not show a significant variation at the total level (Figure 33A and 33B), most of them are modulated at a polysome level (Figure 32B and 33B), confirming the role of DHX30 and PCBP2 in modulating mostly translation-linked responses.

In opposition to RNA sequencing data, *ARFGAP1* and *CTIF* modulation at polysome level were not confirmed for shPCBP2 and shDHX30 clones by qPCR. Notably, the expression of targets identified only in shPCBP2 (*FOSL2*) or in shDHX30 (*TRIM35*) on the RNA sequencing is significantly enhanced in both clones, possibly because the two proteins may be part of the same complex and/or share the binding site sequence, as previously shown in Figure 23. Surprisingly, HEYL and RHOB expression were changed also at the total RNA level in shDHX30 clones and in both clones for TRIM35 compared to shNT control.

Figure 34 reports the expression of some of the targets belonging to the SJSA1 translationally enhanced DEGs containing at least one instance of the CG-motif. Despite the high variability of expression between control and silenced clones at a total level, only *FKBP8* was statistically significant upregulated in the silenced clones (Figure 34A). On the other hand, the expression of *BAK1*, *BCL2L1*, *CAPNS1*, *EIF5A*, *NOL3* and *TAF10* was significantly enhanced in shDHX30 translatome, with *TAF10* being translationally upregulated also in shPCBP2 clones (Figure 34B) compared to the shNT control. Of all the targets considered in this set, only BCL2L1 was identified

in the shDHX30 polysome upregulated mRNAs, confirming RNA-seq data. These results suggest that DHX30 might have a predominant role in the expression of mRNAs containing the CG-motif with respect to PCBP2. As a general observation, the enhanced polysome association of CG-motif-containing mRNAs follows DHX30 perturbation.

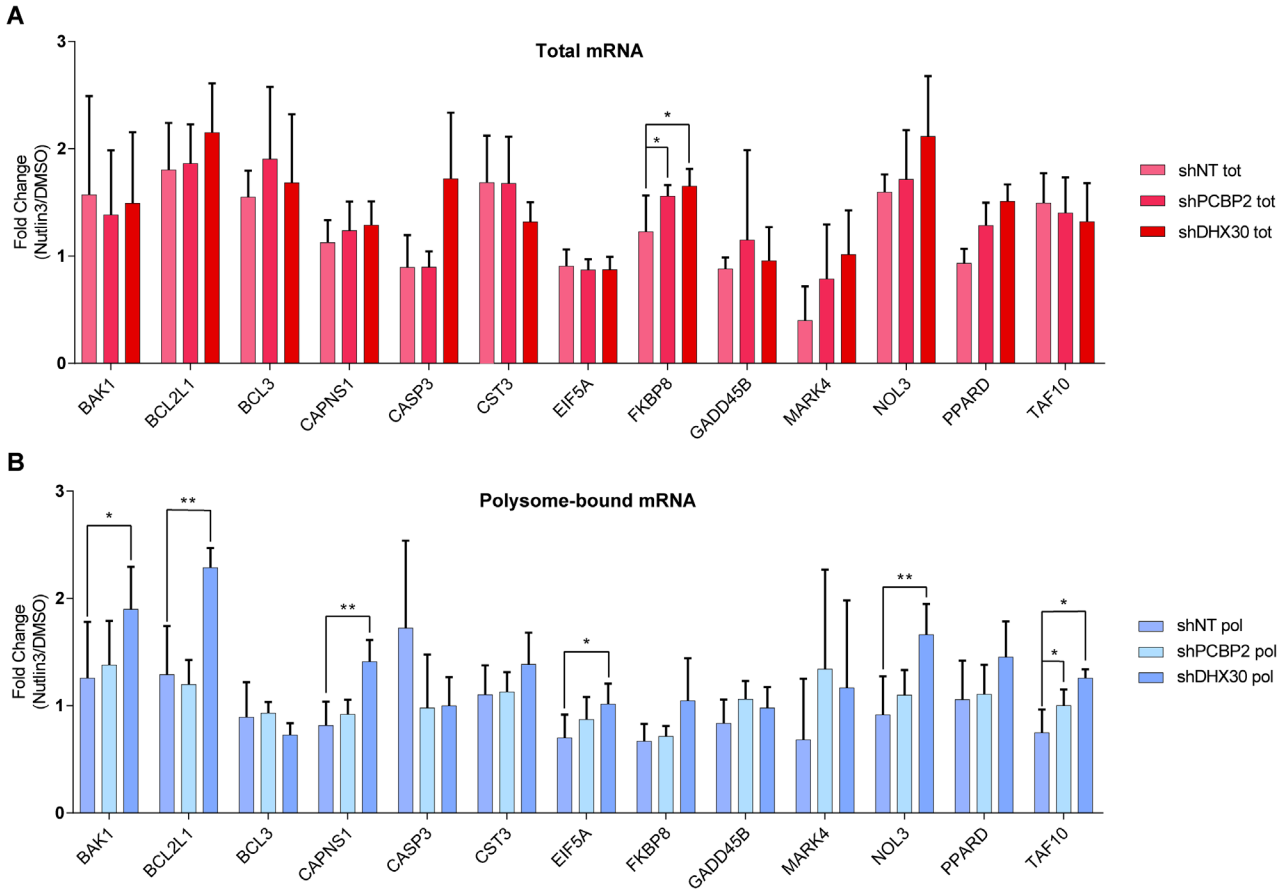


Figure 34: qPCR results on some CG-containing targets originally identified as polysome-upregulated DEGs in SJSA1. Panel A: expression at total mRNA level; Panel B: expression at polysome level. * = p-value < 0.05; ** = p-value < 0.01

Finally, we analysed the tendency of a few p53-target genes at total and polysome level (Figure 35A and 35B respectively). As expected, none of them shows enhanced transcriptional or translational boost in silenced clones compared to shNT control, with TNFRSF10A being slightly translationally downregulated in shDHX30 clones.

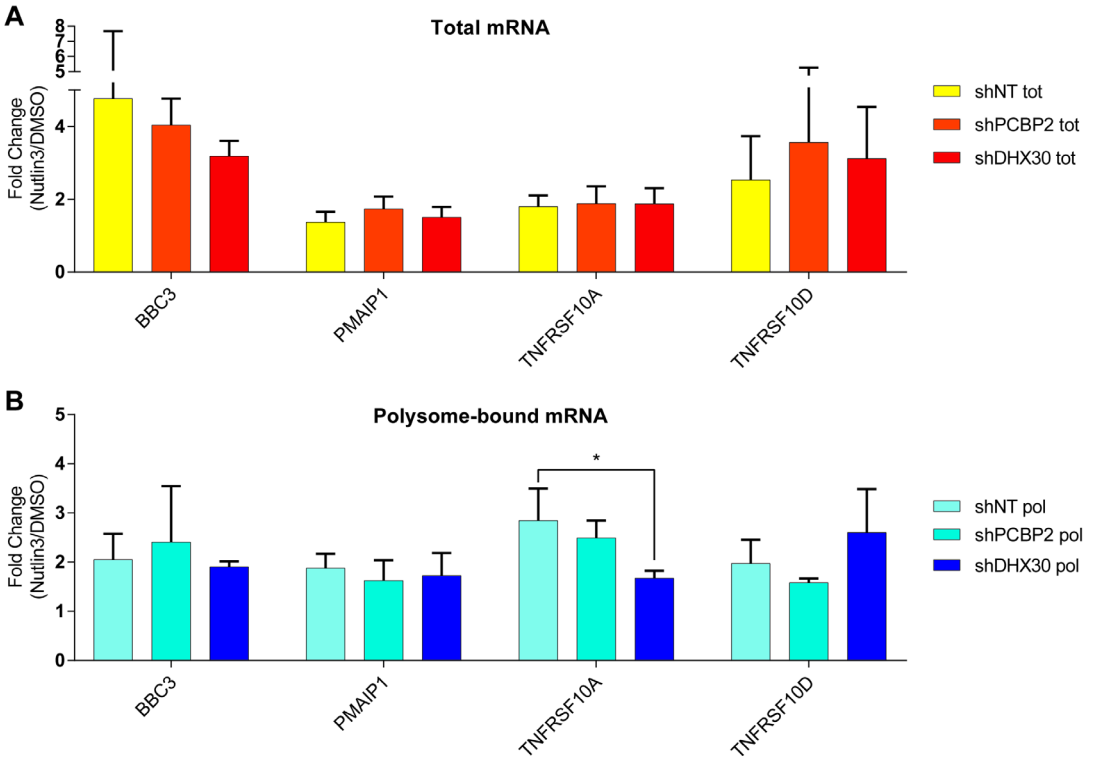


Figure 35: Expression level of some p53 target genes as detected by qPCR. Panel A: expression at total mRNA level; Panel B: expression at polysome level. * = p-value < 0.05; ** = p-value < 0.01

Collectively, the qPCR analysis confirmed the RNA sequencing and the expectations according to our model: polysome-associated mRNAs identified by Illumina sequencing show enhanced polysome association in either one or both knock-down clones, that is usually not matched by increased total mRNA level. Some of the targets containing the CG-motif that were not identified from with the RNA-seq show enhanced polysome association when a more sensitive technique is used. These observations confirm the model derived from the luciferase assay results, where DHX30 (and slightly also PCBP2) have a determinant role in the polysome loading of CG-motif-containing mRNAs, with a consequently enhanced expression of the protein.

Following this idea, we tested the protein expression of some of the genes containing the motif identified in SJSA1. Also in this case, we focused on 24 and 48 hours after the treatment and not at 12 hours like in the RNA seq and qPCR hypothesizing that the effect on protein levels could be noticed later compared to mRNA polysome loading (Figure 36). BAK1 expression is slightly induced at 24 hours after Nutlin-3 treatment in shPCBP2 and shDHX30 and more evidently upregulated at 48 hours in shDHX30 clones, reflecting the expectations coming from the qPCR. In disagreement with qPCR data, EIF5A protein expression is very unstable and no clear upregulation can be seen at 24 or 48 hours. Caspase3, on the other hand, seem to be unaffected by treatment and silencing, as predicted by qPCR.

We also checked whether PCBP2 and DHX30 perturbation could have an impact on the induction of apoptosis in this cellular context by looking at PARP cleavage. DHX30 silencing seems to favour apoptotic responses in HCT116 cells after p53 activation, as a band corresponding to the cleaved

PARP appeared only in shDHX30 clones treated with Nutlin-3, somehow recapitulating SJSA1 behaviour after the same treatment.

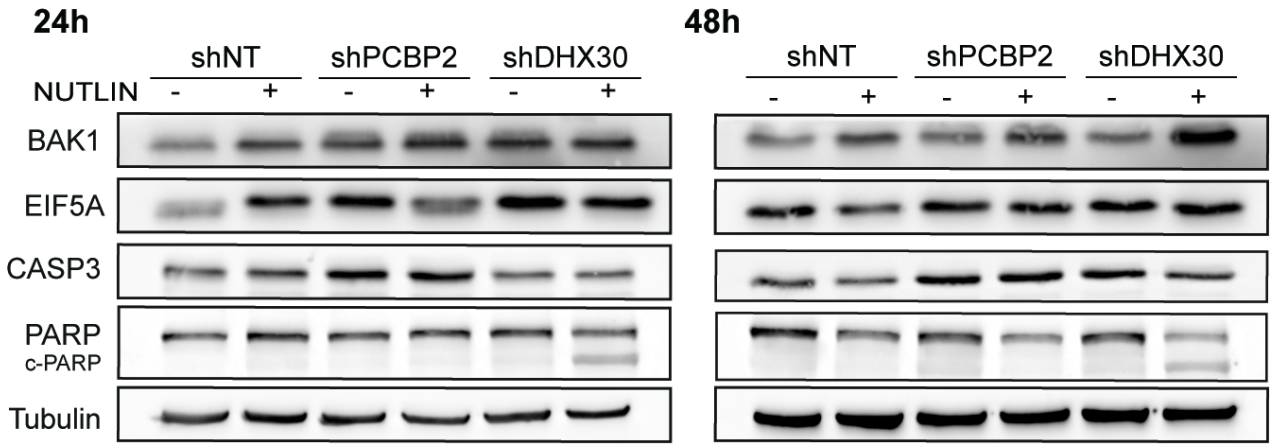


Figure 36: Western blot results of the expression of some proteins containing the CG-motif in control (shNT) and PCBP2 or DHX30 silenced cloned after Nutlin-3 treatment for 24 or 48 hours. Parp and its cleaved form were detected to check for active apoptosis in these clones. Tubulin was used as a loading control

4.3.4 - Apoptotic response of HCT116 knockdown clones

Experiments were performed by me and Sara Zaccara. Sample acquisition at FACS was performed by the CASF facility.

PARP cleavage indicates that shDHX30 clones are more sensitive to Nutlin-3 treatment, proceeding with the apoptotic cascade after the initial cell cycle arrest. To quantify the extent of this response we performed annexin V and propidium iodide (PI) double staining followed by FACS analysis after 24 or 48 hours of Nutlin-3 treatment on the control and PCBP2- or DHX30-silenced HCT116 clones. An example of the FACS analysis of the different clones at 48h is reported in Figure 37A while Figure 37B reports the sum of FITCH-annexin V-positive cells (quadrants Q2 and Q4 of the plots reported in Figure 37A).

Contrary to the western blot for PARP cleavage, HCT116 shDHX30 clones do not show increased apoptotic behaviour as compared to controls after 24 hours of treatment. Conversely, at 48 hours it is possible to notice a significant increase of the annexin V positive population in shDHX30 cells treated with Nutlin-3, in line with the expectations coming from the western blot analysis of PARP. These clones reach ≈30% of apoptotic cells, which is almost twice the amount in shNT controls. Also in this context, PCBP2 role seems to be marginal compared to DHX30 in the induction of apoptosis: despite a slight increase in the annexin V positive population at 48 hours, shPCBP2 clones do not reach statistical significance compared to shNT controls with matching treatment.

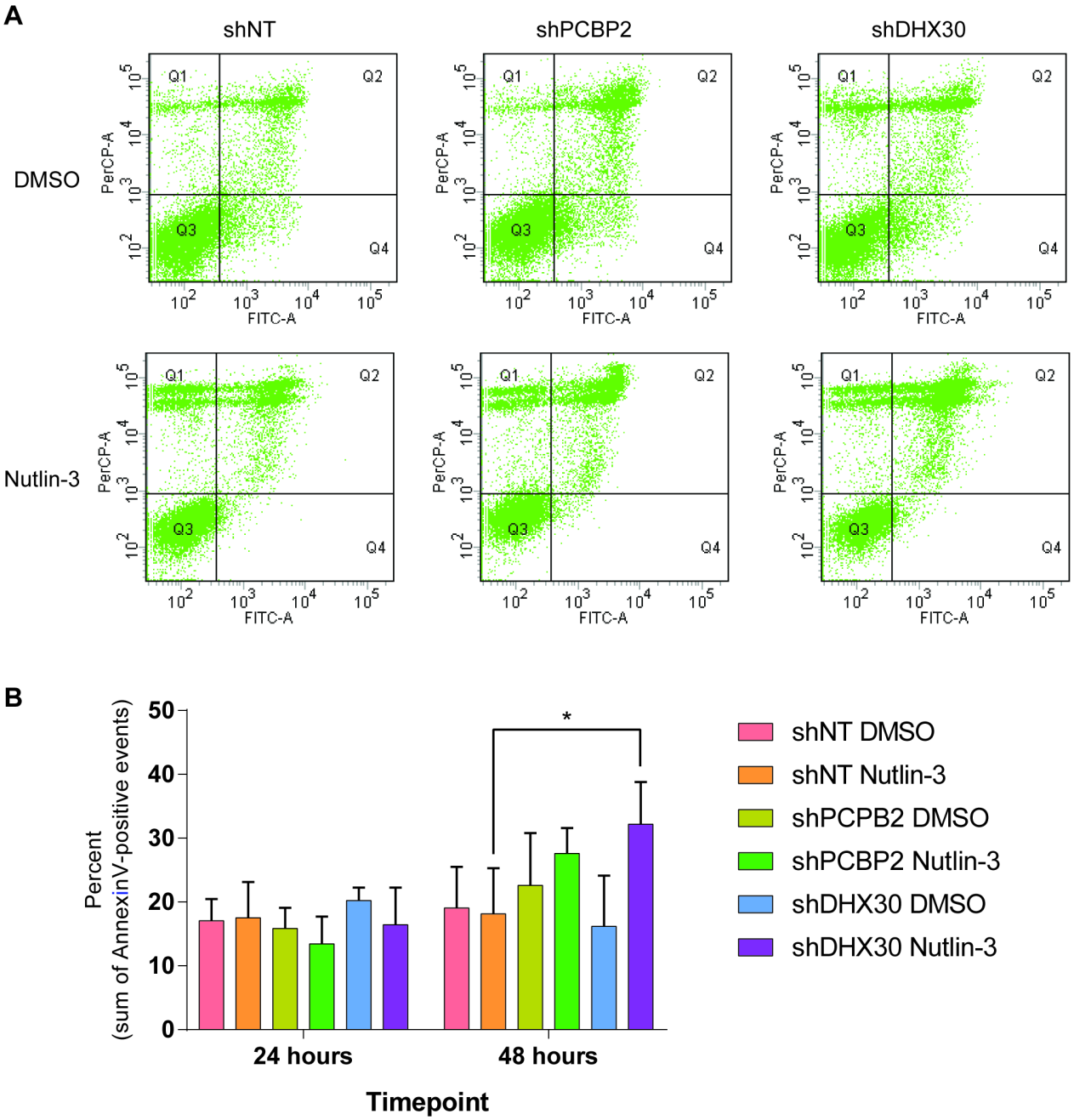


*Figure 37: Annexin V positivity of HCT116 clones. Panel A: example of one replicate after 48 hours of DMSO or Nutlin-3. PerCP indicates the propidium iodide intensity while FITC-A indicates annexin V fluorescence. Panel B: summary of the annexin V positive cells in the different clones at 24 or 48 hours after the treatment. * = p-value < 0.05*

From this analysis we finally conclude that DHX30 levels seem to be predictive of the response of cell lines to non-genotoxic p53 activation with Nutlin-3: cell lines expressing constitutive low levels of the protein (like SJSA1) have a strong tendency towards apoptosis while others expressing high levels of the helicase, (like HCT116) do not proceed to cell death after the initial cell cycle arrest. Knocking down DHX30 using RNAi in HCT116 affects the propensity of HCT116 cell line towards apoptosis, with a mechanism that involves the translation upregulation of some genes containing a CG-motif in the 3'UTR. However, not all the translational regulated, CG-motif containing apoptotic

target identified in SJSA1 have a similar fate in the context of HCT116 shDHX30 and shPCBP2 clones.

Surprisingly, PCBP2 have a marginal role in the regulation of CG-motif-containing mRNAs (and consequently apoptosis) despite being able to interact with them (see Figure 24B). Finally, the apoptotic response in HCT116 clones is not as dramatic as in SJSA1, possibly because other proteins may interact with the motif in place of PCBP2 and DHX30 when they are silenced. Moreover, the apoptotic response could be milder because of a widespread, yet less strong variation in the expression of many target mRNAs, as suggested by the GSEA analysis. The slight upregulation of many apoptotic targets could make HCT116 shDHX30 clones more susceptible to apoptosis, but not dramatically more sensitive.

4.4 - Expanding the model

4.4.1 - Silencing DHX30 in U2OS

Natthakan Thongon helped me in clone screening and performed some experiments reported in this section.

Given the promising results obtained in HCT116, we decided to expand and generalize our findings by focusing our attention on another cell line. We chose to use U2OS because it is a p53 wild type cell line that responds with persistent cell cycle arrest to the Nutlin-3 treatment, despite sharing the same tissue derivation with SJSA1 (osteosarcoma). According to our previous findings in HCT116, DHX30 seems to have a more determinant role in the choice between life and death and for this reason we decided to concentrate our efforts mostly on this protein.

In accordance with our model, U2OS express a higher level of DHX30 protein compared to SJSA1, like HCT116 (Figure 38). Similarly to what we had previously done, we opted for RNAi approach to knock-down DHX30 expression using a lentiviral vector, but in this case we used only one shRNA (TRCN0000052032), corresponding to the one achieving better silencing levels in HCT116. An example of the successful knockdown of the protein is reported in Figure 38, together with a matching shNT control. Also in this case, one clone was used for most of the experiments while some selected replicates were performed on other clones to account for clone-specific effects.

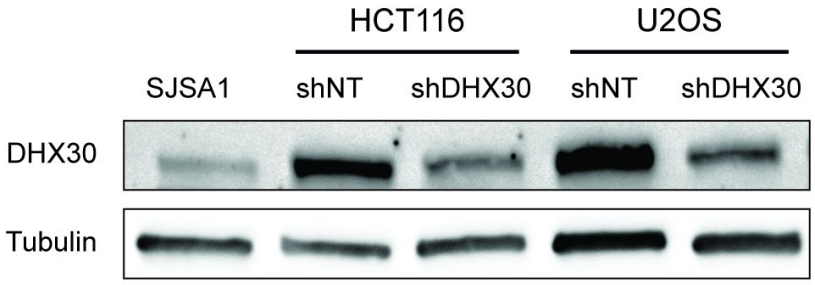


Figure 38: Western blot showing the expression of DHX30 in different cell lines in control (shNT) or silenced conditions (shDHX30). Tubulin was used as a loading control

We analysed the silencing-dependent effects using cell count in digital phase contrast (DPC) using Operetta high content imaging system (PerkinElmer) after the treatment with Nutlin-3 over a 72h time course (Figure 39). DHX30 silencing seems to decrease the number of cells attached to the plate at the final time point, suggesting that Nutlin-3 response changes also in this cell model after the silencing, but with different kinetics compared to HCT116 or SJSA1. We confirmed the observation by performing the same growth curve with the addition of propidium iodide staining (PI) at end-point (72h). As reported in Figure 39B, DHX30-silenced clones show increased PI positivity compared to control clones after p53 activation with Nutlin-3, compatible with a cell death-related process ongoing in these cells, too.

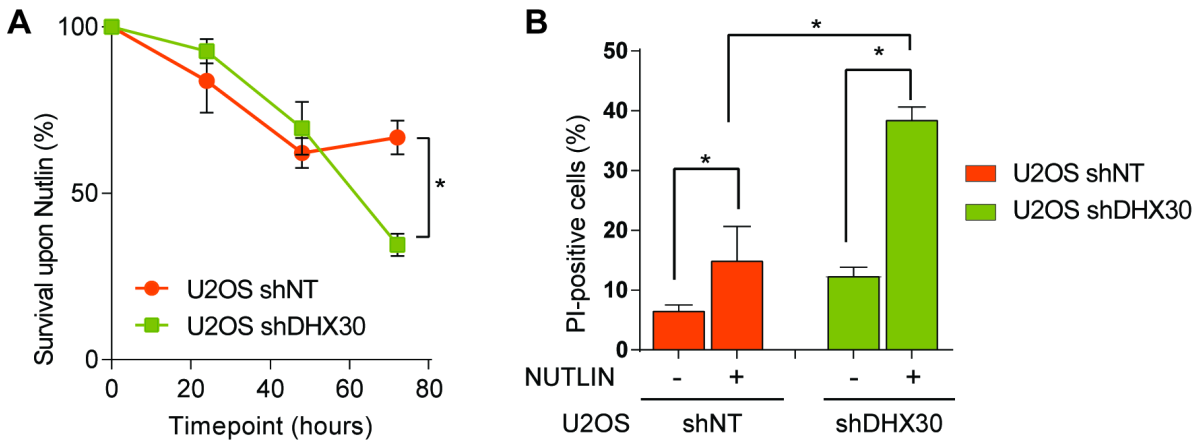


*Figure 39: A= cell count of U2OS control (shNT) or DHX30-silenced (shDHX30) clones over a 72h time course after Nutlin-3 treatment using Operetta in Digital Phase Contrast mode. B = number of PI-positive cells at 72 hours after Nutlin-3 as detected by Operetta. * = p-value < 0.05*

From a morphological point of view, cells with the combination of silencing and treatment show a different morphology and tend to detach from the plate (Figure 40). This phenomenon is visible already at 48 hours but becomes more evident at 72 hours.

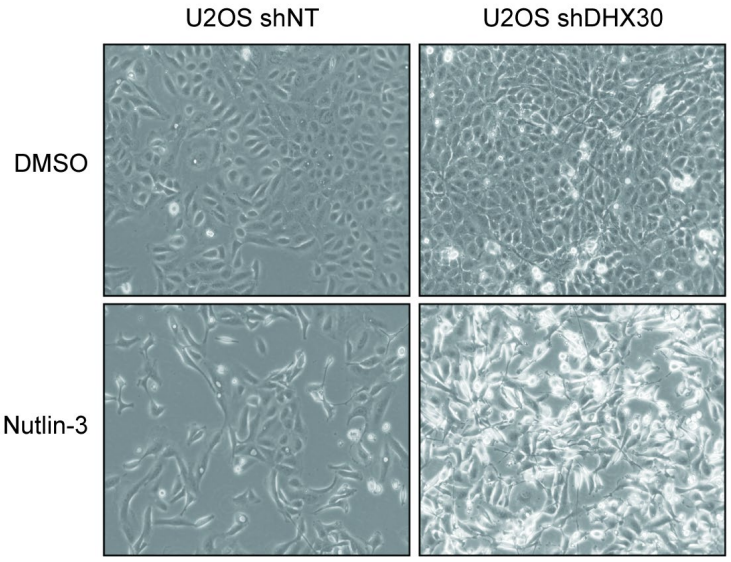


Figure 40: U2OS morphology after 72 hours of DMSO or Nutlin-3 treatment in control (shNT) or DHX30-silenced conditions (shDHX30)

To exclude that the detachment of the cells could be linked to epithelial-to-mesenchymal transition (EMT), we checked the expression of EMT targets using qPCR (Figure 41) in a time-course dependent manner. EMT promoting factors (Vimentin, ZEB and TWIST) are comparable to the control at 24 and 48 hours while decreasing at 72 hours. Conversely, E-cadherin expression (CDH1) increases in Nutlin-3 treated DHX30 silenced clones, indicating that shDHX30 cells change in morphology is not due to ongoing EMT.

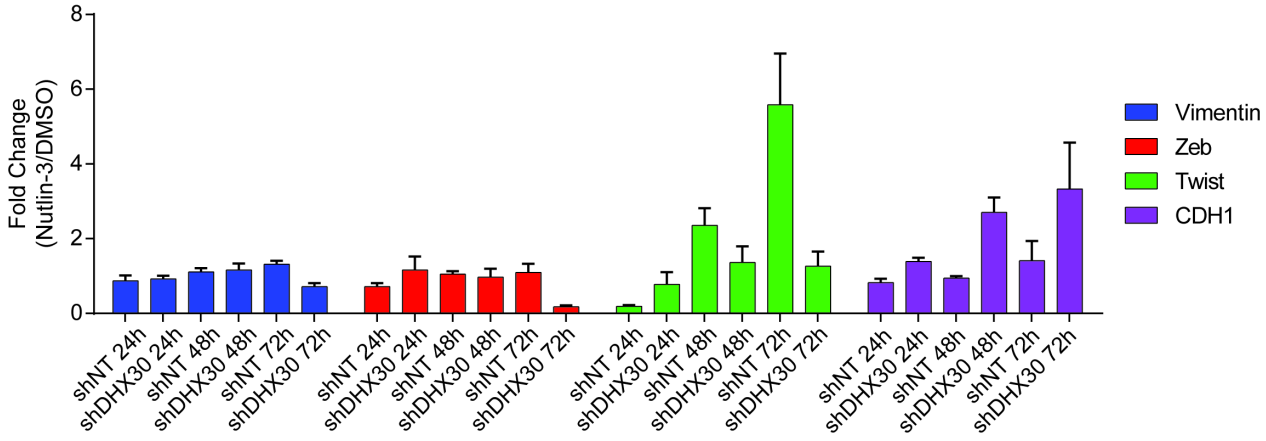


Figure 41: qPCR analysis of the expression of mRNAs involved in Epithelial to Mesenchymal transition (EMT) over a 72-hour time-course. Results are presented as fold change over control treatment (DMSO)

Altogether, these data suggest that DHX30 silencing in U2OS impacts on their response to p53-activating treatment with Nutlin-3. The kinetics of the response seems to be slower compared to SJSA1 or DHX30-silenced HCT116 clones, possibly reflecting the slower growth rate of U2OS compared to the other cell lines. From a rough estimation of the cell morphology or growth, the response seems to be cell death related and we can exclude EMT-like processes being involved in this model.

4.4.2 - Impact on CG-motif mRNA translation and protein production

To confirm that the silencing of DHX30 in U2OS cells enhances translation of CG-containing mRNAs like in HCT116 clones and SJSA1 we used the same luciferase reporter assay described previously. To match for the slower kinetics of the response we performed the assay at 48 hours after the initial treatment, which is delayed compared to HCT116 and SJSA1 but earlier to the overt phenotype we observed at 72 hours. Bar graphs reported in Figure 42 seem to support the model in which low DHX30 levels permit an increased expression of CG-containing mRNAs also in this cell line. We excluded transcription-dependent effects by normalizing relative luciferase units (RLU) to the *luc2* mRNA level detected by qPCR, as it was done in the other assays.

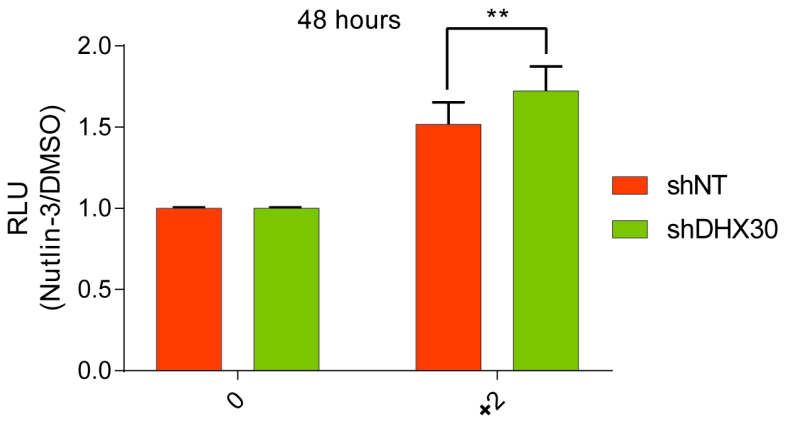


*Figure 42: Luciferase reporter assay in U2OS clones after 48 hours of Nutlin-3. 0 refers to the plasmid containing the wild-type bGlob 3'UTR downstream the luc2 gene while +2 refers to the bGlob 3'UTR added with 2 copies of the consensus CG-motif. ** = p-value < 0.01*

To see if the knockdown had a similar effect on the expression of natively CG-motif containing mRNAs, we performed polysome profiling of control or shDHX30 clones after a 24 hours-long treatment. The choice of a delayed time point compared to the previous experiments in HCT116 and SJSA1 was made to account for the slower response kinetics of the cell line compared to those studied previously. Matching total RNA extracts were obtained from Trizol-based extraction of a fraction (10%) of the cells subsequently employed for polysome separation or from treatment-matched clones subsequently processed with column-based RNA purification kits.

The results of the qPCR are reported in Figure 43 and are presented as 2^{-dCt} over the housekeeping gene (YWHAZ). This unusual choice of presenting the data was made necessary by the fact that in this cell line the silencing has a constitutive effect on the expression of the targets that we considered, which would have been flattened if data had been presented as fold change over DMSO (2^{-ddCt}), like in the other qPCR experiments. The downside of this choice is the intrinsic noise in the data, which did not lead to statistical significance except for Caspase 3.

Overall, it seems that the silencing enhances polysome association of all the targets we considered (Figure 43B), although some degree of upregulation can be seen also in the controls after Nutlin-3 (*BAK1* and *BCL2L1*). Contrary to SJSA1 and HCT116 shDHX30 total mRNA levels are not unchanged after the treatment. We can speculate that this phenomenon can be due to an accumulation effect of the mRNA, given the later time point investigated for this experiment in this cell line. It is hard however to imagine that the effect on the total RNA can be only transcription-dependent since none of the mRNAs considered is a direct p53 target.

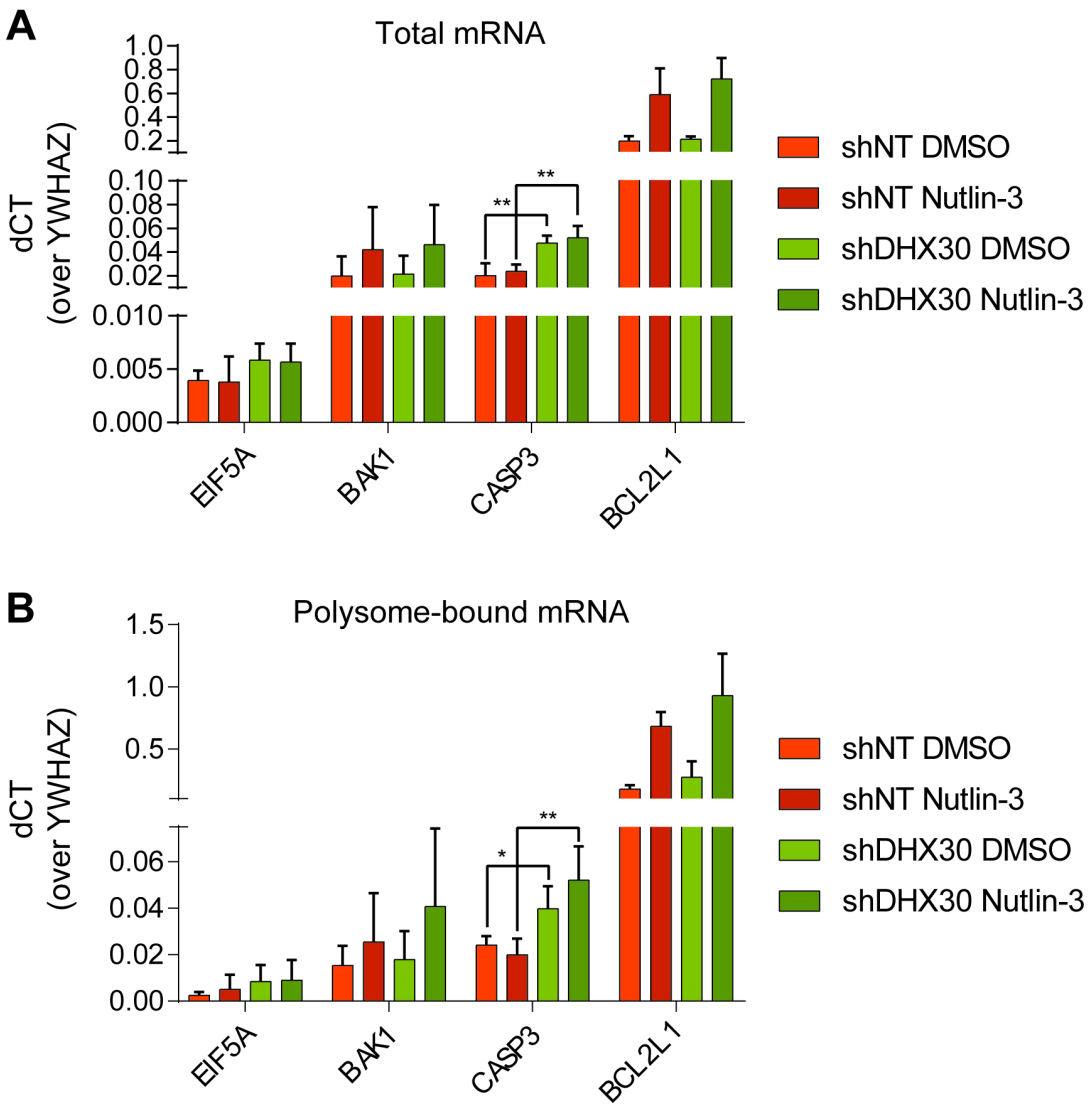


*Figure 43: qPCR results on the total (panel A) or polysome-bound (panel B) mRNAs in U2OS control or shDHX30 clones after 24 hours of treatment. * = p-value < 0.05; ** = p-value < 0.01*

Protein expression levels were also investigated via western blot (Figure 44). Meeting the expectations from qPCR analysis, all the proteins taken in consideration whose mRNA contains a CG-motif (CASP3, BAK and EIF5A), show a marked overexpression compared to shNT controls. Additionally, the boost on protein production seems to be shDHX30 dependent, since increased BAK, CASP3 and EIF5A a is visible in the DMSO condition, matching the observation from the qPCR analysis.

Altogether, the luciferase assay, qPCR and western blot data suggest that also in this cell line CG-motif containing mRNA undergo a different fate in a DHX30-downregulated context. Contrary to the previous experiment in SJSA1 and HCT116 clones, the silencing seems to affect the total level of mRNA and not only the polysome-bound one and not in a strictly Nutlin-3 dependent manner.

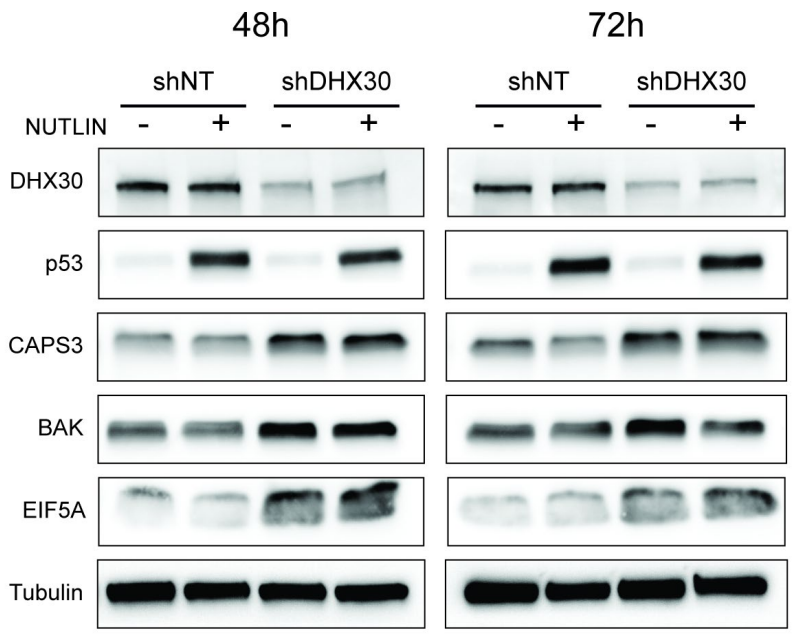


Figure 44: western blot analysis of U2OS clones exposed to DMSO or Nutlin-3 for 48 or 72 hours. DHX30 and p53 were detected to test for the successful silencing and treatment respectively; Tubulin was used as a loading control

4.4.3 - U2OS shDHX30: is it apoptosis?

The caspase activation experiment was performed by Natthakan Thongon.

If our model holds true, to the shDHX30-dependent enhanced CG-motif mRNA translation would correspond an increased propensity (or sensitization) towards cell death. As noticed by others, U2OS parental cells are refractory to apoptosis after Nutlin-3, but they rather undergo a persistent cell-cycle arrest [83]. We checked whether the knockdown of DHX30 had a similar impact on this cell line as it did in HCT116. Unfortunately, a reproducible PARP or Caspase 3 cleavage could not be noted on this cell line via western blot, not even after 72 hours of treatment. We decided to try anyway the annexin V – PI double staining followed by FACS analysis, but also in this case only marginal variations between control and knock-down clones could be noticed (data not shown). Given the morphological changes observed in the shDHX30 clones exposed to Nutlin-3, we suspected that the cells undergoing apoptosis could be lost during the staining procedures preceding FACS analysis. For this reason, we decided to use different assays to measure, either directly or indirectly, the propensity of these clones to undergo apoptosis.

The initial evidence that shDHX30 clones have a different behaviour compared to controls after Nutlin-3 exposure derived from the growth curves presented in Figure 39: knockdown clones show a significantly lower cell count and increased PI positivity at endpoint (72 hours) compared to matching shNT controls.

We used also two luciferase-based assays to test for caspase activation and phosphatidylserine exposition: *Caspase-Glo 3/7 Assay System* and *RealTime-Glo Annexin V Apoptosis Assay* (Promega) respectively. According to these assays, U2OS shDHX30 show a significantly higher level

of caspase activation also in the DMSO condition compared to scrambled controls, which corresponds to a higher level of annexin V positivity starting from 48h after the treatment (Figure 45).

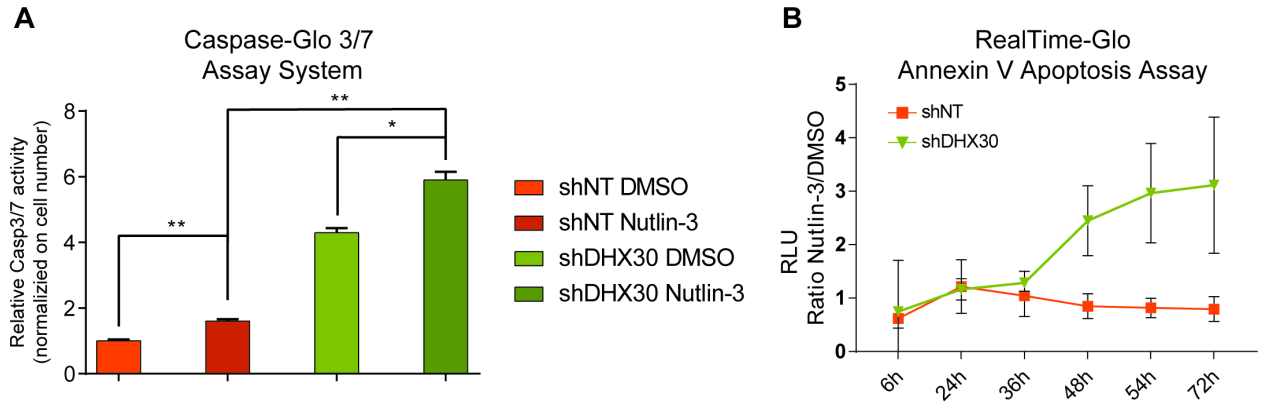


*Figure 45: Results of the Caspase-Glo 3/7 Assay System (Panel A) and RealTime-Glo annexin V Apoptosis Assay (Panel B). * = p-value < 0.05; ** = p-value < 0.01*

Contrary to annexin V staining followed by FACS analysis, RealTime-Glo assay does not require trypsinization and processing of the cells for the staining, thus excluding the chances of losing cells during handling.

Despite the lack of a clear activation of Caspase 3 and/or PARP detected by western blot, the surrogate assays that we used seem to point in the direction of an increased propensity toward cell death in shDHX30 clones after p53 activation. Unfortunately, we cannot firmly state that U2OS are undergoing overt apoptosis or speculate on the magnitude of the effect based on the results of these assays. Moreover, the exact nature of the phenomenon to which U2OS are exposed after silencing is still not completely elucidated, as other types of cell death, besides apoptosis, could be involved in activating caspases (Figure 45A) or in increasing the number of PI-positive cells on the plate (Figure 39B).

Although the results obtained in U2OS do not completely match those obtained in HCT116 and in SJSA1, we can confirm that DHX30 protein level plays a relevant role in the choice between life and death in the contexts we analysed. The mechanism is still not completely clear, but translation control of some apoptotic mRNA is strongly influenced by the presence (or reduction) of this protein.

4.5 - Modulating the responses in SJSA1

4.5.1 - DHX30 overexpression

To consolidate our model, we decided to try the rescue experiment on SJSA1 cell by means of DHX30 overexpression. We initially tried a transient overexpression using the wild type sequence of the protein cloned inside a pCMV6-AC-Myc-His vector (OriGENE) (see section 4.7.1 and 7 for the details of the cloning). Unfortunately, this overexpression was not efficient nor reproducible in SJSA1 cells, despite the successful production of the protein in HEK293-T cells (section 4.7.1). For this

reason, we decided to move to an inducible overexpression system in order to have a time-controlled overexpression of DHX30. For this purpose, we used the pCW57.1 vector (developed by David Root and purchased from Addgene), which is a TetON, all-in-one system containing also the LTR for viral vector production. DHX30 was subcloned starting from the pCMV6-AC-Myc-His plasmid and the destination vector (pCW 57.1 with WT DHX30) was used to produce viral vectors in HEK293-T cells, that were subsequently used to transduce different pools of SJSA1 cells (the map of the plasmid is reported in Figure 46A). Cells were puromycin-selected and subsequently tested for the ability to overexpress DHX30 in presence of 2,5µg/ml doxycycline (Doxy). As a control, ethanol was used (EtOH). Two pools of cells were selected for their ability to induce DHX30 overexpression in a doxycycline-dependent manner and the stability of the induction was tested over 48 hours (Figure 33B).

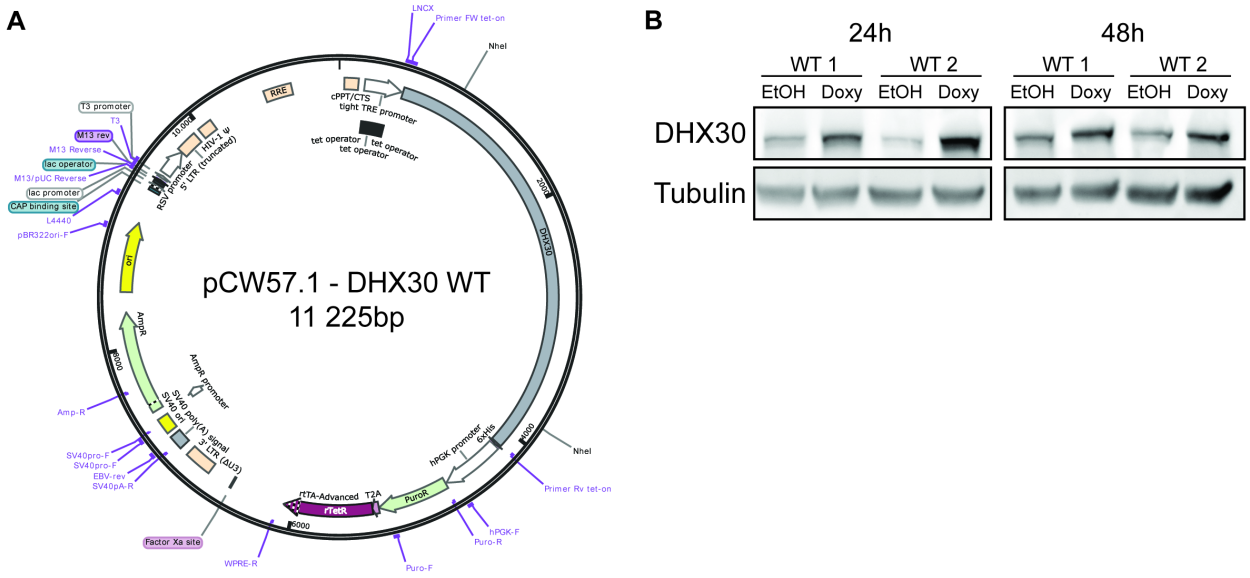


Figure 46: Panel A: map of the pCW57.1 TetON vector containing DHX30 WT coding sequence. Panel B: Western blot showing DHX30 overexpression on two different pools of SJSA1 cells at two time points (24 and 48 hours) after doxycycline addition. EtOH = ethanol control; Doxy = doxycycline. Tubulin was used as a loading control

Having assessed the successful and stable induction of DHX30 overexpression we decided to test SJSA1 overexpressing cells' behaviour to the Nutlin-3 treatment. We tried the luciferase reporter assay at first to assess the impact on the translation of CG-motif containing genes. DHX30 overexpression was induced at least 12 hours before starting a 24-hours Nutlin-3 treatment, in order to activate p53 in a high-DHX30 cellular background. Results presented in Figure 47A show a decreased luciferase activity of the βGlobin 3'UTR added with the consensus CG-motif (+2) in the presence of doxycycline. The result suggests that in this context higher DHX30 levels could lower *luc2* translation with consequent inhibition of luciferase enzyme production, similarly to what was hypothesized from the experiment in parental HCT116 and SJSA1 cells. Also in this case, luciferase units are normalized over the *luc2* mRNA level.

To see if the overexpression has a similar effect on mRNAs with natural-occurring instances of the motif in the 3'UTR, we performed polysome profiling of SJSA1 were DHX30 was induced (Doxy) or

not (EtOH) at least 12 hours before starting a 24 hours long Nutlin-3 (or DMSO) treatment (Pol samples). Cells exposed to matching treatments were used to extract total RNA using Trizol (Tot samples). Figure 47B reports the mRNA expression on the total or polysome-bound mRNAs in a DHX30 normal (EtOH) or overexpressed (Doxy) SJSA1 context. In opposition to the expectations from the luciferase reporter assay, CG-motif containing targets do not show reduced expression in the polysome-bound fractions compared to ethanol-treated controls, but rather similar levels.

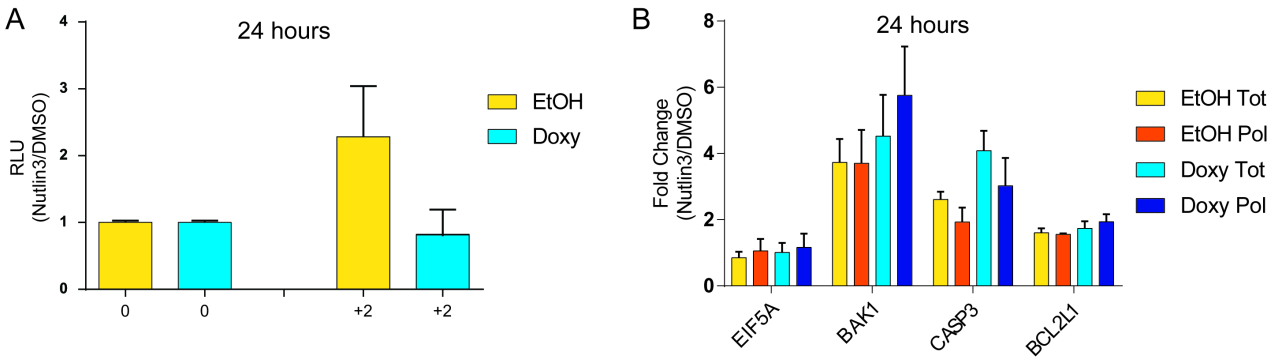


Figure 47: Panel A: results of the luciferase reporter assay on the luc2 gene followed by the bGlobin 3'UTR (0) or the bGlobin 3'UTR added with two copies of the CG-consensus (+2) in SJSA1 treated with ethanol (EtOH) or doxycycline (Doxy). Panel B: expression of mRNAs containing the CG-motif in the total (Tot) or polysome-bound fractions (Pol) of SJSA1 expressing normal (EtOH) or higher levels of DHX30 (Doxy).

We checked also the expression of some CG-motif-containing targets at protein level via western blot (Figure 48). Here the results are more in line with the expected DHX30-dependent downregulation of CG-motif targets. By looking at the 48-hours time point BAK, CASP3 and EIF5A expression are higher in ethanol and Nutlin-3 treated SJSA1. For unknown reasons, doxycycline and Nutlin-3 seem to play a synergistic role in DHX30 overexpression, a phenomenon that is particularly evident at 48 hours after Nutlin-3. From the analysis of the Caspase3 and PARP cleavage, it seems that SJSA1 overexpressing cells are less susceptible to cell death, given the apparently lower levels of both cCASP3 and cPARP at 48 hours.

We then tried to quantify if DHX30 overexpression impacted on SJSA1 apoptosis either quantitatively or in the timing of the response by annexin-V/PI staining and FACS analysis at different time points after the treatment. Doxycycline was added to the culture media at least 12 hours before starting DMSO or Nutlin-3, similarly to what was done in the other experiments. Unfortunately, no differences were noticed between doxycycline or ethanol-treated cells as regards the timing or the magnitude of the apoptotic response (data not presented).

Using this cellular model, we could only partially confirm the importance of DHX30 in lowering apoptotic rates by acting on the translation of mRNAs containing a 3'UTR with the CG-motif.

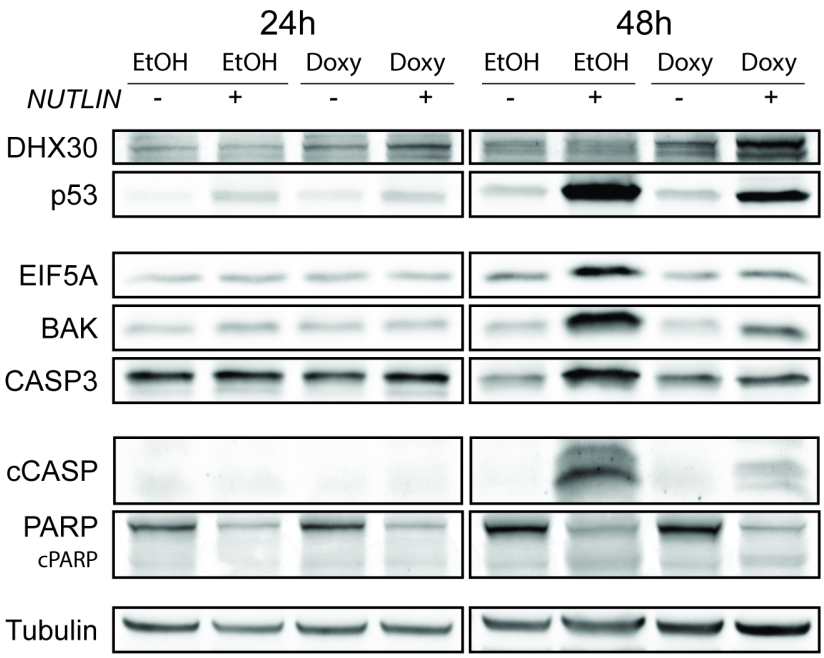


Figure 48: Western blot of the expression of CG-motif-containing targets and apoptosis markers in SJSA1 overexpressing DHX30 cells. DHX30 and p53 were included as control of the successful overexpression and Nutlin-3 treatment, respectively. Tubulin was used as a loading control.

4.5.2 - MYH9: another player in translation control?

According to the mass spectrometry data of the interactors of the CG-motif in HCT116 and SJSA1 cells (Figure 22B), the Myosin Heavy Chain 9 (MYH9, also known as non-muscle myosin IIa) was identified as a specific interactor of SJSA1 cells. This protein has been associated with different types of diseases, including cancer[101]. Its role in oncogenesis is debated, since it may act as an oncogene[102,103] or as a tumour suppressor gene[104,105] in different cancer types. Notably, it was demonstrated that MYH9 is able to stabilize the p53 protein and favour its nuclear retention, thus helping p53's tumour suppressor gene functions in the squamous cell carcinoma[105]. Despite not having recognized RNA binding domains, at least one report[96] have demonstrated how this protein can interact with a long non-coding RNAs (named PTCSC2), favouring tumour suppression when expressed in the wild type form in the context of thyroid cancer.

We decided to explore if this protein could be important also in the context of translation of CG-motif containing mRNAs or if the results obtained by mass spectrometry were due to a spurious interaction given by the higher abundance of the protein in SJSA1 compared to HCT116 (see Figure 49A for the relative protein expression levels in the two cell lines). We performed the pulldown using the consensus CG-motif followed by immunoblot detection of MYH9 to confirm mass spectrometry data (Figure 49). Results seem to confirm the ability of MYH9 to bind the motif, but no striking differences were noticed in the binding of WT vs mutated motifs. Surprisingly, MYH9 seem to interact more strongly in SJSA1 in presence of Nutlin-3.

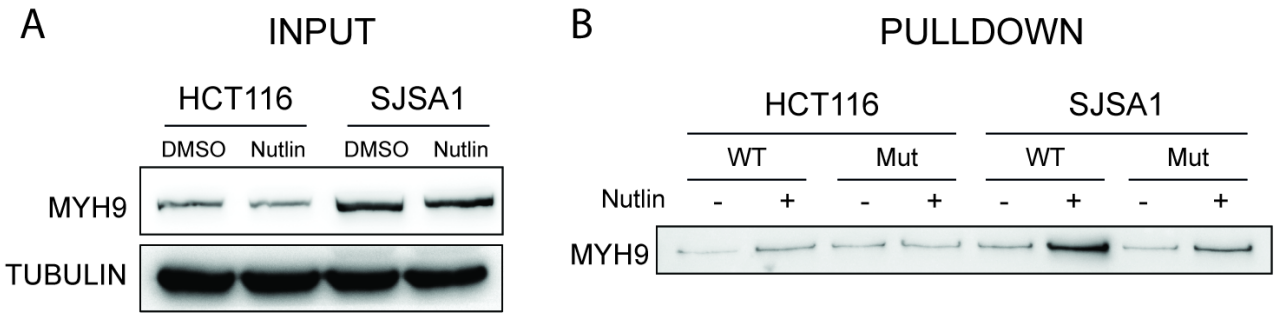


Figure 49: Panel A: expression level of MYH9 in HCT116 and SJSA1 in the input used for the RNA pulldown. Tubulin was used as a loading control. Panel B: binding of MYH9 to the WT or mutated consensus CG-motif using protein lysates from HCT116 and SJSA1 cells

MYH9 expression in SJSA1 cells was perturbed using a lentiviral vector-based shRNA approach similar to what was previously done for DHX30 and PCBP2. We had two different shRNA available, whose identifiers and sequences are reported in section 11.4.

Unfortunately, during plasmid testing, only TRCN0000029465 plasmid showed the correct migration in an agarose gel and thus was the only shRNA that we had available for MYH9 knock-down.

After producing viral vectors in HEK293-T cells, we infected different pools of SJSA1 that were subsequently puromycin-selected before proceeding with single clone dilution of the pools. As a control, we transduced SJSA1 with the same scrambled RNA (shNT) that was used for U2OS and HCT116.

After careful testing of the clones, we picked up those presenting the highest silencing degree for further experiments. Surprisingly, MYH9 silencing did not cause any apparent morphological change in SJSA1 (Figure 50A) and Nutlin-3 treatment did not alter MYH9 expression in control or shMYH9 cells (Figure 50B).

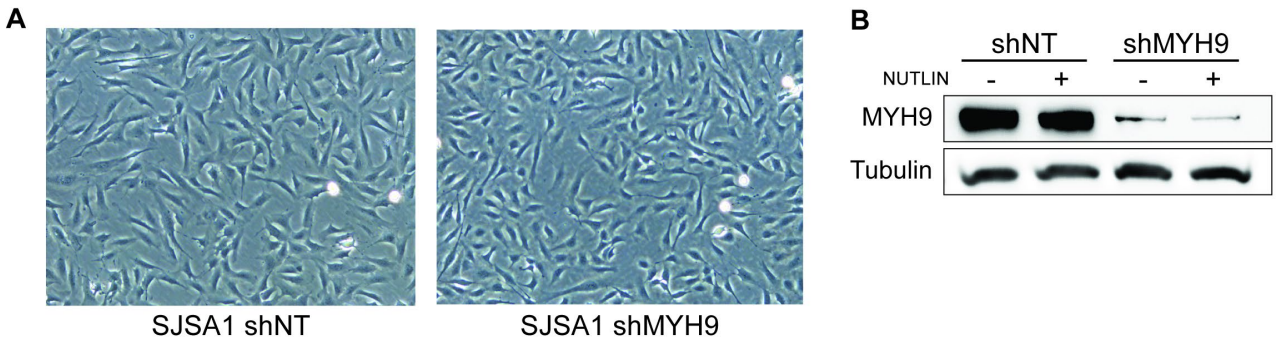


Figure 50: Panel A: Morphology of control (shNT) or silenced (shMYH9) SJSA1 cells. Panel B: MYH9 protein expression levels in control and shMYH9 clones with or without Nutlin-3 treatment. Tubulin was used as a loading control

We aimed next at identifying the effect of the silencing on the translation of CG-motif transcripts using the luciferase reporter assay (Figure 51A) after 24 hours of Nutlin-3 exposure. As expected, control cells (shNT) show enhanced luciferase activity in the presence of the βGlobin 3'UTR added twice with the CG-consensus (+2). Surprisingly this effect is boosted in a shMYH9 context, although it does not reach statistical significance. According to this result, translation (and consequently

protein production) of CG-motif mRNA should be increased in an MYH9-depleted background, and this holds true for BAK expression in shMYH9 at 12 hours, but not for Caspase-3 (Figure 51B). Interestingly, MYH9 knock-down seems to anticipate the apoptotic onset, as CASP3 cleavage was already visible at 24 hours in shMYH9 clones treated with Nutlin-3.

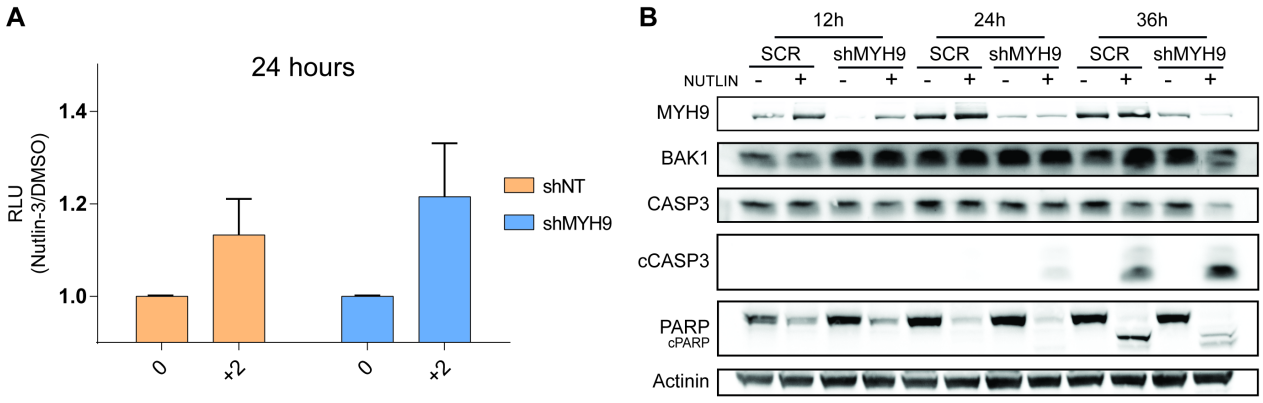


Figure 51: Panel A: luciferase reporter assay of shMYH9 and shNT SJSA1 cells. 0 refers to the WT bGlobin 3'UTR; +2 refers to the bGlobin 3'UTR added with two copies of the CG consensus. Panel B: expression of CG-containing genes and apoptotic markers in at different time points in SJSA1 (control and shMYH9). Actinin was used as a loading control

To confirm the different sensitivity in apoptosis commitment of silenced clones, we repeated the annexin V and PI double staining followed by FACS analysis (Figure 52). It is possible to notice how shMYH9 clone shows a higher fraction of annexin V positive cells already at 24 hours after the treatment and significantly more positive cells at 48 hours, despite the control cell line undergoes massive apoptosis by itself.

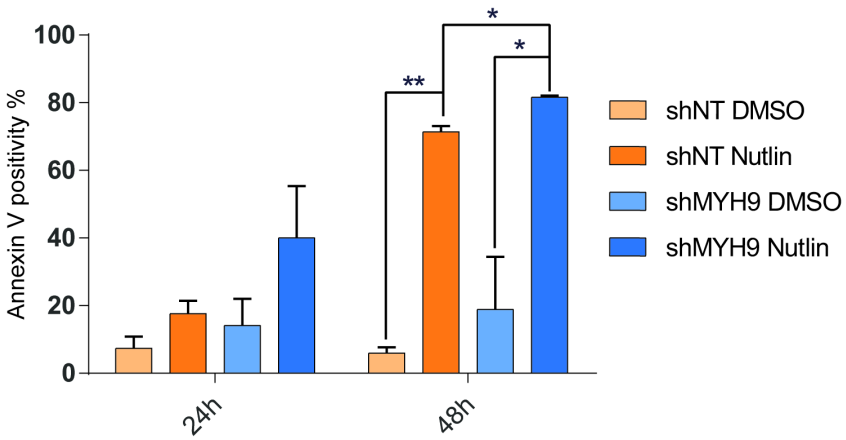


*Figure 52: Annexin V staining of shMYH9 and control (shNT) SJSA1 cell after 24 and 48 hours of exposure to Nutlin-3. * = p-value < 0.05; ** = p-value < 0.01*

Altogether, it seems that also MYH9 has a role in the regulation of the apoptotic response in SJSA1 cells exposed to p53 activation by non-genotoxic means. From these data, it is hard to rule out a clear MYH9-dependent impact on translation, but luciferase assay and BAK protein expression at 12 hours seem to suggest a role in this type of regulation. Notably, BAK expression in shMYH9 clones seem to be constitutively higher, somehow reflecting what was previously noted in U2OS shDHX30 clones. Despite these uncertainties, MYH9 knock-down clearly impacts SJSA1 propensity to undergo apoptosis after Nutlin-3.

4.6 - Analysis of BAK and BAX activation

BAK and BAX are multi-BH-domain proteins that are well known for their capacity to induce mitochondrial apoptosis. They are usually present in the cell in a non-active state thanks to the action of anti-apoptotic members of the BCL2 family, such as MCL1 (acting predominantly on BAK[106]), BCL2 (acting on BAX[107]) and BCL-X_L (inhibiting both BAK and BAX activation[108,]). Apoptotic stimuli that activate other pro-apoptotic, BH3-only proteins, like PUMA or NOXA, disrupt the inhibition provided by anti-apoptotic BCL2 proteins, favouring BAX translocation to the mitochondria, conformational changes and oligomerization of both BAK and BAX and subsequent mitochondrial outer membrane permeabilization (MOMP) and activation of the apoptotic cascade[109]. Additionally, PUMA can exert a direct activation of BAK and BAX besides counteracting the action of anti-apoptotic BCL2 family proteins [110,111].

Since BAX is a known transcriptional target of p53 and BAK is a CG-motif-containing gene, we wanted to see if and how these two proteins are involved in the induction of apoptosis in HCT116 and SJSA1 after Nutlin-3 treatment. We pursued this idea by exploiting BAK and BAX conformation-sensitive antibodies, namely anti-BAK AB1[112] and anti-BAX 6A7[113], followed by flow cytometry analysis of cell after Nutlin-3 treatment, as previously done by other groups[114,115].

Since BAK and BAX activation precedes apoptotic onset, we decided to track their activation at 16 hours post-Nutlin-3 treatment, but no differences could be spotted between HCT116 and SJSA1 (data not shown). We focused then at 24 hours after treatment, when SJSA1 start showing a low level of annexin V positivity. In Figure 53 is reported the result of BAK (panel A) and BAX (panel B) activation of one out of three replicates yielding similar fluorescence distribution. Surprisingly both apoptotic resistant HCT116 and apoptotic prone SJSA1 show activation of BAK/BAX because of Nutlin-3 treatment, with an unexpectedly more marked shift in fluorescence for HCT116. To see if the phenomenon is specific for this time point only, we checked protein activation also at a later time point (36 hours – Figure 54).

To our surprise, also in this case HCT116 show a markedly higher fluorescence shift in Nutlin-3 treated conditions compared to SJSA1. This is particularly striking since SJSA1 have an openly apoptotic phenotype at this time point, contrary to HCT116 which have ≈10% annexin V positive cells at 48 hours post-treatment.

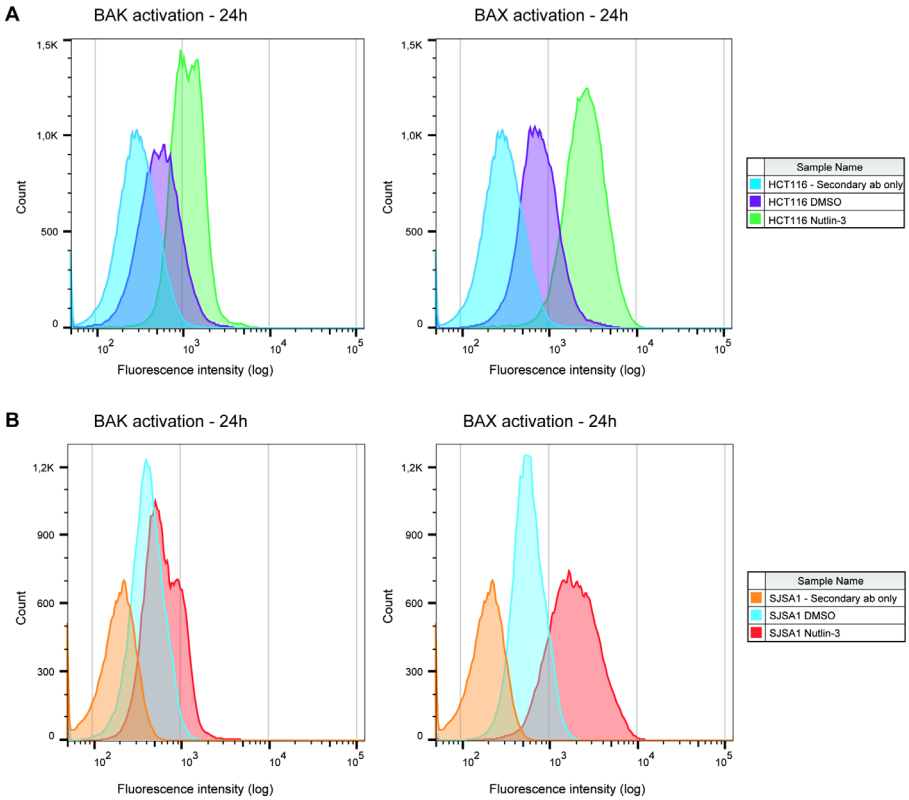


Figure 53: BAK and BAX activation in HCT116 and SJSA1 after DMSO or Nutlin-3 treatment for 24 hours. On the y-axis is reporter the number of events and on the x-axis the fluorescence intensity detected by FACS analysis

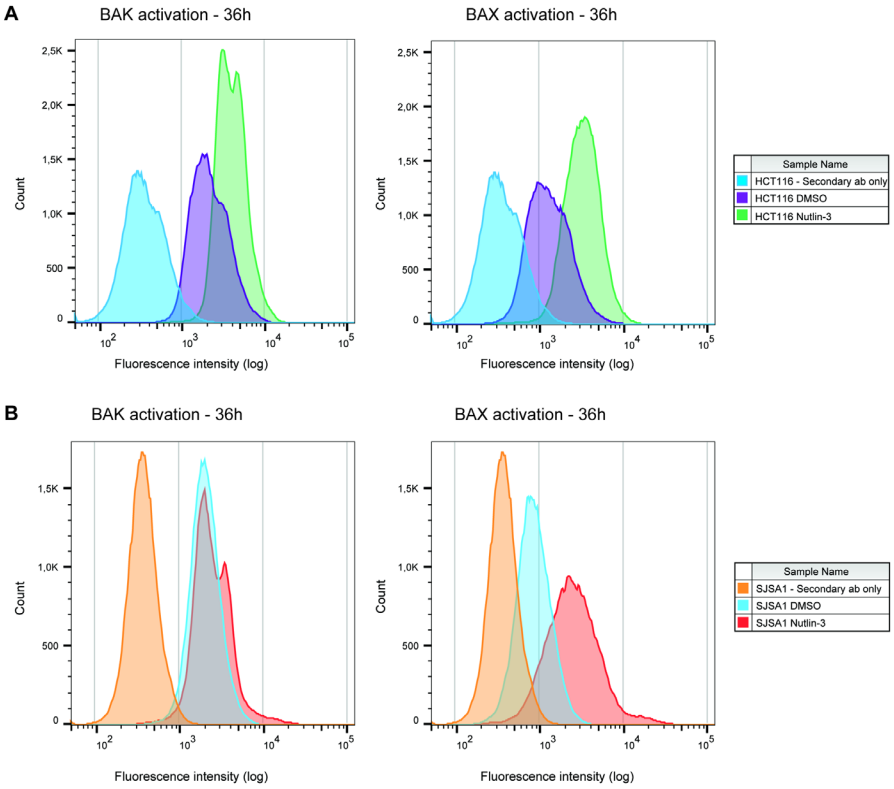


Figure 54: BAK and BAX activation in HCT116 and SJSA1 after DMSO or Nutlin-3 treatment for 36 hours. On the y-axis is reporter the number of events and on the x-axis the fluorescence intensity detected by FACS analysis

To investigate if a temporal difference in the activation (or inactivation) of the proteins could account for the differential phenotype, we decided to follow BAX activation over four time points (12 to 36 hours post-treatment) (Figure 55). We chose to follow only this protein since its activation seems to be more evident with respect to DMSO-treated controls. DMSO-treated HCT116 and SJSA1 show a comparable level of active BAX at all the time points considered. Analysis of the cells treated with the small molecule match the previous results: Nutlin-3 induces BAX activation at early time points (12-16 hours) compared to DMSO, but curves skew towards higher fluorescence values later (24-36 hours). Surprisingly, HCT116 show a more marked shift than SJSA1, which does not diminish over the time frame that we considered.

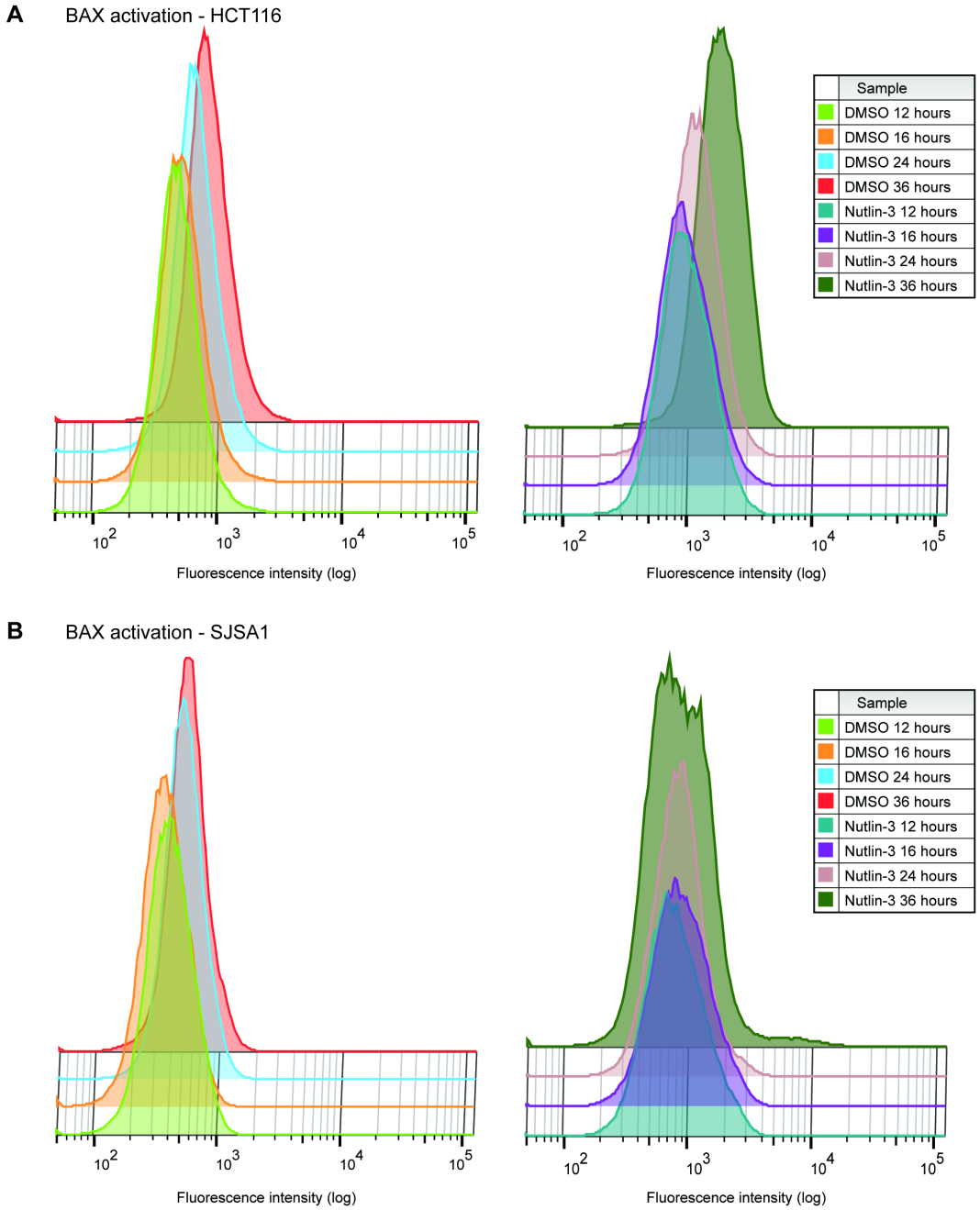


Figure 55: Bak activation at 12, 16, 24 and 36 hours after DMSO or Nutlin-3 treatment in HCT116 (Panel A) and SJSA1 (Panel B) detected by FACS analysis

From the analysis of BAK and BAX activation it is hard to predict if SJSA1 apoptosis depends mainly on the transcriptional (BAX) or translational (BAK) contribution of p53-dependent responses. For this reason, we opted for a CRISPR/Cas9 knock-out approach to remove either one or the other protein or both. We used a lentiviral vector strategy based on the lentiCRISPR v2 backbone (AddGENE, gently provided by the Villunger's lab) and two different antibiotic resistance (puromycin and blasticidin) to help the selection of double knock-out clones.

HCT116 and SJSA1 single or double KO were tested for their responses to Nutlin-3 over 72 hours using Operetta high content imaging system in digital phase contrast mode to study their growth. At endpoint, cells were stained using Hoescht and Propidium Iodide to have a more precise cell count and estimate the number of late apoptotic/necrotic cells. In this experiment two different concentrations of Nutlin-3 were used, the standard 10 μM used in all the previous experiments and 5 μM, beside a vehicle DMSO control.

According to the growth curves (Figure 56), knockout clones derived from HCT116 cells have a similar growth trend to the parental cell line, showing a very slight growth 24 hours post Nutlin-3 (independently from the concentration), which remains almost stable in the subsequent time points for the 10 μM treatment while slightly increasing in the 5 μM wells. None of the knockout clones, either single or double depleted, show differences compared to the parental: HCT116 are in fact apoptosis-refractory when exposed to Nutlin-3.

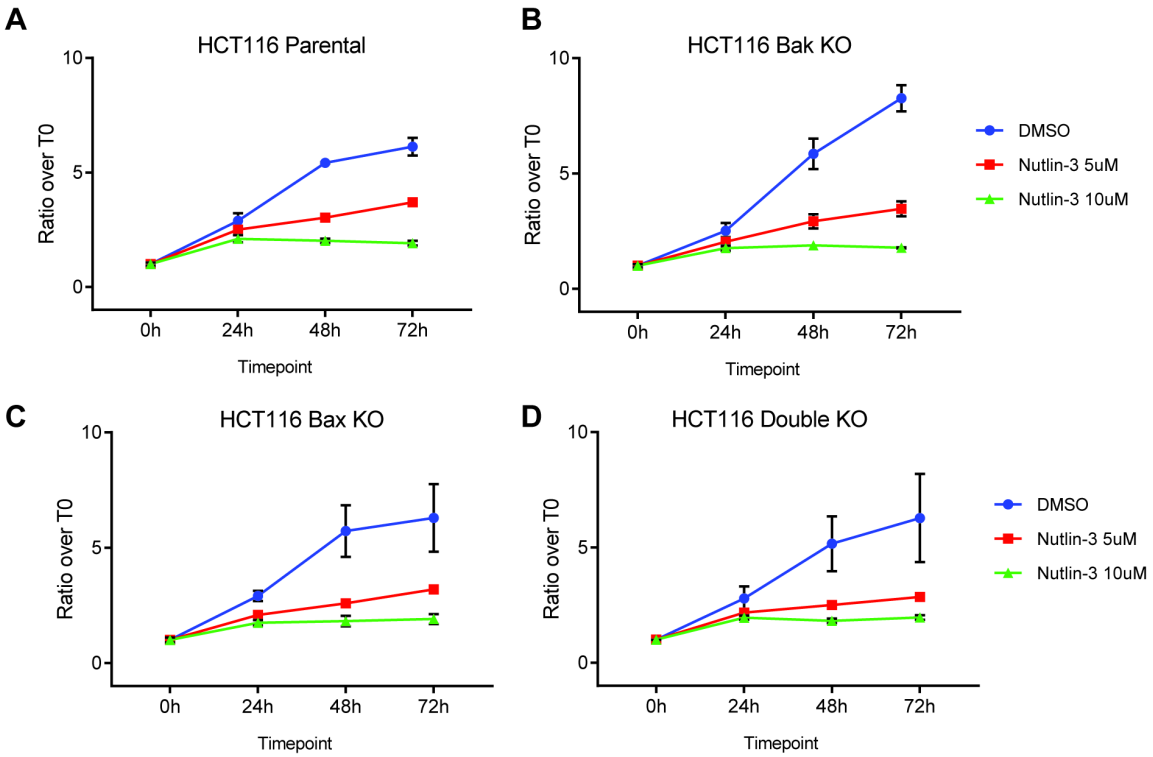


Figure 56: Growth curve of parental (A), Bak (B), Bax (C) and double (D) knock-out of HCT116 cells estimated using Operetta in Digital phase contrast mode. Results are presented as a ratio over the number of cells present in the well at T=0.

Figure 57 shows the growth curves obtained in parental SJSA1 and derivative KO clones. From the cell count estimation using digital phase contrast, depletion of neither BAK, BAX nor both prevents cell number from decreasing, starting from 48 hours after the treatment. Surprisingly, not even a lower dose of Nutlin-3 seems to prevent this phenomenon from happening.

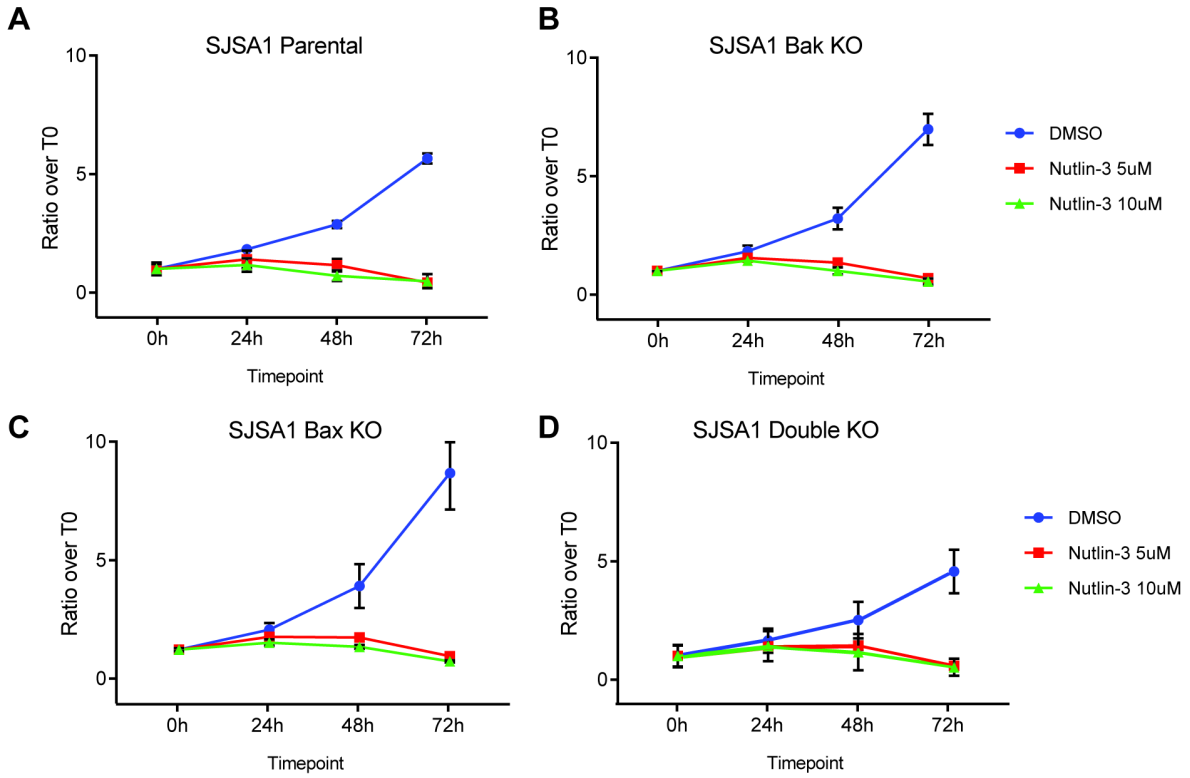


Figure 57: Growth curve of parental (A), Bak (B), Bax (C) and double (D) knock-out of SJSA1 cells estimated using Operetta in Digital phase contrast mode. Results are presented as a ratio over the number of cells present in the well at T=0.

By counting cell number using nuclear count after Hoescht staining, results are in line with what expected from DPC estimation of cell number: HCT116 and SJSA1 show a markedly reduced number of cells in Nutlin-3 treated conditions compared to controls in every genetic background (Figure 58A and 59A). Interestingly, the number of PI-positive events is relatively high in HCT116, especially in the parental and BAK KO conditions, regardless of the treatment and concentration (Figure 45B). BAX and double knock-out clones seem to be more protected from cell death as demonstrated by the low number of PI-positive events in the DMSO controls. In any case, the number of PI-positive events never exceeded 10% of the total nuclear count, in line with the very low apoptotic response of this cell line to non-genotoxic p53 activation.

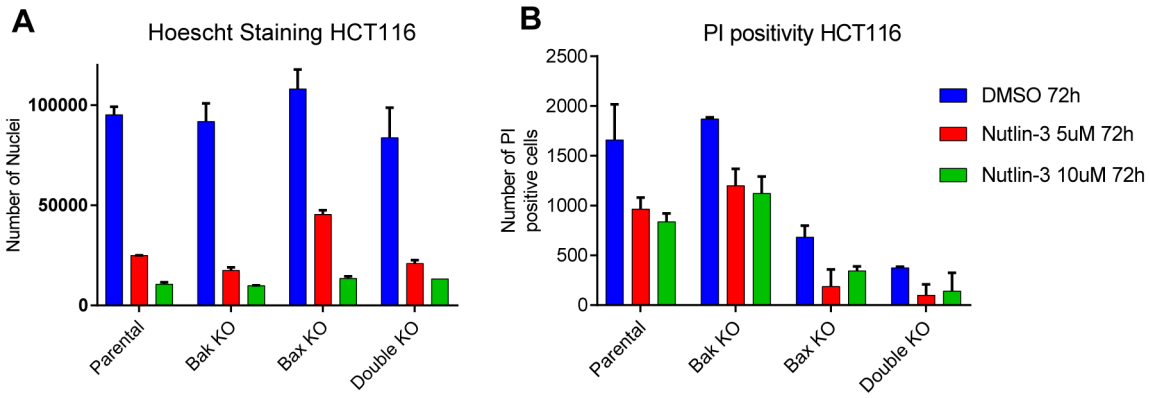


Figure 58: Panel A: Nuclear count after Hoescht staining using Operetta in HCT116 cells after 72 hours of treatment. Panel B: number of PI-positive cells after the treatment at the endpoint

As expected, SJSA1 show a more striking propidium iodide positivity when exposed to Nutlin-3, in particular at the higher dose (Figure 59B). Using PI positivity as a proxy of cell's sensitivity to the treatment, it seems that the knock out of either BAK, BAX or both has a protective effect on SJSA1 death compared to parental controls. Moreover, double knockout seems to be particularly effective at lower doses of the small molecule.

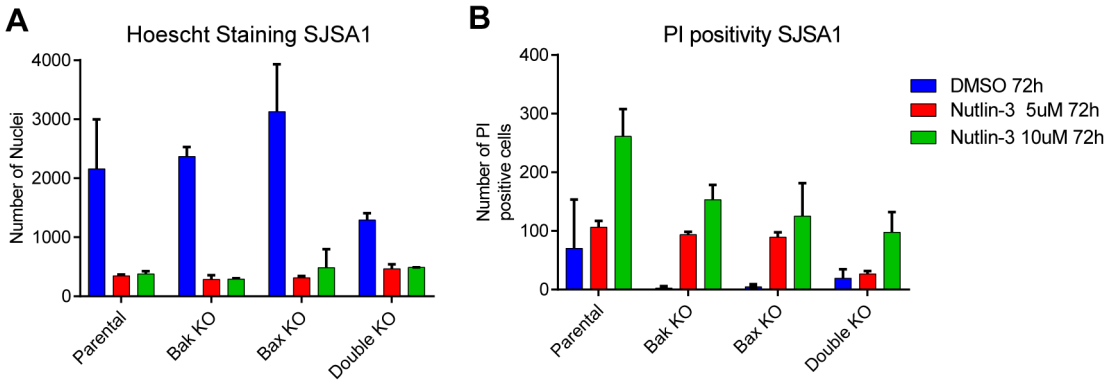


Figure 59: Panel A: Nuclear count after Hoescht staining using Operetta in SJSA1 cells after 72 hours of treatment. Panel B: number of PI-positive cells after the treatment at the endpoint

To confirm the protective effect of the knockout in SJSA1 we opted again for annexin V and PI double staining of SJSA1 after 48 hours of exposure to Nutlin-3 (10μM). As depicted in Figure 60, only the double KO prevents SJSA1 from undergoing apoptosis, as detected by annexin V staining and FACS analysis.

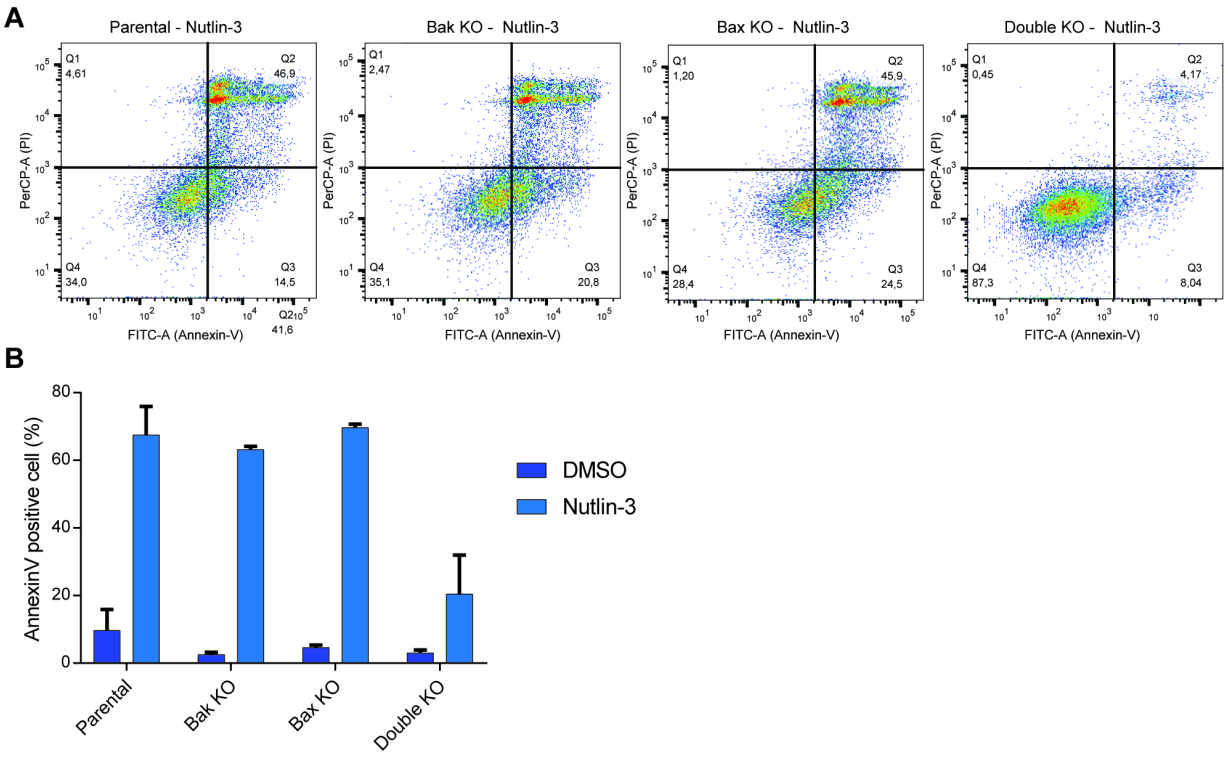


Figure 60: Panel A: example of the FACS analysis of SJSA1 (parental and KO derivative clones) after 48 hours of Nutlin-3. Panel B: summary of the annexin V positive events in SJSA1 cells (sum of Q3 and Q2 of panel A).

Despite the different efforts, it is hard to rule out a differential role of BAK and BAX in inducing apoptosis in SJSA1 cells and thus estimate how the transcription and/or translation control happening in cell lines after Nutlin-3 treatment may determine the phenotypic outcome. In fact, only double knockout has a strong protective effect on SJSA1 cells. PI positivity estimated on the Operetta provides us with some hints on the higher role of BAX in HCT116, although the result must be considered in the light of the apoptosis-refractory phenotype that this cell line has towards p53 activation by Nutlin-3.

4.7 - In vitro studies

4.7.1 - Quantification of DHX30 binding ability to the RNA

Vito D'Agostino helped in planning the cloning strategy, optimizing buffers for DHX30 purification and setting optimal conditions for ALPHA assays.

Having demonstrated the role of DHX30 in favouring apoptotic-like phenotypes in knock-down contexts (HCT116 and U2OS) we wanted to exploit these findings to investigate a possible therapeutic inhibition of this protein in combination with p53 activating drugs.

To pursue this idea by means of a biochemical approach, we decided to purify the protein in order to perform in vitro assay and better understand its RNA binding performance and, possibly, develop a feasible assay to test small molecule inhibitors able to specifically perturb DHX30-RNA interaction.

DHX30 coding sequence was amplified via PCR from a total mRNA extract of HCT116 cells using primers containing the restriction sites for MluI and SgfI, which were used to digest the pCMV6-AC-Myc-His cloning vector (OriGENE, gently provided by Vito D'Agostino and the Provenzani's lab). This cloning strategy allowed us to overexpress a C-terminus Myc-His tagged version of the protein under the control of a strong constitutive promoter (CMV promoter) (see Figure 61A for a map of the plasmid after cloning). We tested the effective DHX30 overexpression by transiently transfecting different plasmids in HEK293-T cells. To verify the successful overexpression, we transfected different quantities of the plasmids and we incubated the blotting membranes with antibodies recognizing the endogenous DHX30 protein and the Myc-Tag derived from the overexpression plasmid (Figure 61B).

Having demonstrated the functionality of the expression vector (linear response to increasing doses of transfected plasmid and absence of Myc-Tag bands in Hek293-T control lysates), we scaled up the system to attempt protein purification exploiting nickel-chelate sepharose beads (GE Healthcare). After different attempts, we obtained a successful composition of the buffers to maximize purified protein yield and we proceeded with dialysis and size exclusion of the imidazole-eluted protein using Amicon Ultra 15 100K filters (Millipore).

Purified protein was used in Amplified Luminescent Proximity Homogeneous Assay (ALPHA) technology using the c-Myc Detection Kit (PerkinElmer). Briefly, donor beads are coated with streptavidin that can interact with the biotinylated RNA probes previously used for pulldown assays (see section 4.2.2) while acceptor beads have anti-Myc antibodies that can recognize the Myc-tag at the C-terminus of purified DHX30 protein. When the RNA probe and the protein are interacting, donor beads are in close proximity to the acceptors. When donor beads are excited by a 680nm laser, a proportional number of oxygen singlets is released, which can reach acceptor beads only if they are < 200nm away. If so, acceptor beads will emit light that is detected as ALPHA count units.

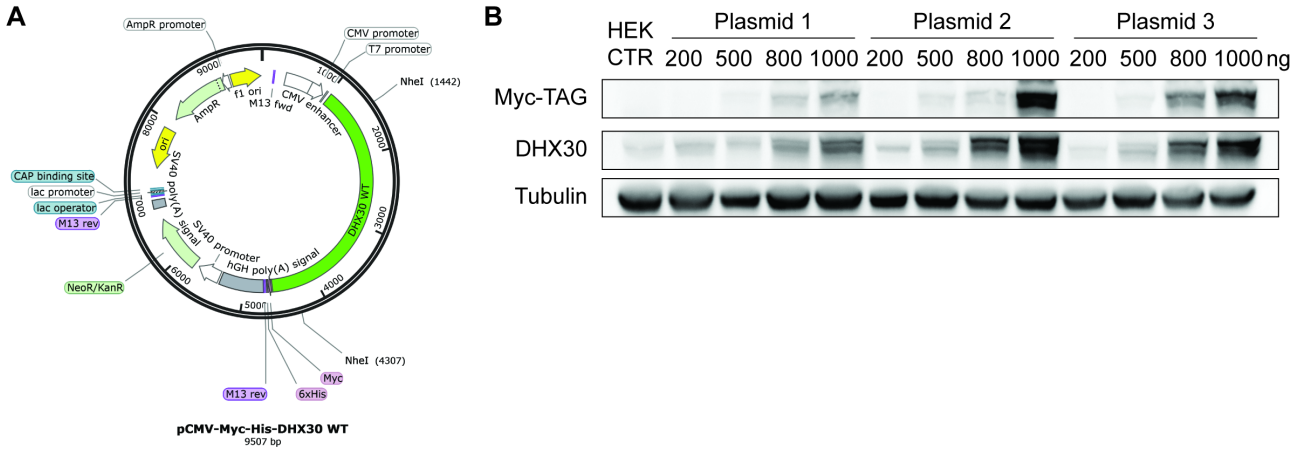


Figure 61: Panel A: DHX30 overexpression vector. Panel B: Western blot testing the successful overexpression of DHX30. Three different plasmids at four different concentrations (200-1000ng) each were used. Both the native protein (DHX30) and the Myc-tagged version (Myc-TAG) were tested. Untransfected HEK293-T cells lysate was used to discriminate overexpressed from native protein (Hek CTR). Tubulin was used as a loading control.

To quantify the strength of DHX30 binding to the RNA we used ALPHA assay to measure the K_d of the interaction. Figure 62 reports an example of DHX30 interaction with different concentrations of biotinylated RNA probe of the wild type CG-motif, the mutant variant (used also for pulldown experiments), an AU-rich element and a negative control probe (the sequences of the probes are reported in section 7). For this assay, protein concentration was kept fixed (130nM; this molarity was identified in previous titrations as corresponding to the hooking point, that indicates the best ratio among beads and test ligands in the assay) and below the top binding capacity of the acceptor beads. We processed ALPHA count measurements using GraphPad Prism using a non-linear regression model (binding-saturation: one-site binding specific equation) to calculate the binding affinity of DHX30 to the different RNA baits.

DHX30 showed the strongest affinity towards the wild type CG-motif (K_d 1.2 nM; R^2= 0.9971) followed by the mutated motif (K_d 5.7 nM; R^2= 0.9499), reflecting the results observed in the immunoblot after pulldown using the same probes (see Figure 23B). A negative probe composed mainly of adenosines had a significantly weaker binding (K_d 9.7 nM; R^2= 0.7826) while the AU-rich probe (containing the consensus binding motif for HuR) was not bound by DHX30 in those experimental conditions.

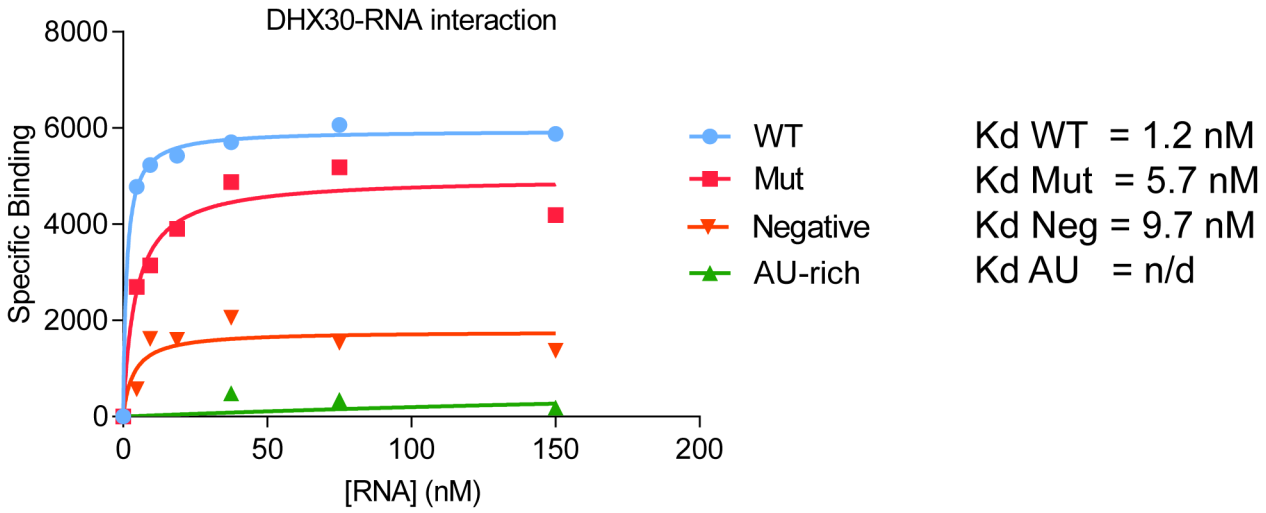


Figure 62: DHX30 affinity (Kd) towards different RNA probes containing wild type CG-motif (WT), a mutated CG-motif (Mut), a negative probe (Neg) and HuR binding sequence (AU-rich).

4.7.2 - Effect of DHX30 mutations

Vito D'Agostino helped in performing some experiments and in data interpretation.

Having proved the effective binding of DHX30 to the RNA we wanted to test if identified mutations of the protein could impact on its binding capacity. The group of prof. Davor Lessel gently provided us with the plasmids containing two mutant versions of the DHX30, that they discovered as de-novo

mutations associated with a severe form of mental retardation in humans [94]. Figure 63 reports the position of the mutations in the coding sequence and the domains of DHX30 where they fall. Mutant Arg493His (from now on referred to as R4) falls within the RNA binding domain of the protein while mutant His562Arg (from now on H5) affects the Walker B motif responsible for ATP hydrolysis. As demonstrated by Lessel et al, R4 mutants show impairment in the RNA binding capacity, while H5 mutant has impaired ATPase activity.

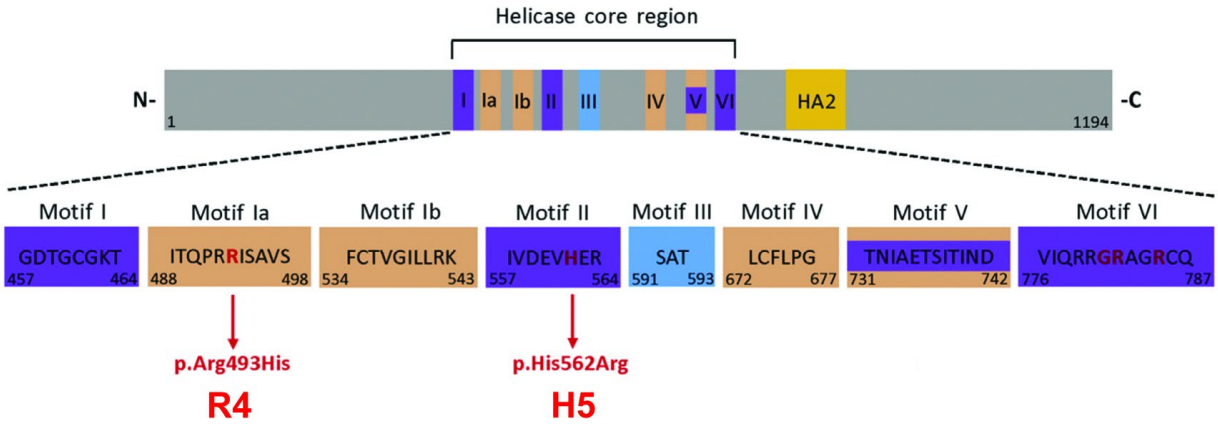


Figure 63: Position of the mutations in the DHX30 domains (adapted from Lessel et al, AJHG, 2017 [94])

After cloning of the mutants in the pCMV6-AC-Myc-His vector and proceeding with protein purification, we measured the binding of the corresponding proteins to the wild-type CG-consensus RNA probe (Figure 64). Also in this case, we kept a fixed protein concentration and titrated the RNA probe concentration. Curve fitting was performed using GraphPad Prism using non-linear fitting (specific binding with Hill slope).

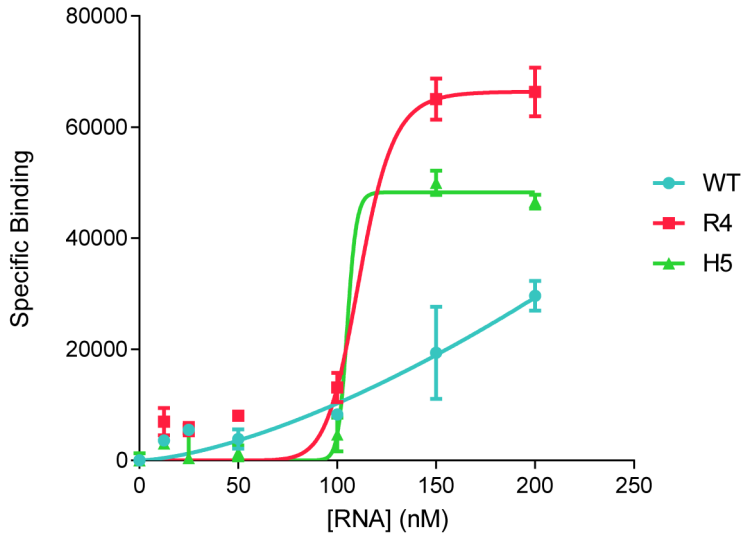


Figure 64: Curve fitting (specific binding with Hill slope) of WT and mutant DHX30 proteins binding to different concentrations of the CG-motif RNA probe

The Hill slope (h) indicates the tendency of a receptor to bind the ligand in a cooperative way. When h = 1 it means that binding is not cooperative (a single monomeric “receptor” binds the ligand independently) while when h > 1 a cooperative binding is implicated (binding of one ligand facilitates the binding of subsequent ligands).

According to the fitting, WT protein showed an h coefficient of 1.417 (R^2 = 0.9197) suggesting very little cooperativity in the binding to the RNA. Conversely, both mutants showed a strong cooperative behaviour: R4 had h=12.58 (R^2=0.9639) while H5 h≈44.56 (R^2=0.9878). This suggests that many ligand (RNA) molecules are needed for successful protein binding or, conversely, many proteins are bound to the same RNA molecule. In the second case, proteins can oligomerize in an RNA-dependent manner or can oligomerize independently before binding to the RNA.

From this assay, we cannot determine whether mutant proteins oligomerize in an RNA dependent or independent way, since the assay exploits the myc-tag and not an epitope on the protein to interact with the acceptor beads. In any case, the different oligomerization behaviour of the mutant proteins hints that their functionality is somehow changed by the mutations, although we cannot appreciate any differences given by the different domains subjected by the mutation.

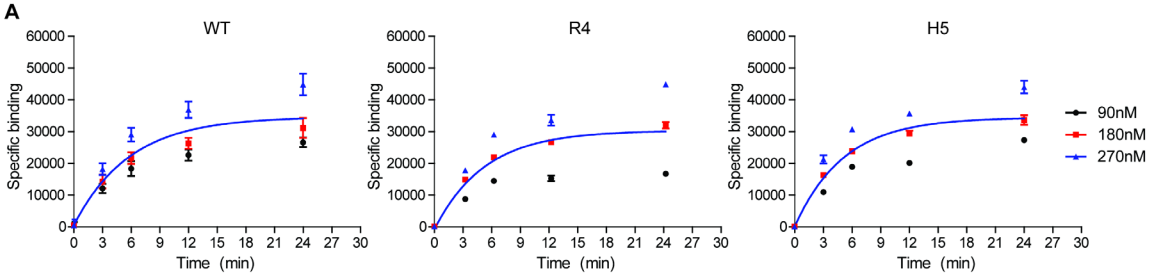
To confirm the previous results, we performed the inverse experiment, by keeping RNA concentration fixed (200nM) and changing the concentration of the proteins (90, 180, 270 nM). In this experiment, we monitored the kinetics of the binding at different time points (0,3,6,12 and 24 minutes after putting protein and RNA in contact) in order to calculate the association (K_{on}) and dissociation (K_{off}) rate constants of the binding. The ratio of K_{off} and K_{on} was then used to estimate the K_d of DHX30 (WT or mutant) binding to the probe. Fitting was performed on GraphPad Prism using the equation for association kinetics of two or more concentrations of ligand.

As reported in Figure 65, the binding of DHX30 to the WT probe is similar for all the protein variants considered, with estimated K_d ranging from 0.67nM to 0.72nM (WT and mutants respectively). These values were similar to those obtained previously, where the K_d was estimated at 1.26 nM for the WT motif probe (see Figure 62).

According to this experiment, the fluctuations in avidity among proteins are minimal, but mutant proteins have a slightly more rapid association to the RNA probe than the WT (K_{on}) paralleled by a slower dissociation (K_{off}), impacting the soluble protein fraction and possibly the protein-RNA turnover rate.

According to the results in Figure 64 and 65, the observation made by Lessel et al find a partial confirmation: mutant proteins have a higher tendency to oligomerize (Hill slope > 1) and have a lower protein-RNA turnover, possibly impacting translation as they noticed. Through this technique, it was not possible to appreciate any mutant-specific phenotype in the binding efficiency or the kinetics to the RNA probe.

A



B

	WT	R4	H5
K_{off}	0.069 min^{-1}	0.0802 min^{-1}	0.08173 min^{-1}
K_{on}	1.028 x 10^8 min^{-1}	1.11 x 10^8 min^{-1}	1.133 x 10^8 min^{-1}
K_d	0.6709 nM	0.7224 nM	0.7216 nM
R^2	0.89	0.98	0.98

Figure 65: Panel A: Binding kinetics of the WT and mutant DHX30 proteins to the WT CG-motif RNA probe. Panel B: Koff and Kon values obtained from the fitting; Kd was estimated as Koff/Kon; R2 indicates the goodness of fitting

Overall, we have proven that the ALPHA screen is a sensitive and reliable way to study the interaction of DHX30 and the RNA. We quantitatively demonstrated the specific and effective binding of the protein to CG-motif RNAs. The assay that we developed can be potentially used to test if small molecules are able to disrupt DHX30-RNA interactions in vitro for subsequent optimization as new therapeutic strategies in combination with p53-activating drugs to favour cell death responses over cell cycle arrest.

5) DISCUSSION

The complex array of responses that follow p53 activation depend on a multifactorial interplay between transcription of target genes, genetic and epigenetic background of the cell, nature of the event that triggered p53 activation and timing/magnitude of p53 accumulation. On top of this complexity, post-transcriptional gene regulation may refine the transcriptional-dependent responses by promoting or inhibiting the translation of certain subsets of mRNAs. This additional layer of control can depend on p53 as well, relying on the plethora of RBPs and miRNAs whose transcription in p53-dependent.

In order to investigate the extent post-transcriptional responses in the choice of cell fate, we exploited two cell lines which show an antithetic phenotype after non-genotoxic p53 activation with the small molecule Nutlin-3. Indeed, HCT116 colon-cancer cell line undergoes persistent cell cycle arrest while the osteosarcoma-derived SJSA1 cells proceed to apoptosis after an initial cell cycle arrest. For our purpose, we used the inhibitor of the p53-MDM2 interaction Nutlin-3 given its high specificity in triggering p53-dependent transcriptional responses [79,83], limiting the presence of confounders related to the activation of concomitant pathways, such as DNA damage responses, that follow p53 activation by genotoxic means. Notably, phosphorylation of p53 on key serine residues does not occur when cells are treated with Nutlin-3[116]. These post-translational modifications are instead associated with DNA-damage induced p53 activation, strongly supporting the non-genotoxic nature of Nutlin-mediated p53 activation.

Nonetheless, off-target and p53 transcription-independent effects elicited by Nutlin-3 cannot be completely excluded. As reported by Vaseva et al, apoptosis occurring after p53 activation by Nutlin-3 relies mainly on cytoplasmic p53 accumulation and translocation to the mitochondria, where it promotes MOMP [117]. In their model, it seems that p53 transcriptional program is dispensable and even able to reduce apoptosis. Additionally, by binding to the MDM2 pocket, Nutlin-3 can alter the functions that the ubiquitin-ligase has besides the regulation of p53. It was demonstrated that MDM2 can mono- or poly-ubiquitinate other proteins, such as other transcription factors and histones to regulate gene transcription[118] and its amplification is associated with increased polyploidy and genome instability [119]. Nonetheless, strictly MDM2-dependent responses that follow the treatment with Nutlin-3 have very little impact on cells' transcriptional programs or viability, as demonstrated in p53 null cell lines[59,83]. For these reason, the use of Nutlin-3 is perhaps the cleanest way to study p53 dependent phenotypes, thanks to the non-genotoxic nature of its activation.

To assess the relative contribution of the transcriptional and translational responses in the two model cell lines, we performed a comparative analysis of the total and polysome-bound differentially expressed mRNAs, respectively. Total RNA was used to account for transcription-dependent variations (transcriptome) while polysome-bound mRNAs were considered for translation-dependent differences (translatome). The analysis of transcriptome and translatome was performed after a

relatively short exposure (12 hours) of the cells to 10μM Nutlin-3 to exclude differences induced by delayed and p53-independent responses, like apoptosis initiation.

Despite the different tissue of origin, HCT116 and SJSA1 share a fair amount of coupled (ie: equally regulated in transcription and translation) genes, both up- and down-modulated (Figure 15B and 15D). Coupled genes encode for cell cycle arrest and apoptosis factors, without a striking bias towards the first or the latter in HCT116 and SJSA1 respectively (Figure 16A and 16C). Conversely, the analysis of polysome-bound regulated mRNAs show that the cell lines have already started a differential translation program (Figure 15C and 15E), showing only a marginal overlap of translationally up- or down-regulated mRNAs. If we consider the translatome as a good proxy of the proteome, this divergence anticipates and explains the phenotype difference that will become evident only a few hours later: SJSA1 have already started the translation of many apoptotic mRNAs contrary to HCT116 (Figure 16B and 18C). This observation reveals that the choice of the phenotype occurs early after p53-dependent responses started and it is heavily affected by post-transcriptional events that shape the translatome.

Previous reports identified a threshold mechanism as responsible for the choice between cell cycle arrest and apoptosis after p53 activation [82,120]. This threshold depends basically on the interplay between the expression of pro- and anti-apoptotic factors, which in turn depends on the levels of p53 protein. High level of p53 favour the transcription of weak p53 response elements (such as those found in the promoter of apoptotic factors)[48] and in the context of MDM2-overexpressing cells (like SJSA1), higher levels of active p53 can be reached faster compared to HCT116 after Nutlin-3, providing a plausible explanation for the phenotype difference. However, this does not seem to reflect the observation coming from the RNA sequencing, since pro- and anti-apoptotic genes are transcribed in both cell models at similar levels (Figure 17). Our analysis rather suggests that transcription alone cannot account completely for the difference between HCT116 and SJSA1 and it is rather post-transcriptional events that determine if the threshold needed to commit to apoptosis is reached or not.

We aimed next at investigating how features in the mRNAs can influence their translation starting from the motif enrichment analysis of the UTRs of the genes transcriptionally (total RNA) or translationally (polysome-bound mRNAs) regulated in both HCT116 and SJSA1. We identified the enrichment for a motif composed mainly by Cs and Gs (CG-motif) starting from the list of translationally upregulated genes in the apoptotic-prone cell line SJSA1 (Figure 18). We confirmed that genes containing the motif are actually more represented in the translatome of SJSA1 and that to a higher polysome association corresponds increased protein production (Figure 20). Moreover, a significant number of CG-motif containing mRNAs are involved in apoptosis, providing an explanation for the phenotype observed in SJSA1 cells.

We assessed how the presence of this motif in the 3'UTR is able to enhance the translation of the mRNAs containing it, as it was demonstrated by the luciferase assay, but only in SJSA1 and in a

p53-dependent manner (Figure 19). This prompted us to look for interactors that may promote or inhibit polysome loading of CG-motif containing mRNAs in a cell-specific context.

Proteins belonging to the PCBP family were investigated at first given their affinity for RNA sequences composed by a stretch of cytosines. They have proven roles in regulating gene expression at transcriptional[121] and post-transcriptional level [122]. PCBP2 was chosen for the differential protein expression levels in HCT116 and SJSA1. This protein has been implicated in promoting tumorigenesis in a variety of contexts [123,124,125] involving also the control of apoptosis by targeting miRNA 34a[126] and by promoting IRES-dependent translation of XIAP (in concert with PCBP1)[127]. In this view, higher expression levels of PCBP2 in HCT116 could account for their apoptosis-refractory phenotype after Nutlin-3 and thus being, in principle, a good candidate protein responsible for the phenotype choice. Additionally, PCBP2 promotes p73 stability[128], providing a functional link between the transcription program started by p53 (and other members of its family) and translation responses. Of note, p73 has been implicated in translation control and proteins synthesis in the context of cellular responses to oxidative stress cell responses to oxidative stress as a promoter of mRNA translation and protein synthesis[129].

We proved that PCBP2 can bind the CG-motif via pulldown using the consensus sequence (Figure 23B). We can speculate that the binding is sequence-specific since changing the rC triplet at the 3' of the consensus into rArGrA strongly reduces PCBP2 binding to the RNA probe. The interaction with CG-motif containing mRNAs was proved via RIP (Figure 24B) although the binding seems to be poorly Nutlin-3 treatment dependent.

Surprisingly, PCBP4 did not show the expected p53-dependent activation in either SJSA1 or HCT116 despite previous studies reporting its gene as containing p53-response elements [87]. Various factors can be responsible for the different observation, as the type of p53 activation (non-genotoxic with Nutlin-3 or DNA damage-induced with camptothecin) or the timing of the observation (12 hours after treatment vs 24 hours). We tried to analyse PCBP4 ability to interact with the motif after the pulldown, but only marginal differences in the binding could be noted via immunoblot between the cell lines (data not shown).

DHX30 was identified by mass spectrometry as a protein able to interact with the CG-consensus RNA probe in HCT116 cells specifically (Figure 22B). Subsequent experiments confirmed this specificity (Figure 23B), which can be strongly dependent on the differential expression levels of the protein, being more expressed in HCT116 cells compared to SJSA1. Similarly to PCBP2, the binding of DHX30 to the RNA probe depends on the sequence, since the mutated consensus strongly reduces DHX30 signal as revealed by western blot analysis after the pulldown. The binding affinity of DHX30 to the WT or mutated consensus probes was quantified in vitro using ALPHA assay, confirming a stronger binding of the protein to the wild-type consensus (K_d=1.2nM; Figure 62). Additionally, DHX30 binds CG-motif containing mRNAs, as demonstrated by RIP experiments (Figure 24C).

Helicases have been implicated in virtually all the steps of RNA biology, including splicing, translation and/or degradation, miRNA mediated silencing and their cytoplasmic storage in stress granules [130]. DHX30 has been poorly studied and characterized, but a few pieces of evidence of possible roles in translation have been discovered. Indeed, DHX30 is part of the mammalian ribo-interactome[95] and mutant protein overexpression results in spontaneous stress granules formation in the cell with a concomitant decrease in global translation[94]. A 5' alternative transcription start site in the DHX30 gene results in the use of an alternative first exon, which contains a mitochondria localization signal. It was demonstrated that a specific DHX30 isoform localizes to mitochondria[131] where it works as part of an RNP complex necessary for mtDNA transcription[89]. Additionally, in the mitochondrion DHX30 participates in post-transcriptional gene regulation by favouring mitochondrial ribosome assembly and translation[90].

We demonstrated that PCBP2 and DHX30 can interact in RNA-dependent mode (Figure 24A) and their binding sites on the 3'UTR of CG-motif containing mRNAs are mostly overlapping (Figure 23A). This supports the idea that both proteins are involved in the regulation of CG-motif containing mRNAs, most likely as part of a wider RBP complex (as assessed by mass spectrometry). This complex acts in trans on mRNAs containing the CG-motif by reducing their translation in HCT116 cells, which show constitutive high expression of both DHX30 and PCBP2.

However, the inter-dependency on either one or the other protein for successful binding to the RNA is not completely clear: PCBP2 seems to be necessary for DHX30 to efficiently bind to the CG-motif probe (Figure 25) and not vice-versa. According to this observation, depletion of PCBP2 would have a much stronger effect in abolishing the inhibitory effect of the RBP complex on CG-motif mRNAs translation, but the luciferase assay (Figure 26) and the subsequent experiment on HCT116 knockdown clones rather suggest that is DHX30 to have a determinant role for a successful translation inhibition of the CG-motif mRNAs.

Data coming from the RNA sequencing of the polysome-bound mRNAs in HCT116 clones depleted for PCBP2 or DHX30 support that these proteins have crucial roles in translation control: silencing of either one or the other protein results in a strong alteration of the translatome of HCT116 clones compared to controls (Figure 28). In accordance to our model, apoptosis-related genes are more represented in the translatome of shPCBP2 and shDHX30 clones (Figure 31A) and this occurs via a CG-like motif, in particular in the shDHX30 clones (Figure 31B). The panel of pro-apoptotic and/or CG-motif containing mRNAs investigated via qPCR confirmed that the enhanced polysome-association occurs independently from a transcriptional upregulation of the targets considered (Figure 32,33 and 34) and is followed by enhanced protein production in some cases (Figure 36). A consequence of the enhanced expression of pro-apoptotic genes results in sensitization of HCT116 shDHX30 cells to programmed cell death (Figure 37) but not in shPCBP2 clones.

How Nutlin-3 treatment (and p53) impacts on the phenotypes we have observed are still not completely clear: luciferase assay shows no translation enhancement of the luciferase activity in p53

KO clones (Figure 19C), but neither PCBP2 nor DHX30 show p53-dependent expression changes at RNA level or from total protein extracts. Additionally, the perturbation of one protein does not correlate with expression variations on the other. From the data we collected, the expression level (constitutive or RNAi-modulated) of a specific tissue or cell line for DHX30 and PCBP2 are the only factors responsible for the enhanced or decreased translation of CG-motif containing mRNAs. The absence of alterations in the luciferase assay after Nutlin-3 treatment of p53 null clones of HCT116 and SJSA1 supports once more the p53-dependency of the phenotypes we are investigating and suggests that p53-independent effects of Nutlin-3 have negligible contributions also in our model.

Despite being a good candidate protein in the post-transcriptional control of p53 dependent responses, PCBP2 showed only marginal effects in the ability to alter CG-motif containing RNA translation and to sensitize HCT116 cells towards apoptosis after Nutlin-3 treatment, although its perturbation resulted in shPCBP2 clones being significantly different from control and shDHX30 clones from a translatome point of view. It is tempting to speculate that other members of the PCBP family may act on the CG-motif in place of PCBP2 when the latter is depleted, although this aspect was not investigated.

On the other hand, data on DHX30 silencing revealed the helicase to have a more prominent role in enhancing CG-motif mRNA translation and slightly, but significantly, increasing apoptosis in knockdown cells. Whether the phenotype depends only on the translation of pro-apoptotic mRNAs is still being investigated in the laboratory by other people, who are more focused on the impact of DHX30 (and PCBP2) silencing on cell's metabolism, with particular attention to the aspects linked to mitochondrial biology.

Since HCT116 sensitization to apoptosis was not dramatically higher than control clones, we investigated whether a complete ablation of DHX30 could have a more determinant effect. Despite many trials with different CRISPR/Cas9 gRNAs and vectors, no homozygous knock out clone could be obtained, possibly confirming DHX30 as an essential gene[93] in human cells as well.

The results obtained in HCT116 clones prompted us to focus our effort only on DHX30, which is the most promising candidate protein in our model of translation regulation of apoptosis. Approaches of knock-down and overexpression only partially confirmed our expectations but supported anyway the importance of DHX30 in shaping translation and apoptotic-responses in the cell.

Knock-down of DHX30 in U2OS cells was the first approach that we attempted to generalize our model. In line with our expectation, U2OS show higher expression of DHX30 compared to the cognate cell line SJSA1, consistent with the fact that the first cell line is apoptosis-resistant after Nutlin-3 treatment. U2OS shDHX30 clones showed enhanced luciferase translation compared to shNT clones (Figure 42), but in opposition to the previous assays performed in parental HCT116/SJSA1 and HCT116 silenced clones, the addition of the CG-motif impacts on the luciferase activity already in the control clone. This observation reflects also on the levels of polysome-bound mRNAs that were tested by qPCR, besides the high variability on the data given by the normalization

that was chosen for this experiment (Figure 43B). According to qPCR, BAK1 and BCL2L1 expression is increased also in control shNT clones. Another observation is the constitutive effect of the silencing on CG-motif containing RNA expression: western blot revealed that all the targets we considered are expressed more in DMSO-treated cells of the shDHX30 clones (Figure 44) compared to the control clone. Surprisingly, all the targets that we analysed in U2OS derivative clones seem to be coupled, showing the same pattern of mRNA expression at total and polysome level. The reason why this happens is not known, but we can speculate that the time in which U2OS clones were collected and processed for analysis may be responsible for the loss of uncoupling between transcriptome and translatome. In the original experiment using HCT116 and SJSA1 parental cells, 12 hours-treatment time point was chosen to remove confounders given by secondary and late responses, since SJSA1 start showing signs of apoptosis activation 24 hours after the treatment [59,83]. In the context of U2OS, we were biased by the slower response kinetics of the derivative clones to the treatment with Nutlin-3 (Figure 39) and decided to focus on a later time point (24 hours instead of 12 hours) before collecting RNAs. We can speculate that RNA accumulation might be responsible for the increased quantification of the target mRNAs at total level (Figure 43A), since none of the four genes that we considered (*EIF5A, BAK1, CASP3* and *BCL2L1*) is a known direct p53 target, thus making a transcription-dependent enhancement of the mRNA level unlikely.

Whether the ultimate effect of DHX30 silencing and Nutlin-3 treatment in U2OS is really cell death is still not completely sure. U2OS shDHX30 show an inconsistent pattern of PARP and CASP3 cleavage and annexin V staining did not reveal any differential percentage in the population of healthy or apoptotic cells. The surrogate assays that we used (*Caspase-Glo 3/7 Assay System* and *RealTime-Glo Annexin V Apoptosis Assay*) point in the direction of a cell death-related process, but further elucidation is needed. On the other hand, the reproducible change in U2OS shDHX30 cells morphology that we detected at 48 and 72 hours after Nutlin-3 treatment (Figure 40) does not depend on EMT (Figure 41).

Having demonstrated how DHX30 silencing was effective in changing HCT116 and U2OS cells quite in accordance to the expectations of our model, the overexpression of the helicase in SJSA1 (resulting in decreased apoptosis) would have confirmed the centrality of DHX30 in the phenotype choice. Overexpression was not straightforward in this cell line, contrary to HEK293-T. After establishing doxycycline-inducible overexpression pools of SJSA1 cells, we tested their responses to the Nutlin-3 treatment. In SJSA1 cells overexpressing DHX30, only a partial phenotype rescue was obtained.

Analysis of the CG-motif targets by qPCR did not reveal decreased polysome-association (Figure 47B) as we would have expected and as suggested by the luciferase assay (Figure 47A). Similarly to the previous experiment in U2OS, the timing chosen for collecting the cells after the treatment (24 hours) might be too late to appreciate initial differences between transcriptome and translatome according to the model. In opposition to qPCR data, the expression of CG-motif containing targets

at protein level seems to be lower in DHX30-overexpressing cells compared to SJSA1 expressing constitutive levels of the helicase (Figure 48). To our surprise, doxycycline and Nutlin-3 had a synergistic effect on the overexpression of DHX30 (visible at 24 and more strongly at 48 hours). We wondered if this effect was due to a selection of DHX30-high cells, since we had a heterogeneous population of cells overexpressing the protein (and not single clones like in all the previous cell models). This would have reinforced our model, since cells surviving the treatment were those showing the highest level of DHX30 induction. Unfortunately, annexin V staining did not reveal any variation in the percentage of apoptotic cells also after doxycycline-dependent DHX30 overexpression.

We are now developing a doxycycline-dependent, GFP-tagged DHX30 overexpression system in SJSA1 to clarify if the helicase can really rescue cells fate preventing apoptosis. In fact, the dose of doxycycline that we use to induce DHX30 (2.5µg/ml) may be insufficient or excessive to obtain a biologically relevant phenotype in our model. The possibility to visualize overexpressed DHX30 localization and relative level of induction might help the fine tuning of the overexpression level to reach the levels and localization presented by the cognate cell line U2OS.

Secondly, SJSA1 are sensitive to lower Nutlin-3 doses than we used in these experiments, as the predicted IC50 is 3,44µM (according to the GDSC database[132]). The treatment with 10µM Nutlin-3, which corresponds to the dose used by Tovar et al[83] and Andrysik et al[59] and it is widely used in the field[12,133–135], may flatten weak protective effects exerted by DHX30 overexpression. Indeed, also HCT116 cells do not show a huge increase in the apoptotic population as a consequence of DHX30 silencing (Figure 37).

According to all the experiments and cell models that we used, the most reproducible and consistent results were obtained with the luciferase assay. Fold change differences are not strikingly marked, but they are always statistically significant, supporting the expected CG-motif dependent upregulation of translation. The easiest explanation for this efficacy is given by the optimal design of the artificial 3'UTR that is attached downstream to the *luc2* coding sequence. B-globin 3'UTR is relatively short (196 nucleotides) reducing the chances of confounders effects introduced by the interaction of the UTR with other RBPs. In addition, the consensus of the motif was added in two copies to the UTR. Having noticed that DHX30 (and also PCBP2) are extremely sensitive to variations of the consensus (see pulldown in parental HCT116/SJSA1 and the K_d calculated by ALPHA assay), it is not surprising that in the 3'UTR of endogenous genes the effects of the silencing (or conversely overexpression) are weaker than expected. As an example, *BAK1* has 5 instances of CG-like motifs, but they all vary in length and lack the central "CATGGCC" core that is present in the consensus and identified also in the motif enrichment analysis in HCT116 DHX30 clones (Figure 31B) In addition, longer 3'UTRs might be bound by other RBPs which contribute or alter mRNA fate and/or stability besides DHX30- (and PCBP2-) mediated effects.

In line with this, the luciferase assay suggested that also MYH9 has a role in the inhibitory complex that lowers the expression of CG-motif containing mRNAs (Figure 51A). After MYH9 knockdown in SJSA1, we noticed anticipation of the expression changes of apoptotic targets in response to the treatment (like BAK, whose expression is higher already at 12 hours) and a slightly higher annexin V positivity at 24 hours, which becomes significant after 48 hours (Figure 52). The exact mechanism by which MYH9 influences mRNA translation was not elucidated: the lack of recognized RNA binding motifs suggests that it participates indirectly to RNA binding, providing structural support to the inhibitory RBP complex or perhaps managing RNA localization in the cell. On the other hand, the protein was successfully detected after the pulldown using the WT or mutated RNA probe both in mass spectrometry (Figure 22B) and western blot analysis (Figure 49). As reported by Wang et al[96], MYH9 can successfully bind the lncRNA PTCSC2, so the protein can also have direct RNA binding capacities, besides the known functions in controlling cell migration, adhesion and maintaining cell integrity [136].

One important aspect that may have equal importance in the phenotype choice is how cells manage cell cycle arrest response. This part was not investigated for two reasons mainly, the first being that SJSA1 undergo initial cell cycle arrest as well and the second being that no nucleotide motif nor miRNA seed enrichments were identified in the lists of parental HCT116 cells. This complicates the analysis of cell cycle arrest responses in the perspective of a post-transcriptional contribution, basically because no common functional features could account for a similar fate of functionally-related mRNAs, like in the case of pro-apoptotic mRNAs containing the CG-motif. Nevertheless, we recently acquired HCT116 clones depleted for p21 and we plan to investigate if and how the lack of a strong cell cycle arrest regulator influences cell's propensity towards apoptosis.

The last two sections of the results report experiments that were not directly linked to a generalization or expansion of the model, but rather understanding the contribution of transcriptional and translational responses in the induction of apoptosis (section 4.6) and to test DHX30 ability to bind RNA in vitro, in order to develop a benchmark assay to test disruptors of DHX30-RNA interactions (section 4.7). BAK and BAX are multi-BH domain proteins that convey the signals given by others pro- and anti-apoptotic BCL2-family proteins to induce mitochondrial outer membrane permeabilization and start the apoptotic cascade. p53 controls the transcriptional expression of many pro-apoptotic members of the BCL2 family, like PUMA, NOXA, BID and also BAX[50]. On the other hand, *BAK1* was identified as a translationally enhanced gene in SJSA1 cells. We wanted to exploit conformation specific antibodies[114] to test the relative contribution of the proteins in the induction of MOMP, using BAX and BAX as a proxy of transcription-dependent and translation-dependent responses, respectively. Results were surprising since we noticed that both proteins undergo activation in the apoptosis-resistant HCT116 cells as well (Figure 53 and 54). Strikingly, in this cell line, the activation seems to be even slightly stronger and to start earlier than in SJSA1 cells (at least in the case of BAK activation, Figure 55). This suggests that some proteins acting downstream of

BAK/BAX activation and/or downstream of MOMP are differentially expressed or regulated in HCT116 vs SJSA1 cells, determining the different phenotype upon Nutlin-3 treatment. From a transcriptional point of view, pro- and anti-apoptotic targets show very little differences in the two model cell lines (Figure 66), suggesting that differential regulation of BAK/BAX activation occurs at the protein level. New investigation on the composition of the interactome of the active proteins will help understand what induces cell line specific inhibition of apoptosis. According to Paek et al[82], the contribution of the IAP proteins is mainly responsible for the inhibition of apoptosis in HCT116 after genotoxic p53 activation by Nutlin-3. This can explain why HCT116 show BAK/BAX activation without proceeding in the apoptotic cascade since IAPs act on caspases preventing their activation[137]. Nevertheless, MOMP is often considered the point of no return for cells apoptotic commitment[138], thus a deeper investigation of what happens in HCT116 after BAK and BAX activation is warranted.

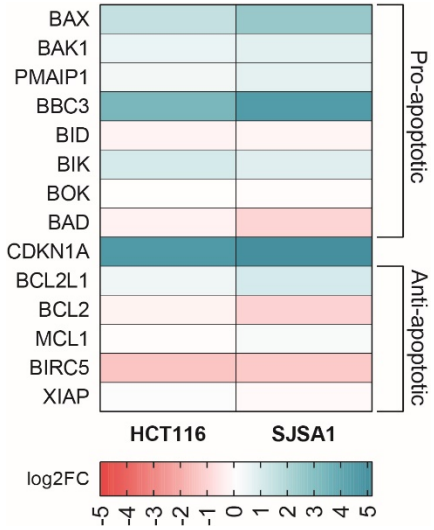


Figure 66: RNA-sequencing of the expression of selected pro- and anti-apoptotic mRNAs after 12 hours of Nutlin-3 in HCT116 and SJSA1 cells. Results are presented as log2 fold change over control (DMSO). CDKN1A was included to assess the successful activation of a p53 -dependent programme

The other approach that we used to weight the relative contribution of BAK and BAX to apoptosis was the knockout of a single or both proteins in HCT116 and SJSA1 using CRISPR/Cas9. As expected, HCT116 cells did not show any increased resistance to apoptosis after Nutlin-3. Only a mild effect was noticed in the BAK and double KO clones in the number of PI-positive cells, regardless of the treatment (Figure 58). Looking at the experiments on SJSA1 we were surprised by the fact that single KO clones showed no protective effect towards apoptosis: derivative clones had a similar number of annexin V positive cells to the parental cell line. Conversely, double KO cells were effectively protected from apoptosis (Figure 60). Although it has been reported that only one among BAK and BAX is sufficient to induce apoptosis [139,140], the two proteins do not overlap completely in their functions in some contexts [114]. As stated before for the DHX30 overexpression, the high dose of Nutlin-3 that we use may flatten subtle protective effects given by a single knockout. It will be interesting to assess how SJSA1 overexpressing DHX30 or with single KO respond to lower, sub-lethal doses of Nutlin-3.

The final topic of this thesis focuses on the in vitro studies of DHX30. Protein cloning and overexpression in HEK293-T cells were successfully obtained. In this cell line, DHX30 could be overexpressed easily using a constitutive and strong promoter, contrary to SJSA1 that required the transduction of an inducible system to obtain reproducible results. ALPHA technology was efficient and sensitive to quantify the strength of the interaction between the helicase and the RNA probes (Figure 62), confirming the higher affinity of the protein towards the WT consensus compared to the mutant version, as previously observed on the immunoblot after the pulldown experiments (Figure

23B). We also confirmed the specificity of DHX30 towards motifs containing mostly Cs and Gs, since the negative (rich in As) and the AU-rich probes strongly impaired the binding. Overall, DHX30 shows a strong binding affinity to the CG-consensus (K_d = 1.2nM), opening the possibility for the discovery and development of efficient inhibitors of the interaction with the RNA. ALPHA technology was already used successfully in a similar context (inhibition of HuR, another RNA binding protein[141]) as a relatively quick and highly sensitive assay to screen for small molecules able to disrupt protein-RNA interactions.

The same experimental setup was exploited also for testing the effects of rare missense mutations associated with a neurological phenotype in humans [94]. Mutated amino acids fall in the RNA binding domain (Arg493His, R4) and on the ATPase domain (His562Arg, H5). According to the experiments performed by Lessel et al, the R4 mutant shows impaired RNA binding ability, although this is not completely abolished, while H5 mutant shows decreased capacity to hydrolyse ATP. Nevertheless, both mutants show increased association to stress granules also in non-stressed conditions (while the wild type protein co-localizes to stress granules in response to exogenous stressors), resulting in lower rates of protein translation in cells overexpressing mutant forms of the helicase.

According to the ALPHA experiments, WT and mutant proteins retain RNA binding ability and have a similar affinity towards RNA in vitro (Figure 65). Slight differences in the protein-RNA turnover were noticed, with mutant proteins having slower turnover rates compared to WT DHX30. Surprisingly, R4 mutant did not show any striking binding impairment to the RNA. We can speculate that the amino acid substitution with another that has the same charge have only mild effects on the protein ability to bind RNA, at least in vitro.

Notably, both mutants showed a higher tendency to oligomerize (Figure 66) compared to the WT. This may resemble the phenotype observed by Lessel et al in the formation of stress granules containing DHX30. Whether this oligomerization is RNA dependent cannot be assessed using this experimental approach: exploiting the exogenous myc-tag allows the protein to oligomerize prior to the RNA binding or, conversely, proteins oligomerize once they have recognized the same RNA probe, resulting in the same amplification of the acceptor beads emission. Using antibodies raised against DHX30 epitopes on acceptor beads can be helpful to clarify how mutant proteins oligomerization occurs.

Overall, DHX30 mutant proteins do not show massive RNA-binding impairments according to this in vitro assay. Their tendency to oligomerize was probably recapitulated in vitro, although further investigation on their behaviour is needed, both in vivo and in vitro. Unfortunately, in the perspective of screening for small molecules, these mutants are not good candidates for counter-screening since their binding affinity to the RNA is comparable to the WT helicase.

6) CONCLUSION AND FUTURE PERSPECTIVES

After 40 years from its discovery and thousands of articles, p53 has not revealed all its secrets. Its perception has dramatically changed over the years as new discoveries were made, most notably it has become the “guardian of the genome” after having been considered an oncogene for a decade. Only recently its roles beyond transcription are starting to emerge, adding new complexity to an already intricated network of gene expression modulation as a sequence-specific transcription factor. Integrating data coming from multi-omics approaches seem to be effective for a comprehensive understanding of p53 dependent responses, taking into considerations multiple layers of control where p53 is implicated. Of the many ways that can induce p53 activity, Nutlin-3 is probably the most effective to analyse p53-dependent responses without bias and confounders introduced, for example, by concomitant DNA damage responses.

Here we report how post-transcriptional gene regulation following p53 activation can produce different phenotypic outcomes to the same treatment with the MDM2 inhibitor Nutlin-3. This regulation depends on a cis-acting element (CG-motif) embedded in the 3’UTR of some apoptosis-related genes and a trans-acting complex of RBPs. We identified DHX30 as one of the proteins mostly responsible for determining polysome loading of CG-motif containing RNAs, ultimately leading to apoptosis (in DHX30-low contexts) or preventing it (DHX30-high context). The exact mechanism of DHX30-dependent translation inhibition, together with the interactions that this helicase has with other RBPs (like PCBP2) or interactors of the CG-motif that we identified (like MYH9) still needs to be dissected.

From our observations, only the expression level of DHX30 seems to be predictive of the outcome to Nutlin-3 treatment. We expect that other RBPs, whose expression might be modulated by p53, are acting in concert with DHX30, PCBP2 and MYH9 to successfully lower (or enhance) polysome loading of CG-motif containing mRNAs, although they are still to be determined. Whether DHX30 levels could be generally used as a predictor for successful induction of p53-dependent apoptotic responses needs to be investigated, but the data we collected so far seems to point in this direction. A graphic summary of the p53-dependent translation regulation of CG-motif containing mRNAs is reported in Figure 67.

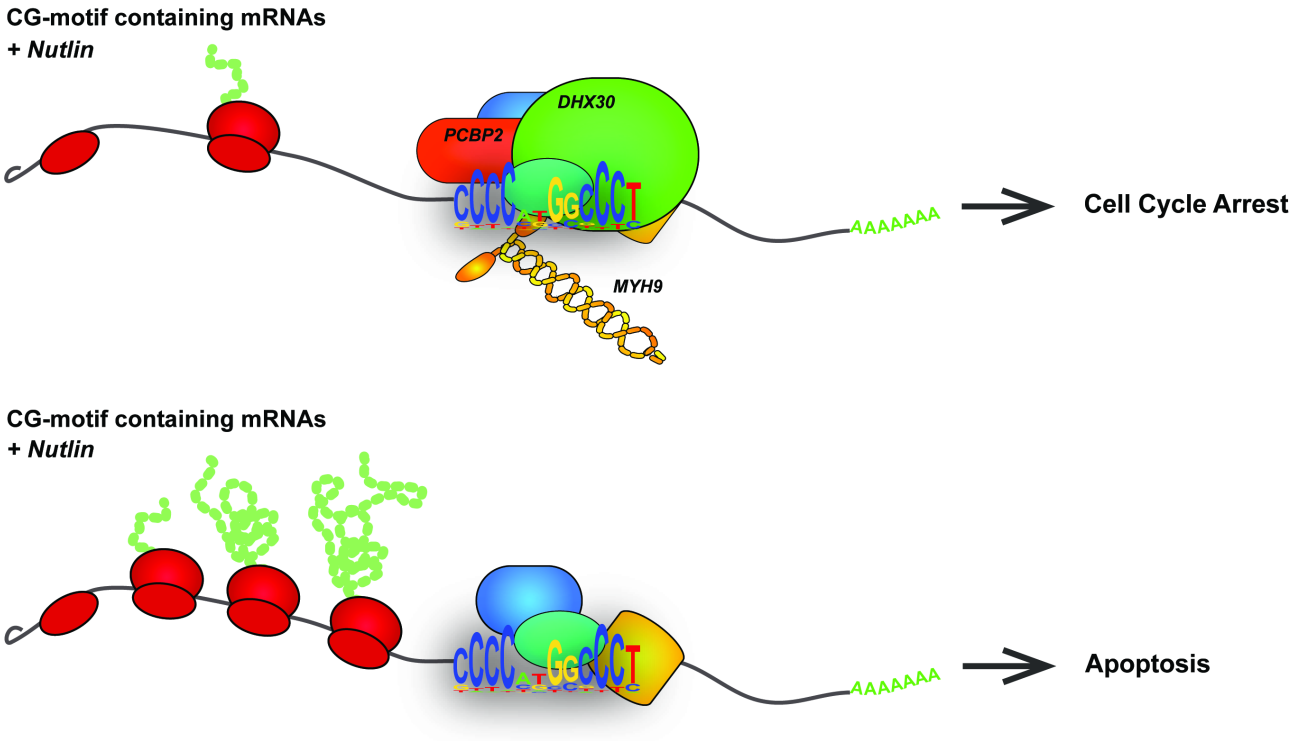


Figure 67: Schematic representation of the post-transcriptional regulation of CG-motif containing mRNAs mediated by DHX30, PCBP2 and MYH9

Nevertheless, the successful induction of apoptosis does not rely only on transcription or translation control of apoptotic proteins: BAK and BAX activation analysis revealed that other factors acting downstream their activation to blocking apoptotic pathway in HCT116 cells and not in SJSA1. Indeed, both BAX and BAX show a similar activation pattern and timing in both cell contexts, despite the different phenotype that they show.

Finally, we developed a sensitive and effective in-vitro assay to test the binding ability of DHX30 to CG-motif RNA probes based on ALPHA technology. This platform will be exploited for screening compounds able to disrupt DHX30-RNA interaction, promoting cell death responses in cell lines after p53 activation with Nutlin-3, as we have obtained by RNAi approaches.

The simultaneous discovery of mechanisms that control apoptosis in translation and downstream of BAK and BAX activation will increase the chances to identify the core regulatory factors able to control apoptotic responses after p53 activation. In the perspective of cancer treatment, a combination of non-genotoxic p53 activating therapies (with Nutlin-3 derivatives) and concomitant promotion of apoptotic responses (DHX30 inhibition and/or promotion of BAK-BAX dependent MOMP) would be beneficial for successful cancer remission, minimizing the risk of disease relapse.

7) MATERIALS AND METHODS

Cell line culture and treatments

HCT116 and SJSA1 cells were grown in RPMI medium added with 10% FBS, 100 units/ml penicillin, 100 mg/ml streptomycin and 2mM L-glutamine. U2OS and HEK293-T cells were grown in DMEM medium added with FBS, antibiotics and glutamine as reported above. Cells were kept at 37°C in a humified atmosphere with 5% CO_2.

Cells were treated with 10μM Nutlin-3 (Sigma Aldrich; Cayman Chemical) when they reached 60-70% confluence. Nutlin-3 stock (10mM) was previously diluted in FBS-free media before adding it to the cell culture. An equal volume of vehicle DMSO was diluted in FBS-free media and added to the cells as control.

Doxycycline treatment was performed on SJSA1 cells when they reached 50%-60% confluence and at least 12 hours before Nutlin-3 treatment. The volume corresponding to 2.5μg/ml doxycycline or an equivalent volume of ethanol was added directly to the cell culture media.

Polysome profiling and RNA extraction

After 12 or 24 hours of DMSO/Nutlin-3 treatment cells were treated with cycloheximide (1:1000; stock 10mg/ml) for 10 minutes at 37°C. Media was removed and cells were washed twice with cold 1X PBS added with 1:10000 cycloheximide and collected using a scraper. Cells were pelleted (200g for 5 minutes) and the supernatant was replaced by a variable volume of lysis buffer (1X sodium deoxycholate; 1X Triton-X; 10mM Tris-HCl pH 7.5; 10mM NaCl, 10mM $MgCl_2$, 0.2U RNase inhibitors, 10mM DTT, 100μg/ml cycloheximide). Lysates were left 5 minutes on ice and centrifuged for 5 minutes at 16000g to pellet the nuclei. Supernatants were loaded on top of sucrose gradients before starting ultracentrifugation.

Sucrose gradients were prepared by solubilizing 50% weight/volume sucrose (Sigma) in Buffer A (30mM Tris-HCl pH 7.5; 100mM NaCl, 10mM MgCl2 in DEPC H_2O). 15% sucrose solution was obtained by diluting 50% solution in Buffer A. Polyallomer tubes (Thermo Fischer Scientific) were soaked in 3% H_2O_2 in DEPC H_2O and dried and then 6 ml of 50% sucrose solution were poured on the tube. 15% sucrose solution was carefully layered on top of the 50% solution and the tube was closed using parafilm. Tubes were moved to a custom-designed gradient maker, that automatically tilted the tubes to reach the tubes in a horizontal position in 15 minutes, held the tubes for 90 minutes in horizontal position and returned them to a vertical position in 15 minutes. After loading the lysates, volume differences between the tubes were normalized using Buffer A. Gradients were gently moved to buckets of the SW41Ti rotor (Beckman Coulter) and ultracentrifuged for 100 minutes at 40 KRPM at 4°C.

Gradients were fractionated using a Teledyne ISCO Foxy fraction collector in 1ml fractions following absorbance spectrum at 260nm. Fractions corresponding to the sub-polysomal or the polysome were collected together and processed as described in Chapter 4.3.4 of the following manual [142].

Briefly, pooled fractions were supplemented with 100ug/ml proteinase K and 1% SDS and left at 37°C for 2 hours. Phenol:chloroform 5:1 pH 4.7 (Ambion) was added to the samples (1:4 of the volume), vortexed and centrifuged at 16000g for 5 minutes. The aqueous phase was collected and combined with an equal volume of isopropanol + 1µl Glycoblue (Ambion), and left at -80°C overnight. Samples were centrifuged for 30-40 minutes at 16000g at 4°C and the supernatant was removed. Pellets were resuspended in fresh cold 70% EtOH and centrifuged at 16000g at 4°C for 15 minutes. Supernatants were removed, pellets were air-dried and resuspended in 20-40 µl nuclease-free water (Thermo Fischer Scientific).

Data analysis (HCT116 and SJSA1 dataset)

Samples (two biological replicates for each condition) were sequenced using Ion Proton System (Thermo Fischer Scientific), producing 60-80M raw reads per sample. Sequencing data were quality filtered and trimmed (minimum read length: 30nt; maximum read length: 150nt) by FASTX-Toolkit. Mapping to the hg19 build of the human genome (February 2009 GRCh37, NCBI Build 37.1) was performed using GSNAP [143]. Mapping parameters allowed for 0.03% mismatches. To compute the per-gene reads counts, we used HTSeq[144]. We chose the intersection nonempty mode with reads overlapping more than one exonic feature. The DESeq R package [145,146] was used to call Differentially Expressed Genes (DEGs) starting from two replicates for each condition and comparing each fraction with itself across conditions (e.g. Nutlin-3 total versus DMSO total; Nutlin-3 polysomal vs DMSO polysomal). For all analyses on DEGs, two thresholds were set: (1) log2FC > 1 and < -1 for up-regulated and down-regulated genes, respectively; (2) FDR-corrected p-value <0.1. We defined coupled DEGs as genes which met the indicated thresholds at both the total and polysomal RNA level. Translationally enhanced or reduced genes are genes that met the indicated thresholds in the polysomal fraction but are not significantly modulated at the total RNA level. Unchanged in translation DEGs are genes that follow the reported thresholds in the sub-polysomal fraction, but not in the polysomal one independently from their expression change at the total RNA level.

Library preparation and data analysis (HCT116 clones dataset)

Polysome-bound mRNAs extracted as described above were quantified using Qubit RNA BR Assay (Thermo Fischer Scientific) and RNA quality was assessed using the Agilent RNA 6000 Nano kit with an Agilent Bioanalyzer. Library preparation was started from 1.5 µg of RNA from samples scoring an RNA integrity number (RIN) ≥ 8.9 using TruSeq RNA Library Prep Kit v2 (Illumina). Samples (four biological replicates each condition) were sequenced using Illumina HiSeq 2500, producing 25-28M raw reads per sample. After quality filtering and trimming with trimmomatic [147] (minimum quality 30, minimum length 36nt) we quantified each Gencode v25 (http://www.gencodegenes.org/releases/) transcript by means of Salmon[148]. edgeR[149] was used to

call Differentially Expressed Genes (DEGs), considering a significance threshold on the corrected p-value <0.05 and a $\log_2$ Fold Change (Nutlin-3/DMSO) > 0 (upregulated) or < 0 (downregulated).

Metascape and GSEA analysis

DEGs belonging to the total and polysome-bound lists were prepared to be uploaded as a unique Microsoft Excel file (different samples were grouped in different columns of the spreadsheet) to the Metascape[84] portal (http://metascape.org/gp/index.html). Analysis was performed as Multiple Gene List, selecting Reactome Gene Sets, Canonical Pathways, GO Biological Processes and Hallmark Gene Sets as references for enrichment. Significance of the enrichment was set at 0.05.

Motif discovery

To search for common sequence motifs in each category of coupled or uncoupled DEGs, we used Weeder[85], setting the following parameters: (1) longest 5' or 3'UTR of each gene as input, (2) motifs had to be found in at least 25% of all input sequences, (3) the motif could have a length ranging from 6 to 12nt, (4) motif search had to be performed on the sense strand only. Background for Weeder consisted in the frequency of all possible n-mers (with n=6,8,10,12) in the human genome, as provided by the Weeder package. When the analyses were completed, the "adviser" program was started to select the best motifs, especially according to their redundancy. Resulting motifs were compared by computing the Pearson correlation of their positional weight matrices by means of the TFBStools R package[150].

Luciferase assay

Full-length β-globin 3'-UTR, β-globin 3'-UTR with the addition of a CG-motif at the 3'UTR beginning and/or at the 3'UTR end were cloned at the Firefly luciferase end of a pGL4.13 vector using XbaI as restriction enzyme. Dual-luciferase reporter assay was done according to the manufacturer's instructions (Promega).

Briefly, cells were plated on a 24-well plate and then transfected with 30µg of Firefly luciferase containing vectors (control pGL 4.13, β-globin 3'-UTR or β-globin 3'-UTR with the addition of the CG-motif) using FugeneHD (Promega) in a 1:3 ratio (DNA weight : transfectant volume). As transfection control 5µg of pRL Renilla Luciferase vector was used in each well. After 24 hours, cells were treated with DMSO or Nutlin-3 for 12, 24 or 48 hours. Cells were then washed twice with 1X PBS, lysed using passive lysis buffer and 10µl of each lysate were used for the assay. Luciferase activity was measured in triplicate wells using Infinite 200 PRO microplate reader (TECAN). Firefly luciferase values were normalized on the Renilla luciferase units (RLU). To highlight changes upon treatment, data are then normalized on the DMSO condition. RLUs of the CG-motif containing vector were normalized over the RLU of the pGL empty vector and over β -globin vector. From the same samples, RNA was also extracted using Trizol-based extraction (see the details in the following sections). After

qPCR, luc2 quantification was used to further normalize RLUs. Statistical significance was calculated using Student t-test with Microsoft Excel.

Pulldown

Two in vitro synthesized RNA probes with a 5'-UTR biotinylation were purchased by IDT. Their sequences match the wild-type motif sequence (5'-GAAGGGCCCU CCCCAUGGCCCU GGAGAGUGGG-3'), or the mutant one (5'-GAAGGGCCCU CCCCAUGGAGAU GGAGAGUGGG-3'). These probes were used to perform the pull-down experiment as previously reported[151]. Briefly, after washing and immobilization of 0.50mg/reaction of streptavidin-conjugated beads (Dynabeads M-280, Thermo Fisher Scientific) according to the manufacturer's instruction, the binding of 80 pmol RNA probes to the beads was obtained by incubation for 20 minutes at room temperature using gentle rotation. After washing (Wash buffer: 200mM Tris-HCl pH 7,5; 500mM NaCl; 20mM $MgCl_2$; 1% Tween 20 in DEPC H_2O), 500-600 µg of protein extract was added and incubated for 1h at 4°C to the streptavidin beads now binding the biotinylated RNA probe. After washing, the bound complex was detached from the beads heating at 95°C for 7 minutes and then loaded on an 8% or 4-12% polyacrylamide gel.

In the case of mass spectrometry analysis, the gel was stained by Colloidal Coomassie[152]. The visual comparison of bands from the SJSA1 and HCT116 cell pull-down upon DMSO or Nutlin treatment was used to select a range of protein size for quantitative mass spectrometry that showed the most evident difference between the two cell lines. Label-free quantification was performed at the IFOM Proteomics facility (Milan). We filtered the list of obtained proteins according to two parameters: 1) number of matching peptides greater than 2; 2) sequence coverage greater than 10%. When the LFQ LC-MS experiment was performed at the proteomics facility at CIBIO, the same parameters were used for peptide call and selection.

In the case of pulldown followed by immunoblot, standard blotting procedure was followed (see next sections). Input samples (30 µg) were loaded as control.

RNA extraction

Total RNAs were extracted with Illustra RNAspin Kit (GE Healthcare) according to the manufacturer instructions. When column-based extraction was not possible, Trizol-based methods were used.

Cells lysates were added with equivalent or higher volumes of Trizol reagent (Thermo Fisher Scientific) and left at room temperature for 3 minutes. Chloroform was added (1:5 of the Trizol volume), samples were vortexed and centrifuged at 4°C for 10 minutes at 16000g. The upper aqueous phase was collected and transferred to a new tube. An equal or higher volume of isopropanol was added along with 1.5µl of Glycoblue co-precipitant (Ambion). Samples were centrifuged again at 4°C for 10 minutes at 16000g, supernatant was removed and a washing in EtOH

was performed. Pellets were air-dried and resuspended in 20-40µl RNase free water. Quantification was performed using Nanodrop 1000 (Thermo Fisher Scientific).

RT-qPCR

RNAs extracted from polysome profiling, Trizol or column-purified were converted to cDNA using the RevertAid First Strand cDNA Synthesis Kit (Thermo Fisher Scientific), following manufacturer's instructions. Depending on the quantity of RNA (determined using Nanodrop) 0.3 to 2µg of RNA were converted. Samples were diluted to the same final concentration (12,5 or 10ng/µl) assuming 100% RNA to cDNA conversion efficiency.

KAPA SYBR® FAST qPCR Master Mix (Kapa Biosystems) or qPCRBIO SyGreen Mix (PCR Biosystems) were used as master mix for qPCR adding 200nM forward and reverse primers and 25 or 20ng of cDNA. The reaction was performed on CFX96 or CFX384 thermal cyclers (Biorad) in technical triplicate for each gene/sample. Raw data were processed with Biorad CFX Manager software to calculate Ct. Normalization was calculated using 2^{-dCt} method over the housekeeping gene[153] (YWHAZ or GAPDH) and the fold change was calculated as 2^{-ddCt} using Microsoft Excel. Statistical significance was determined using Student T-test in Microsoft Excel or GraphPad Prism of at least two independent biological replicates (three technical replicates for each biological replicate). The list of primers used for qPCR is reported as supplementary materials (Section 11.2).

Western blot

Cell pellets were collected from the plates using trypsin and washed once with 1X PBS before proceeding with lysis using RIPA-buffer (50 mM Tris HCl pH8, 150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS) supplemented with 1X protease inhibitors (cOmplete EDTA-free Protease Inhibitor Cocktail, Roche/Sigma Aldrich). Lysis was performed for 1 hour in rotation at 4°C and debris were collected by centrifugation (12 minutes, 13500g, 4°C). Supernatants containing proteins were quantified using BCA Assay (Thermo Fischer Scientific/Euroclone) according to the manufacturer's protocol. A defined quantity of protein (30 or 20 µg) was mixed with NuPAGE LDS Sample Buffer + 10% DTT (Thermo Fischer Scientific) before denaturation (95°C for 5 minutes) and loading on a polyacrylamide gel. SDS-PAGE was performed using Tris-glycine buffer (25mM Tris, 192mM glycine) containing 0.5% SDS.

Blotting was performed on nitrocellulose membranes using Towbin buffer (25mM Tris, 192mM glycine, 20% methanol) for 2 hours at 300mA or 3 hours at 200mA at 4°C. Membranes were stained with Ponceau S solution (Sigma Aldrich) before blocking in 5% milk powder in 1X PBS + 0,1% Tween20 (PBS-T) for 1 hour with gentle rotation. Primary antibodies were diluted in 1% milk in PBS-T and incubated overnight at 4°C in rotation. After four washing steps with PBS-T, membranes were incubated with HRP-conjugated secondary antibodies (diluted in 1% milk in PBS-T) for 1 hour at room temperature. Membranes were washed four times before proceeding with detection with

Amersham ECL Select (GE Healthcare). Pictures were acquired using Uvitec Alliance L2 acquisition system and processed with Alliance software.
Images reported in the results are representative of at least two independent experiments yielding similar results. The list of antibodies used and the dilutions are reported in Section 11.5.

Viral vectors production

HEK293-T cells were seeded on P150 dishes at 7.5 x 10^6 cells density. The following day media was replaced with 15ml OptiMEM (Gibco) containing 17.5µg pCMV-Δ8.9, 7.5µg pCMV-VSV-g and 25µg of the desired plasmid (based on the pLKO or pCW 57.1 backbones) and 1:2 weight:volume PEI 2X. After 24 hours OptiMEM was removed and replaced with 18ml complete DMEM. After 48 hours media were collected, centrifuged (500g for 5 minutes) and filtered using 0.45µm syringe filters. The filtrate was aliquoted and stored at -80°C. 10µl of the filtrate were used for quantification of the viral vector load (RTU) using qPCR [154].

DHX30, MYH9 and PCBP2 silencing

HCT116, U2OS and SJSA1 were seeded on a 6-well plate and left 24 hours in the incubator to allow attachment. The following day 1 ml of media was removed and replaced with a volume of viral vectors corresponding to 1 RTU. Cells were spin-inoculated for 2 hours at 1600g and left in the incubator for 24 hours. Spin-inoculation was not performed on SJSA1 and U2OS cells. Lentivirus-containing media was removed, and cells were left to recover for 48 hours before proceeding with puromycin selection (1:4000, stock 1mg/ml), that was kept until the non-transduced control was completely dead. Puromycin dose was progressively lowered but used to maintain selection at low doses (1:10000). Pools were subsequently tested for the successful silencing before proceeding with limiting dilution and single cell-clone selection. The vectors used for silencing are reported in the supplementary materials (section 11.4).

RNA immunoprecipitation (RIP)

Cells were seeded on P100 dishes and processed when they reached 80-90% confluence (either with or without treatment. Crosslinking was performed by exposing plates to UV light using a (0.150 J/m^2 for 15 seconds) keeping the plates on ice, before proceeding with cells washings (twice with cold 1X PBS on ice) and collection using a scraper. After pelleting and removing the supernatant, 500µl Lysis buffer (100mM KCl, 5mM MgCl2, 10mM HEPES pH7, 0,5% NP-40, 1mM DTT, 1U/µl RNase Inhibitors, 1X Protease Inhibitor Cocktail) were added to the pellets. Lysates were transferred in a falcon tube, left at least 2 hours at -80°C, centrifuged at 10000rpm for 30 minutes and supernatants were collected in a new tube. Dynabeads ProteinA or ProteinG (depending on the antibody species) (Thermo Fisher Scientific) were prepared by washing them twice with NT2 Buffer (50mM Tris-HCl pH7.4, 150mM NaCl, 1mM MgCl2, 0.05%NP40) and resuspended in NT2 buffer

with 1% BSA. Beads were distributed in different tubes, diluted adding twice their initial volume of NT2 Buffer + 1% BSA. 4-5µg of the specific antibody (PCBP2 or DHX30) or IgGs were added. Beads were left for 2 hours on a wheel at 4°C, washed three times on a magnet with 1ml NT2 buffer and resuspended in 200µl NT2 buffer + 200U RNase Inhibitors + 1 mM DTT. Lysates were pre-cleared by adding a mix containing ProteinA and ProteinG Dynabeads in equal amount and left for 20 min at 4°C on a wheel. After placing the tubes on a magnet, supernatants were collected and 10% of their volume was used as input to be directly
extracted with Trizol. The remaining supernatant was mixed with the antibody-coated beads and incubated overnight on a wheel at 4°C. Samples were centrifuged at 2000rpm for 1min at 4°C, 100µl of the supernatant were used for WB analysis of the unbound and the rest discarded. Beads were resuspended in 1ml NT2 buffer, transferred in a new tube and washed with 1ml NT2 buffer for 10 minutes on a wheel at 4°C. Three more washes were performed with 1ml NT2 Buffer added with 0,1% Urea + 50mM NaCl (10 minutes, 4°C on a wheel). Beads were washed one more time in 500µl of NT2 buffer, 50µl were collected for WB analysis and the remaining supernatant was discarded. All the washes were performed on a magnet. RNA was extracted by adding Trizol to the beads, the RNA pellet was resuspended in 15µl DEPC water.

Co-immunoprecipitation (CO-IP)

Protein G Dynabeads (Thermo Fisher Scientific) were washed twice with 200µl of CHAPS buffer. Then, beads were resuspended in 200µl of CHAPS buffer in the presence of proteases and RNase inhibitors and 3 µg of PCBP2 antibody. Incubation was performed at 4°C for 1 hour to let the antibody attach to the beads. Beads-conjugated antibodies were washed with 1 mL CHAPS buffer to discard excess of unbound
antibody. 1 mg of cell lysate previously obtained using CHAPS buffer in the presence of protease inhibitors and RNase inhibitors and gently rotating at 4°C for 30 minutes was added to the beads-antibody conjugates. Incubation was performed overnight at 4°C with gentle rotation of the samples. To test whether the binding of the antibody was RNA-dependent, before the incubation, one of the sample lysates was also pre-treated with 10 µg/mL RNase A for 10 minutes at 37 °C. The complexes were then washed three times with CHAPS buffer rotating them for 5 minutes at 4°C. The samples were then resuspended in 30µl of loading buffer with 2% Urea and 20% DTT. After denaturation at 95°C for 8 minutes, the eluate was loaded on the 8% acrylamide gel. As a comparison, 30 µg of input cell lysate were also loaded using the same buffer conditions. Blotting and detection were performed as reported above.

Annexin V staining

Cells were detached from the plates using trypsin, washed once with 1X PBS and counted. $1.5x10^5$ cells were resuspended in 100µl Annexin V binding buffer 1X (10mM HEPES pH 7.4; 140mM NaCl;

2.5mM $CaCl_2$) and incubated with 2µl FITC-Annexin V antibody and 5µl PI solution, provided by the FITC Annexin V Apoptosis Detection Kit (BD). Cells were incubated for 15 minutes at room temperature in the dark before adding 400µl Annexin V binding buffer 1X. Acquisition of at least 10000 events was performed using BD FACS CantoA. Data analysis was performed using BD FACS Diva or FlowJo software. Statistical significance was calculated using Microsoft Excel applying Student T-test on at least two independent biological replicates.

Growth curves and quantification of cell death markers using high-content imaging.

U2OS cells (shNT and shDHX30 clones) were seeded in a 96-well plate format (5000 cells per well, in triplicate). HCT116 and SJSA1 were seeded in 24-well format (3000 cells per well, in duplicate). Cells were imaged using digital phase contrast (DPC) by Operetta High Content Imaging System (Perkin Elmer) 24 hours after seeding, right after the treatment with 5 or 10µM Nutlin-3 or DMSO solvent control. Plates were then imaged again on the same fields at different time points. For HCT116 and SJSA1 growth curves, PI (0.25µg/ml) and Hoescht 33342 (10µg/ml) were added to the media 30 minutes hour before endpoint image acquisition. Images were processed using Harmony® High Content Imaging and Analysis Software (Perkin Elmer) to calculate the number of cells per well at each time point. DPC-count at time point 0 (before the treatment) was used to normalize data.

PI positivity of U2OS cells was quantified by seeding cells in a 96-well plate format (3000 cells per well, in quadruplicate), stained with Propidium Iodide, Calcein, and Hoechst and images acquired by Operetta were processed using Harmony High Content Imaging and Analysis Software to quantify the numbers of total objects, the mean cell size, the number and proportion of PI-positive cells.

RealTime-Glo Annexin V Apoptosis Assay

U2OS were seeded on a 96 well format (2000 cells/well, in duplicate). Reagents for apoptosis reaction were part of the RealTime-Glo Annexin V Apoptosis Detection Kit (Promega). Media containing the apoptosis detection reagents and 10µM Nutlin-3 or an equivalent volume of DMSO were added to the wells, according to the manufacturer instructions. Luminescence was quantified using Tecan Infinite M200Pro at T=0 (right after the treatment) and at subsequent time points. Relative luciferase units (RLUs) are normalized over DMSO treated control at each time point.

DHX30 cloning

DHX30 coding sequence was amplified using the primers reported in section 11.3 starting from the cDNA of HCT116 cells. PCR product was purified using Illustra GFR PCR DNA and Gel Band Purification Kits (GE Healthcare). Both the PCR product and the pCMV6-AC-Myc-His backbone were digested using SgfI and MluI enzymes (Fermentas), purified and ligated. Ligation products were transformed into DH5α competent cells and grown on ampicillin-containing plates. Colonies were

inoculated in LB+ampicillin and subsequently tested after MINI-prep plasmid extraction (EuroGold Plasmid Miniprep Kit – EUROCLONE). Promising clones, confirmed by Sanger sequencing (Eurofins Genomics) were inoculated for MIDI-prep (PureYield Plasmid Midiprep System – Promega) and plasmid DNA was used for HEK293-T overexpression.

Cloning in the pCW-57.1 backbone was performed by amplification of DHX30 coding sequence from the previously obtained pCMV6-AC-Myc-His+DHX30 vector using the primers reported in section 11.3. PCR product was digested using AgeI and XbaI and destination vector was linearized using NheI and AgeI. Ligation was performed exploiting NheI/XbaI compatible ends and the correct insertion was verified by digestion and Sanger sequencing.

Mutant DHX30 sequence was inserted in both pCMV6-AC-Myc-His and pCW-57.1 vector by performing digestion of the destination vector with NheI. Mutant DHX30 sequence was obtained from the plasmids provided by Davor Lessel and his group by digestion of the pEGFP+DHX30 plasmids with NheI. Destination plasmid and donor fragment were gel purified and ligated. The correct orientation was checked via digestion with NheI and StuI and confirmed by Sanger sequencing.

DHX30 overexpression in SJSA1

The pCW-57.1+DHX30 vector was used for the production of lentiviral vectors (see above) that were used for transduction of SJSA1 cells. Transduction procedure was the same described for gene silencing (see above) but in this case we tested the transduction with 1 and 2 RTUs. Pools showing the best overexpression pattern after puromycin selection were chosen for subsequent experiments without proceeding to single clone dilution.

DHX30 purification

HEK293-T cells were seeded on P150 culture dishes at $7.5x10^6$ concentration and left growing for 24 hours. Transfection was performed the following day using 12µg plasmid of interest and PEI 2X in 1:2 ratio (weight:volume) with DNA. After 48 hours cells were washed once with 1X PBS and collected using Equilibration Buffer (25mM Tris-HCl pH 8, 100mM NaCl, 0,05% NP40; 1mM DTT, 3mM MgCl2; 10mM Imidazole in DEPC H_2O) into a falcon tube and frozen at -80°C for at least 2 hours. Thawed lysates were supplemented with protease inhibitors and sonicated for 50 seconds – 1 minute (10-second pulses, power ≥ 200W) on ice. The sonication was repeated until the solution was clear. Lysates were centrifuged (13000g for 20 minutes at 4°C) and transferred in a new Falcon tube, where fresh Equilibration buffer was added to reach 15 ml. 250 µl of Ni Sepharose High-Performance beads (GE Healthcare) were added to the solution and incubated overnight at 4°C with gentle rotation. Beads were pelleted (500g for 1 minute), the supernatant was discarded and beads were moved to 1.5ml Eppendorf tubes using 1 ml Equilibration buffer, that was removed after a centrifugation step (500g 1 minute).

Beads were washed with four different Washing buffers, whose composition is reported in the table below this paragraph. After each buffer additions, beads were resuspended, gently rotated for 5 minutes in the cold and centrifuged for 1 minute at 500g before discarding the supernatant and adding the new buffer. Protein elution was performed by incubating the beads for 15 minutes in rotation using 500µl Elution buffer (25mM Tris-HCl pH8, 250mM NaCl, 0.05% NP40, 1mM DTT, 3mM $MgCl_2$, 400mM imidazole). After centrifugation the supernatant was collected, added with glycerol (final concentration 10%) and stored at -80°C.

Dialysis and size exclusion was performed using Amicon Ultra 15 100K filters (Millipore). The protein solution was loaded onto the column and centrifuged for 20 minutes at 16000g at 4°C. The flow-through was discarded and the rest of the solution was collected, quantified using Nanodrop 1000 and used for ALPHA assays.

	Wash 1	Wash 2	Wash 3	Wash 4
Tris-Cl pH 8	25 mM	25 mM	25 mM	25 mM
NaCl	200 mM	300 mM	300 mM	500 mM
NP40	0.05%	0.05%	0.05%	0.05%
DTT	1 mM	1 mM	1 mM	1 mM
MgCl2	3 mM	3 mM	3 mM	3 mM
Imidazole	50 mM	50 mM	100 mM	70 mM

Composition of the washing buffers used for DHX30 protein purification

ALPHA assay

Active fraction of the protein was assessed by diluting the protein at different concentrations in ALPHA buffer (25mM Hepes pH 8, 50mM NaCl, 2mM MgCl2, 50µM $ZnCl_2$, 100µM ATP, 0.05% BSA) and dispensed in the wells of a 384-well OptiPlate (Perkin Elmer). RNA was diluted at a fixed concentration (200nM) and added to the wells. Acceptor beads previously diluted in ALPHA buffer to a final concentration of 100ng/µl were added to the wells, the plate was centrifuged (150g for 30 seconds) and left on an orbital shaker for 15 minutes at RT in the dark. Diluted donor beads (final concentration 100ng/µl) were added to the wells, left 45 minutes in the dark with shaking at room temperature. ALPHA signal was detected using Perkin Elmer EnSpire Multimode plate reader.

Protein-RNA binding at increasing concentration of RNA was performed similarly to what is described above, but protein concentration was kept fixed and different dilutions of the RNA probe were prepared.

For kinetics experiment, a unique mix with ALPHA buffer, RNA (at fixed concentration), donor and acceptor beads was prepared and added to the wells of the OptiPlate. Protein dilutions were added to the wells at different time points before proceeding with the detection using EnSpire.

Data were processed with Microsoft Excel and GraphPad Prism was used for curve fitting.

The sequence of the RNA probes used is reported in section 11.6.

Intracellular BAK and BAX staining

Cells were seeded on a 6-well plate or 60mm dishes and treated as described above. For the staining $1.5x10^5$ cells were counted, washed in 1X PBS and fixed with 2% PFA for 30 minutes at 4°C. After centrifugation (500g for 4 minutes) cells were washed twice in 1X PBS before incubation of the primary antibody (0.5 µg, BAK 6A7 – Santa Cruz and BAX AB1 - Millipore) previously diluted in Buffer S (1% FCS, 1% Saponin in PBS). A washing step was performed using Buffer W (1% FCS in PBS) before secondary antibody incubation (1:200, Goat anti-Rabbit IgG Alexa-488, Thermo Fisher Scientific). Two additional washing steps in Buffer W were performed before resuspending cells in 1X PBS + 2nM EDTA. Data were acquired using Attune NxT Flow Cytometer (Thermo Fisher Scientific) and processed using FlowJo.

8) APPENDIX: lncEPR

During my PhD, I worked on a side project focused on the characterization of a murine long non-coding RNA (lncRNA) named EPR (epithelial cell program regulator). This lncRNA has been discovered by Rossi et al[155] as an interactor of KHSRP in the murine mammary gland cell line NMuMg. KHSRP is an RNA binding protein involved in mRNA splicing and decay, miRNA maturation and lncRNA-mediated gene expression regulation[156].

The expression of lncEPR is strongly and quickly downregulated when cells are exposed to TGF-β. This downregulation depends on SMAD3 binding to the EPR promoter with a corepressor complex. As a consequence of TGF-β-dependent downregulation, the binding of the lncEPR to KHSRP is decreased. It was recently discovered by the same group that KHSRP functions as a hub to reduce TGF-β-mediated epithelial to mesenchymal transition in the same cell line[157].

In order to determine the role of lncEPR in preventing TGF-β-mediated EMT, NMuMg clones stably overexpressing lncEPR were obtained. Overexpression resulted in upregulation of epithelial markers, reduction of migratory potential and morphological changes to the cells, suggesting an important role in preventing the mesenchymal transformation of the cells under the influence of TGF-β.

Since the lncRNA contains a putative 213 nucleotide-long open reading frame, we were asked to analyse the presence of EPR on the polysomes of NMuMg parental cells (Figure A1 and Supplementary Fig 3a of the paper). Contrary to the previous experiments presented in the main text of this thesis (sections 4.1.1 and 4.3.3), here we pooled the fractions corresponding to free RNA and 40S (a), sub-polysome fractions 60S and 80S (b), light polysomes (c) and heavy polysomes (d). Results of the qPCR analysis showed that indeed lncEPR is polysome-associated (both light and heavy polysomes) suggesting its capability to produce a small polypeptide. It was later demonstrated by Rossi and coworkers that the protein is actually produced and it is likely to function in epithelial junctions.

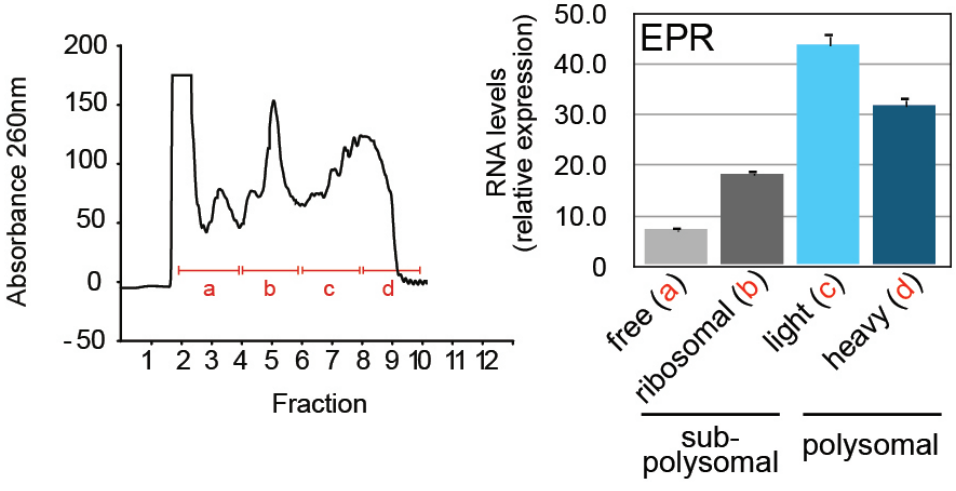


Figure A1: Polysome fractionation of NMuMg cell extracts. Right panel, quantification via qPCR of EPR presence in the different pools of fractions collected.

At the same time, the authors analysed the lncEPR-mediated functions independently from the peptide. For this reason, they obtained an NMuMg clone overexpressing a mutated version of EPR containing a stop codon (EPRSTOPE). Analysis of the transcriptome of the WT or Stop-codon variants of the lnc EPR revealed little differences, suggesting that the peptide has marginal roles in transcription. In fact, both EPR variants induced significant variations at a transcriptome level, particular of those transcripts involved in the regulation of epithelial cell proliferation (like *Cdh1, Cdh2, Tjp1, Tnc, Twist2* and *Ocln,* see Figure A3A and Fig.3B of the paper). Of note, EPR overexpression enhanced *Cdkn1a* expression, resulting in a strongly reduced proliferation of the cells. We quantified the impact of EPR overexpression on cell proliferation using Operetta high content imaging system (Figure A2A and Fig.3H in the paper) and confirmed the p21-dependent cell cycle arrest by means of FACS analysis (Figure A2B and Fig.3I in the paper). To reach a higher sensitivity on establishing the percentage of cells in S phase we used Click-IT Plus EdU Flow Cytometry Assay (Invitrogen). Both techniques confirmed a slower proliferation rate when NMuMg overexpress lncEPR (WT or mutated) and the increased proportion of G1 cells strongly suggests a p21-dependent arrest of the cell cycle.

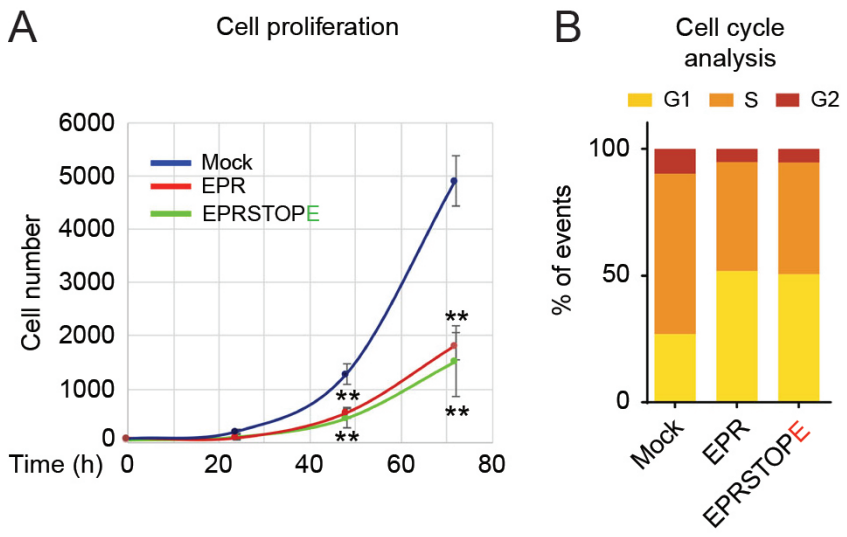


*Figure A2: A) Growth curve of the control (mock), lncEPR WT overexpressing (EPR) or lncEPR mutated (EPRSTOPE) NMuMg cells. B) Cell cycle distribution detected by EdU staining of proliferating cells and subsequent FACS analysis. ** = p-value < 0,01 (Student T-test)*

To confirm that the modulation of epithelial reprogramming factors was not due to the EPR-mediated p21 accumulation and higher fraction of cells in G1, we performed nuclei staining of NMuMg cells (control and overexpressing WT and Stop-mutated lncEPR) using vital Hoescht 33342 followed by FACS sorting of the cells in G1 phase. qPCR results confirmed that the level of expression of the genes involved in epithelial reprogramming are equally expressed in the whole population (Figure A3A and paper Fig.3B) or in the G1-arrested subpopulation (Figure A3B and supplementary Figure 4 of the paper).

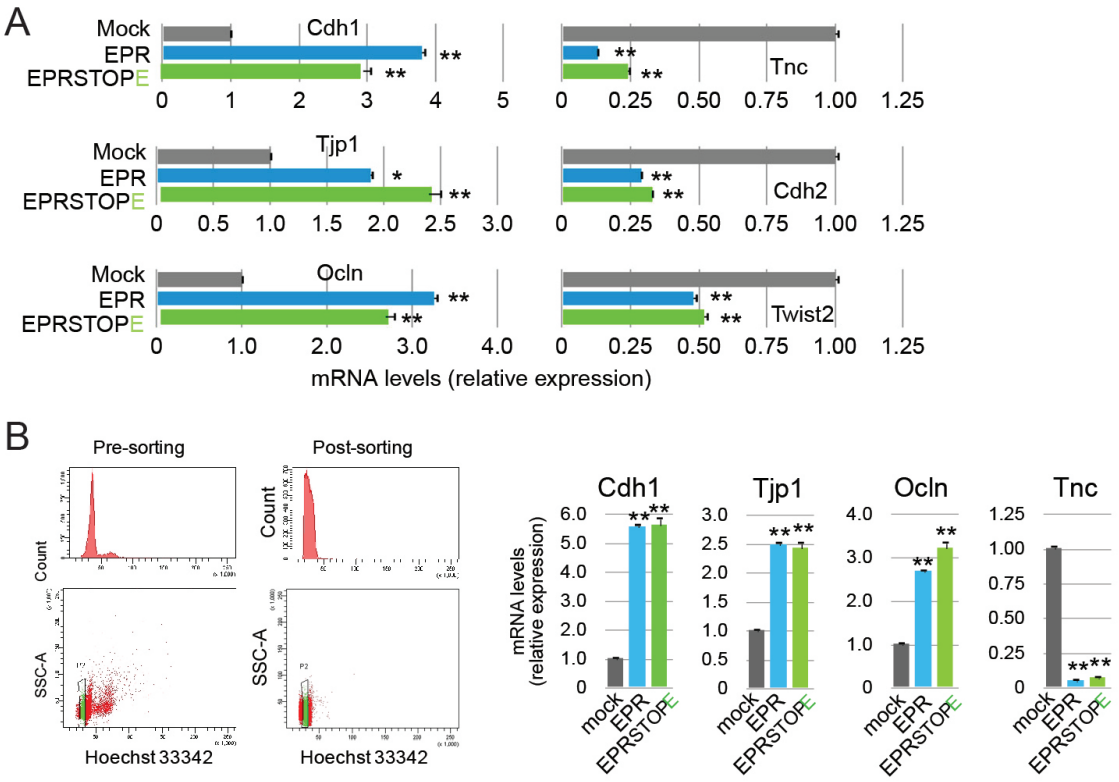


Figure A3: A) Expression profiles of a panel of gene involved in epithelial reprogramming on NMuMg control cells (mock), overexpressing WT lncEPR (EPR) or stop codon mutated lncEPR (EPRSTOPE). B) Sorting of NMuMg after Hoescht staining of the nuclei. On the right panel example of cells profiles pre and post sorting, confirming the good quality of sorting for G1 arrested cells; on the right panel expression of selected genes involved in epithelial reprogramming. Expression profiles match those of the whole population (Panel A)

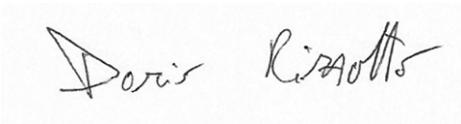
The co-authors then confirmed the role of lncEPR in interacting with *Cdkn1a* promoter and sustaining its transcription. Moreover, it lncEPR when overexpressed counteracts TGF-β downregulation of *Cdkn1a*. On top of this, the interaction of lncEPR with KHSRP allows Cdkn1a mRNA to be less subject to KHSRP degradation, overall increasing p21 expression.

Altogether, results suggest that lncEPR functions by dampening TGF-β driven EMT stimulation and proliferation through enhanced expression and stabilization of Cdkn1a mRNA. According to public available dataset, the expression of the human variant of this lnc (h.EPR) positively correlates with less aggressive subtypes of breast cancer and is highly underrepresented in triple negative cancers. h.EPR ectopic overexpression in human cell lines causes slower growth rates and reduced tumor growth when cells are injected in BALB/C mice, confirming its tumor suppressive capabilities.

More information can be found on the paper, reported in section 11.7

9) DECLARATION OF ORIGINAL AUTHORSHIP

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



10) REFERENCES

1. Lane, D. P. & Crawfor, L. V. T antigen is bound to a host protein in SV40-transformed cells. *Nature* **278**, 261–263 (1979).

2. Linzer, D. I. H. & Levine, A. J. Characterization of a 54K Dalton cellular SV40 tumor antigen present in SV40-transformed cells and uninfected embryonal carcinoma cells. *Cell* **17**, 43–52 (1979).

3. Oren, M., Maltzman, W. & Levine, A. J. Post-translational regulation of the 54K cellular tumor antigen in normal and transformed cells. *Mol. Cell. Biol.* **1**, 101–110 (1981).

4. Crawford, L. V, Pim, D. C., Gurney, E. G., Goodfellow, P. & Taylor-Papadimitriou, J. Detection of a common feature in several human tumor cell lines--a 53,000-dalton protein. *Proc. Natl. Acad. Sci. U. S. A.* **78**, 41–5 (1981).

5. Finlay, C. A., Hinds, P. W. & Levine, A. J. The p53 proto-oncogene can act as a suppressor of transformation. *Cell* **57**, 1083–1093 (1989).

6. Lane, D. P. p53, guardian of the genome. *Nature* **358**, 15–16 (1992).

7. Sullivan, K. D., Galbraith, M. D., Andrysik, Z. & Espinosa, J. M. Mechanisms of transcriptional regulation by p53. *Cell Death Differ.* **25**, 133–143 (2018).

8. Lakin, N. D. & Jackson, S. P. Regulation of p53 in response to DNA damage. *Oncogene* **18**, 7644–7655 (1999).

9. Levine, A. J., Hu, W. & Feng, Z. The P53 pathway: What questions remain to be explored? *Cell Death and Differentiation* **13**, 1027–1036 (2006).

10. Kastenhuber, E. R. & Lowe, S. W. Putting p53 in Context. *Cell* **170**, 1062–1078 (2017).

11. Horn, H. F. & Vousden, K. H. Coping with stress: multiple ways to activate p53. *Oncogene* **26**, 1306–1316 (2007).

12. Williams, A. B. & Schumacher, B. p53 in the DNA-damage-repair process. *Cold Spring Harb. Perspect. Med.* **6**, a026070 (2016).

13. Vousden, K. H. & Lu, X. Live or let die: the cell's response to p53. *Nat. Rev. Cancer* **2**, 594–604 (2002).

14. Harper, J. W., Adami, G. R., Wei, N., Keyomarsi, K. & Elledge, S. J. The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. *Cell* **75**, 805–16 (1993).

15. el-Deiry, W. S. *et al.* WAF1, a potential mediator of p53 tumor suppression. *Cell* **75**, 817–25 (1993).

16. Chen, J. The Cell-Cycle Arrest and Apoptotic Functions of p53 in Tumor Initiation and Progression. *Cold Spring Harb. Perspect. Med.* **6**, a026104 (2016).

17. Georgakilas, A. G., Martin, O. A. & Bonner, W. M. p21: A Two-Faced Genome Guardian. *Trends Mol. Med.* **23**, 310–319 (2017).

18. Hermeking, H. *et al.* 14-3-3sigma is a p53-regulated inhibitor of G2/M progression. *Mol. Cell* **1**, 3–11 (1997).

19. Zhan, Q. *et al.* Association with Cdc2 and inhibition of Cdc2/Cyclin B1 kinase activity by the p53-regulated protein Gadd45. *Oncogene* **18**, 2892–2900 (1999).

20. Jayasurya, R. *et al.* Phenotypic alterations in Rb pathway have more prognostic influence than p53 pathway proteins in oral carcinoma. *Mod. Pathol.* **18**, 1056–1066 (2005).

21. Villunger, A. p53- and Drug-Induced Apoptotic Responses Mediated by BH3-Only Proteins Puma and Noxa. *Science (80-.).* **302**, 1036–1038 (2003).

22. Hofseth, L. J., Hussain, S. P. & Harris, C. C. p53: 25 years after its discovery. *Trends Pharmacol. Sci.* **25**, 177–181 (2004).

23. Letai, A. Apoptosis and Cancer. *Annu. Rev. Cancer Biol.* **1**, 275–294 (2017).

24. Adrain, C. Apoptosis-associated release of Smac/DIABLO from mitochondria requires active caspases

and is blocked by Bcl-2. *EMBO J.* **20**, 6627–6636 (2001).

25. Zou, H., Li, Y., Liu, X. & Wang, X. An APAF-1·Cytochrome c Multimeric Complex Is a Functional Apoptosome That Activates Procaspase-9. *J. Biol. Chem.* **274**, 11549–11556 (1999).

26. Müller, M. *et al.* p53 Activates the CD95 (APO-1/Fas) Gene in Response to DNA Damage by Anticancer Drugs. *J. Exp. Med.* **188**, 2033–2045 (1998).

27. Liu, X., Yue, P., Khuri, F. R. & Sun, S.-Y. p53 Upregulates Death Receptor 4 Expression through an Intronic p53 Binding Site. *Cancer Res.* **64**, 5078–5083 (2004).

28. Sheikh, M. S. *et al.* p53-dependent and -independent regulation of the death receptor KILLER/DR5 gene expression in response to genotoxic stress and tumor necrosis factor alpha. *Cancer Res.* **58**, 1593–8 (1998).

29. Aubrey, B. J., Kelly, G. L., Janic, A., Herold, M. J. & Strasser, A. How does p53 induce apoptosis and how does this relate to p53-mediated tumour suppression? *Cell Death Differ.* 1–10 (2017). doi:10.1038/cdd.2017.169

30. Mirza, A. *et al.* Human survivin is negatively regulated by wild-type p53 and participates in p53-dependent apoptotic pathway. *Oncogene* **21**, 2613–2622 (2002).

31. Altieri, D. C. Survivin and IAP proteins in cell-death mechanisms. *Biochem. J.* **430**, 199–205 (2010).

32. Mihara, M. *et al.* P53 Has a Direct Apoptogenic Role At the Mitochondria. *Mol. Cell* **11**, 577–90 (2003).

33. Chipuk, J. E. *et al.* Direct activation of Bax by p53 mediates mitochondrial membrane permeabilization and apoptosis. *Science* **303**, 1010–4 (2004).

34. Leu, J. I. J., Dumont, P., Hafey, M., Murphy, M. E. & George, D. L. Mitochondrial p53 activates Bak and causes disruption of a Bak-Mcl1 complex. *Nat. Cell Biol.* **6**, 443–450 (2004).

35. Tomita, Y. *et al.* WT p53, but not tumor-derived mutants, bind to Bcl2 via the DNA binding domain and induce mitochondrial permeabilization. *J. Biol. Chem.* **281**, 8600–8606 (2006).

36. Youle, R. J. & van der Bliek, A. M. Mitochondrial Fission, Fusion, and Stress. *Science (80-.).* **337**, 1062–1065 (2012).

37. Wang, W. *et al.* Mitofusin-2 is a novel direct target of p53. *Biochem. Biophys. Res. Commun.* **400**, 587–592 (2010).

38. Karbowski, M., Norris, K. L., Cleland, M. M., Jeong, S.-Y. & Youle, R. J. Role of Bax and Bak in mitochondrial morphogenesis. *Nature* **443**, 658–662 (2006).

39. Bieging, K. T., Mello, S. S. & Attardi, L. D. Unravelling mechanisms of p53-mediated tumour suppression. *Nat. Rev. Cancer* **14**, 359–370 (2014).

40. Vousden, K. H. & Prives, C. Blinded by the Light: The Growing Complexity of p53. *Cell* **137**, 413–431 (2009).

41. Tisato, V., Voltan, R., Gonelli, A., Secchiero, P. & Zauli, G. MDM2/X inhibitors under clinical evaluation: perspectives for the management of hematological malignancies and pediatric cancer. *J. Hematol. Oncol.* **10**, 133 (2017).

42. Shieh, S.-Y., Ikeda, M., Taya, Y. & Prives, C. DNA Damage-Induced Phosphorylation of p53 Alleviates Inhibition by MDM2. *Cell* **91**, 325–334 (1997).

43. Lozano, G. Mouse Models of p53 Functions. *Cold Spring Harb. Perspect. Biol.* **2**, a001115–a001115 (2010).

44. Wang, S., Zhao, Y., Aguilar, A., Bernard, D. & Yang, C. Targeting the MDM2–p53 Protein–Protein Interaction for New Cancer Therapy: Progress and Challenges. *Cold Spring Harb. Perspect. Med.* **7**, a026245 (2017).

45. Joerger, A. C. & Fersht, A. R. Structural Biology of the Tumor Suppressor p53. *Annu. Rev. Biochem.* **77**, 557–582 (2008).

46. May, P. & May, E. Twenty years of p53 research: structural and functional aspects of the p53 protein.

Oncogene **18**, 7621–7636 (1999).

47. Joerger, A. C. & Fersht, A. R. Structure–function–rescue: the diverse nature of common p53 cancer mutants. *Oncogene* **26**, 2226–2242 (2007).

48. Menendez, D., Inga, A. & Resnick, M. A. The expanding universe of p53 targets. *Nat. Rev. Cancer* **9**, 724–737 (2009).

49. Nguyen, T.-A. T. *et al.* Revealing a human p53 universe. *Nucleic Acids Res.* **46**, 8153–8167 (2018).

50. Fischer, M. Census and evaluation of p53 target genes. *Oncogene* **36**, 3943–3956 (2017).

51. Ciribilli, Y. *et al.* The Coordinated P53 and Estrogen Receptor Cis-Regulation at an FLT1 Promoter SNP Is Specific to Genotoxic Stress and Estrogenic Compound. *PLoS One* **5**, e10236 (2010).

52. Bisio, A. *et al.* Cooperative interactions between p53 and NF-kB enhance cell plasticity. *Oncotarget* **5**, 12111–12125 (2014).

53. Collavin, L., Lunardi, A. & Del Sal, G. p53-family proteins and their regulators: hubs and spokes in tumor suppression. *Cell Death Differ.* **17**, 901–911 (2010).

54. Mack, D. H., Vartikar, J., Pipas, J. M. & Laimins, L. A. Specific repression of TATA-mediated but not initiator-mediated transcription by wild-type p53. *Nature* **363**, 281–283 (1993).

55. Fischer, M., Quaas, M., Steiner, L. & Engeland, K. The p53-p21-DREAM-CDE/CHR pathway regulates G 2 /M cell cycle genes. *Nucleic Acids Res.* **44**, 164–174 (2016).

56. Hoffman, W. H., Biade, S., Zilfou, J. T., Chen, J. & Murphy, M. Transcriptional Repression of the Anti-apoptotic survivin Gene by Wild Type p53. *J. Biol. Chem.* **277**, 3247–3257 (2002).

57. Hermeking, H. MicroRNAs in the p53 network: micromanagement of tumour suppression. *Nat. Rev. Cancer* **12**, 613–626 (2012).

58. Engeland, K. Cell cycle arrest through indirect transcriptional repression by p53: I have a DREAM. *Cell Death Differ.* **25**, 114–132 (2018).

59. Andrysik, Z. *et al.* Identification of a core TP53 transcriptional program with highly distributed tumor suppressive activity. *Genome Res.* **27**, 1645–1657 (2017).

60. Menendez, D. *et al.* Diverse stresses dramatically alter genome-wide p53 binding and transactivation landscape in human cancer cells. *Nucleic Acids Res.* **41**, 7286–7301 (2013).

61. Bykov, V. J. N., Eriksson, S. E., Bianchi, J. & Wiman, K. G. Targeting mutant p53 for efficient cancer therapy. *Nat. Rev. Cancer* **18**, 89–102 (2018).

62. Kandoth, C. *et al.* Mutational landscape and significance across 12 major cancer types. *Nature* **502**, 333–339 (2013).

63. Liu, Y. *et al.* Deletions linked to TP53 loss drive cancer through p53-independent mechanisms. *Nature* **531**, 471–475 (2016).

64. Leroy, B., Anderson, M. & Soussi, T. TP53 Mutations in Human Cancer: Database Reassessment and Prospects for the Next Decade. *Hum. Mutat.* **35**, 672–688 (2014).

65. Shirole, N. H. *et al.* TP53 exon-6 truncating mutations produce separation of function isoforms with pro-tumorigenic functions. *Elife* **5**, 1–25 (2016).

66. Giacomazzi, J. *et al.* TP53p.R337H is a conditional cancer-predisposing mutation: further evidence from a homozygous patient. *BMC Cancer* **13**, 187 (2013).

67. Freeman, J. A. & Espinosa, J. M. The impact of post-transcriptional regulation in the p53 network. *Brief. Funct. Genomics* **12**, 46–57 (2013).

68. Hermeking, H. p53 Enters the MicroRNA World. *Cancer Cell* **12**, 414–418 (2007).

69. Lucchesi, C., Zhang, J. & Chen, X. Modulation of the p53 family network by RNA-binding proteins. *Transl. Cancer Res.* **5**, 676–684 (2016).

70. Zaccara, S. *et al.* P53-Directed Translational Control Can Shape and Expand the Universe of P53 Target Genes. *Cell Death Differ.* **21**, 1–13 (2014).

71. White, R. J. RNA polymerases I and III, non-coding RNAs and cancer. *Trends Genet.* **24**, 622–629 (2008).

72. Marcel, V. *et al.* p53 Acts as a Safeguard of Translational Control by Regulating Fibrillarin and rRNA Methylation in Cancer. *Cancer Cell* **24**, 318–330 (2013).

73. Marcel, V., Catez, F. & Diaz, J. J. P53, a translational regulator: Contribution to its tumour-suppressor activity. *Oncogene* **34**, 5513–5523 (2015).

74. Loayza-Puch, F. *et al.* p53 induces transcriptional and translational programs to suppress cell proliferation and growth. *Genome Biol.* **14**, R32 (2013).

75. Chiappinelli, K. B. *et al.* Inhibiting DNA Methylation Causes an Interferon Response in Cancer via dsRNA Including Endogenous Retroviruses. *Cell* **169**, 361 (2017).

76. Brummelkamp, T. R. *et al.* An shRNA barcode screen provides insight into cancer cell vulnerability to MDM2 inhibitors. *Nat. Chem. Biol.* **2**, 202–206 (2006).

77. Chonghaile, T. N. *et al.* Pretreatment Mitochondrial Priming Correlates with Clinical Response to Cytotoxic Chemotherapy. *Science (80-.).* **334**, 1129–1133 (2011).

78. Vassilev, L. T. *et al.* In Vivo Activation of the p53 Pathway by Small-MoleculeAntagonists of MDM2. *Science (80-.).* **303**, 844–848 (2004).

79. Andreeff, M. *et al.* Results of the Phase I Trial of RG7112, a Small-Molecule MDM2 Antagonist in Leukemia. *Clin. Cancer Res.* **22**, 868–876 (2016).

80. Ding, K. *et al.* Structure-Based Design of Spiro-oxindoles as Potent, Specific Small-Molecule Inhibitors of the MDM2−p53 Interaction. *J. Med. Chem.* **49**, 3432–3435 (2006).

81. Bykov, V. J. N. *et al.* Restoration of the tumor suppressor function to mutant p53 by a low-molecular-weight compound. *Nat. Med.* **8**, 282–288 (2002).

82. Paek, A. L., Liu, J. C., Loewer, A., Forrester, W. C. & Lahav, G. Cell-to-Cell Variation in p53 Dynamics Leads to Fractional Killing. *Cell* **165**, 631–642 (2016).

83. Tovar, C. *et al.* Small-molecule MDM2 antagonists reveal aberrant p53 signaling in cancer: implications for therapy. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 1888–1893 (2006).

84. Zhou, Y. *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **10**, 1523 (2019).

85. Pavesi, G., Mereghetti, P., Mauri, G. & Pesole, G. Weeder Web: discovery of transcription factor binding sites in a set of sequences from co-regulated genes. *Nucleic Acids Res.* **32**, W199–W203 (2004).

86. Makeyev, A. V. & Liebhaber, S. A. The poly(C)-binding proteins: A multiplicity of functions and a search for mechanisms. *RNA* **8**, S1355838202024627 (2002).

87. Zhu, J. & Chen, X. MCG10, a novel p53 target gene that encodes a KH domain RNA-binding protein, is capable of inducing apoptosis and cell cycle arrest in G(2)-M. *Mol. Cell. Biol.* **20**, 5602–18 (2000).

88. Tanner, N. K. & Linder, P. DExD/H Box RNA Helicases. *Mol. Cell* **8**, 251–262 (2001).

89. Minczuk, M. *et al.* TEFM (c17orf42) is necessary for transcription of human mtDNA. *Nucleic Acids Res.* **39**, 4284–4299 (2011).

90. Antonicka, H. & Shoubridge, E. A. Mitochondrial RNA Granules Are Centers for Posttranscriptional RNA Processing and Ribosome Biogenesis. *Cell Rep.* **10**, 920–932 (2015).

91. Zhou, Y. *et al.* The packaging of human immunodeficiency virus type 1 RNA is restricted by overexpression of an RNA helicase DHX30. *Virology* **372**, 97–106 (2008).

92. Ye, P., Liu, S., Zhu, Y., Chen, G. & Gao, G. DEXH-Box protein DHX30 is required for optimal function of the zinc-finger antiviral protein. *Protein Cell* **1**, 956–964 (2010).

93. Zheng, H.-J. *et al.* The Novel Helicase helG (DHX30) Is Expressed During Gastrulation in Mice and Has a Structure Similar to a Human DExH Box Helicase. *Stem Cells Dev.* **24**, 372–383 (2015).

94. Lessel, D. *et al.* De Novo Missense Mutations in DHX30 Impair Global Translation and Cause a Neurodevelopmental Disorder. *Am. J. Hum. Genet.* **101**, 716–724 (2017).

95. Simsek, D. *et al.* The Mammalian Ribo-interactome Reveals Ribosome Functional Diversity and Heterogeneity. *Cell* **169**, 1051-1065.e18 (2017).

96. Wang, Y. *et al.* MYH9 binds to lncRNA gene PTCSC2 and regulates FOXE1 in the 9q22 thyroid cancer risk locus. *Proc. Natl. Acad. Sci.* **114**, 474–479 (2017).

97. Van Nostrand, E. L. *et al.* Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat. Methods* **13**, 508–514 (2016).

98. Davis, C. A. *et al.* The Encyclopedia of DNA elements (ENCODE): data portal update. *Nucleic Acids Res.* **46**, D794–D801 (2018).

99. Weidensdorfer, D. *et al.* Control of c-myc mRNA stability by IGF2BP1-associated cytoplasmic RNPs. *RNA* **15**, 104–115 (2008).

100. Subramanian, A. *et al.* Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci.* **102**, 15545–15550 (2005).

101. Pecci, A., Ma, X., Savoia, A. & Adelstein, R. S. MYH9: Structure, functions and role of non-muscle myosin IIA in human disease. *Gene* **664**, #pagerange# (2018).

102. Liao, Q. *et al.* LIM kinase 1 interacts with myosin-9 and alpha-actinin-4 and promotes colorectal cancer progression. *Br. J. Cancer* **117**, 563–571 (2017).

103. Ye, G. *et al.* MicroRNA-647 Targets SRF-MYH9 Axis to Suppress Invasion and Metastasis of Gastric Cancer. *Theranostics* **7**, 3338–3353 (2017).

104. Anne Conti, M. *et al.* Conditional deletion of nonmuscle myosin II-A in mouse tongue epithelium results in squamous cell carcinoma. *Sci. Rep.* **5**, 14068 (2015).

105. Schramek, D. *et al.* Direct in Vivo RNAi Screen Unveils Myosin IIa as a Tumor Suppressor of Squamous Cell Carcinomas. *Science (80-.).* **343**, 309–313 (2014).

106. Dunkle, A., Dzhagalov, I. & He, Y. W. Mcl-1 promotes survival of thymocytes by inhibition of Bak in a pathway separate from Bcl-2. *Cell Death Differ.* **17**, 994–1002 (2010).

107. Ku, B., Liang, C., Jung, J. U. & Oh, B. H. Evidence that inhibition of BAX activation by BCL-2 involves its tight and preferential interaction with the BH3 domain of BAX. *Cell Res.* **21**, 627–641 (2011).

108. Michels, J. *et al.* Functions of BCL-X L at the Interface between Cell Death and Metabolism. *Int. J. Cell Biol.* **2013**, 1–10 (2013).

109. Kalkavan, H. & Green, D. R. MOMP, cell suicide as a BCL-2 family business. *Cell Death Differ.* **25**, 46–55 (2018).

110. Ren, D. *et al.* BID, BIM, and PUMA Are Essential for Activation of the BAX- and BAK-Dependent Cell Death Program. *Science (80-.).* **330**, 1390–1393 (2010).

111. Gallenne, T. *et al.* Bax activation by the BH3-only protein Puma promotes cell dependence on antiapoptotic Bcl-2 family members. *J. Cell Biol.* **185**, 279–290 (2009).

112. Griffiths, G. J. *et al.* Cell damage-induced conformational changes of the pro-apoptotic protein Bak in vivo precede the onset of apoptosis. *J. Cell Biol.* **144**, 903–914 (1999).

113. Edlich, F. *et al.* Bcl-xL Retrotranslocates Bax from the Mitochondria into the Cytosol. *Cell* **145**, 104–116 (2011).

114. Kepp, O., Rajalingam, K., Kimmig, S. & Rudel, T. Bak and Bax are non-redundant during infection- and DNA damage-induced apoptosis. *EMBO J.* **26**, 825–834 (2007).

115. Haschka, M. D. *et al.* The NOXA-MCL1-BIM axis defines lifespan on extended mitotic arrest. *Nat. Commun.* **6**, (2015).

116. Thompson, T. *et al.* Phosphorylation of p53 on Key Serines Is Dispensable for Transcriptional Activation and Apoptosis. *J. Biol. Chem.* **279**, 53015–53022 (2004).

117. Vaseva, A. V., Marchenko, N. D. & Moll, U. M. The transcription-independent mitochondrial p53 program is a major contributor to nutlin-induced apoptosis in tumor cells. *Cell Cycle* **8**, 1711–1719 (2009).

118. Marine, J. C. & Lozano, G. Mdm2-mediated ubiquitylation: P53 and beyond. *Cell Death Differ.* **17**, 93–102 (2010).

119. Bouska, A. & Eischen, C. M. Murine double minute 2: p53-independent roads lead to genome instability or death. *Trends Biochem. Sci.* **34**, 279–286 (2009).

120. Kracikova, M., Akiri, G., George, A., Sachidanandam, R. & Aaronson, S. A. A threshold mechanism mediates p53 cell fate decision between growth arrest and apoptosis. *Cell Death Differ.* **20**, 576–588 (2013).

121. Choi, H. S. *et al.* Poly(C)-binding proteins as transcriptional regulators of gene expression. *Biochem. Biophys. Res. Commun.* **380**, 431–436 (2009).

122. Makeyev, A. V. & Liebhaber, S. A. Identification of Two Novel Mammalian Genes Establishes a Subfamily of KH-Domain RNA-Binding Proteins. *Genomics* **67**, 301–316 (2000).

123. Lin, X. *et al.* Interplay between PCBP2 and miRNA modulates *ARHGDIA* expression and function in glioma migration and invasion. *Oncotarget* **7**, 19483–19498 (2016).

124. Zhang, X. *et al.* Overexpression of PCBP2 contributes to poor prognosis and enhanced cell growth in human hepatocellular carcinoma. *Oncol. Rep.* **36**, 3456–3464 (2016).

125. Ye, J. *et al.* Poly (C)-binding protein 2 (PCBP2) promotes the progression of esophageal squamous cell carcinoma (ESCC) through regulating cellular proliferation and apoptosis. *Pathol. - Res. Pract.* **212**, 717–725 (2016).

126. Hu, C.-E. E., Liu, Y.-C. C., Zhang, H.-D. D. & Huang, G.-J. J. The RNA-binding protein PCBP2 facilitates gastric carcinoma growth by targeting miR-34a. *Biochem. Biophys. Res. Commun.* **448**, 437–442 (2014).

127. Holcik, M., Gordon, B. W. & Korneluk, R. G. The Internal Ribosome Entry Site-Mediated Translation of Antiapoptotic Protein XIAP Is Modulated by the Heterogeneous Nuclear Ribonucleoproteins C1 and C2. *Mol. Cell. Biol.* **23**, 280–288 (2003).

128. Ren, C., Zhang, J., Yan, W., Zhang, Y. & Chen, X. RNA-binding Protein PCBP2 Regulates p73 Expression and p73-dependent Antioxidant Defense. *J. Biol. Chem.* **291**, 9629–9637 (2016).

129. Marini, A. *et al.* TAp73 contributes to the oxidative stress response by regulating protein synthesis. *Proc. Natl. Acad. Sci.* **115**, 6219–6224 (2018).

130. Bourgeois, C. F., Mortreux, F. & Auboeuf, D. The multiple functions of RNA helicases as drivers and regulators of gene expression. *Nat. Rev. Mol. Cell Biol.* **17**, 426–438 (2016).

131. Wang, Y. & Bogenhagen, D. F. Human Mitochondrial DNA Nucleoids Are Linked to Protein Folding Machinery and Metabolic Enzymes at the Mitochondrial Inner Membrane. *J. Biol. Chem.* **281**, 25791–25802 (2006).

132. Yang, W. *et al.* Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res.* **41**, D955–D961 (2012).

133. Supiot, S., Hill, R. P. & Bristow, R. G. Nutlin-3 radiosensitizes hypoxic prostate cancer cells independent of p53. *Mol. Cancer Ther.* **7**, 993–999 (2008).

134. Deben, C. *et al.* The MDM2-inhibitor Nutlin-3 synergizes with cisplatin to induce p53 dependent tumor cell apoptosis in non-small cell lung cancer. *Oncotarget* **6**, (2015).

135. Gu, L., Zhu, N., Findley, H. W. & Zhou, M. MDM2 antagonist nutlin-3 is a potent inducer of apoptosis in pediatric acute lymphoblastic leukemia cells with wild-type p53 and overexpression of MDM2. *Leukemia* **22**, 730–739 (2008).

136. Conti, M. A. & Adelstein, R. S. Nonmuscle myosin II moves in new directions. *J. Cell Sci.* **121**, 11–8 (2008).

137. Silke, J. & Meier, P. Inhibitor of Apoptosis (IAP) Proteins-Modulators of Cell Death and Inflammation. *Cold Spring Harb. Perspect. Biol.* **5**, a008730–a008730 (2013).

138. Dewson, G. & Kluck, R. M. Mechanisms by which Bak and Bax permeabilise mitochondria during apoptosis. *J. Cell Sci.* **122**, 2801–2808 (2009).

139. Lindsten, T. *et al.* The Combined Functions of Proapoptotic Bcl-2 Family Members Bak and Bax Are Essential for Normal Development of Multiple Tissues. *Mol. Cell* **6**, 1389–1399 (2000).

140. Degenhardt, K., Sundararajan, R., Lindsten, T., Thompson, C. & White, E. Bax and Bak independently promote cytochrome c release from mitochondria. *J. Biol. Chem.* **277**, 14127–14134 (2002).

141. D'Agostino, V. G., Adami, V. & Provenzani, A. A Novel High Throughput Biochemical Assay to Evaluate the HuR Protein-RNA Complex Formation. *PLoS One* **8**, 1–9 (2013).

142. Walker, J. M. *Post-Transcriptional Gene Regulation*. **1358**, (Springer New York, 2016).

143. Wu, T. D. & Nacu, S. Fast and SNP-tolerant detection of complex variants and splicing in short reads. *Bioinformatics* **26**, 873–881 (2010).

144. Anders, S., Pyl, P. T. & Huber, W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).

145. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).

146. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

147. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).

148. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).

149. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).

150. Tan, G. & Lenhard, B. TFBSTools: an R/bioconductor package for transcription factor binding site analysis. *Bioinformatics* **32**, 1555–1556 (2016).

151. Dassi, E. *et al.* Hyper conserved elements in vertebrate mRNA 3′-UTRs reveal a translational network of RNA-binding proteins controlled by HuR. *Nucleic Acids Res.* **41**, 3201–3216 (2013).

152. Dyballa, N. & Metzger, S. Fast and Sensitive Colloidal Coomassie G-250 Staining for Proteins in Polyacrylamide Gels. *J. Vis. Exp.* (2009). doi:10.3791/1431

153. Pfaffl, M. W. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* **29**, 45e – 45 (2001).

154. Barczak, W., Suchorska, W., Rubiś, B. & Kulcenty, K. Universal Real-Time PCR-Based Assay for Lentiviral Titration. *Mol. Biotechnol.* **57**, 195–200 (2015).

155. Rossi, M. *et al.* LncRNA EPR controls epithelial proliferation by coordinating Cdkn1a transcription and mRNA decay response to TGF-β. *Nat. Commun.* **10**, 1969 (2019).

156. Briata, P. *et al.* Diverse roles of the nucleic acid-binding protein KHSRP in cell differentiation and disease. *Wiley Interdiscip. Rev. RNA* **7**, 227–240 (2016).

157. Puppo, M. *et al.* miRNA-Mediated KHSRP Silencing Rewires Distinct Post-transcriptional Programs during TGF-β-Induced Epithelial-to-Mesenchymal Transition. *Cell Rep.* **16**, 967–978 (2016).

11) ATTACHMENTS

11.1 – RNA sequencing data

The tables containing the analysis of the DEGs coming from both RNA-sequencing experiments can be downloaded from the following link:
https://drive.google.com/drive/folders/1-6qdNq2V62O5Qcs5jo59idMS0PKJFt5P?usp=sharing

11.2 – List of primers used for qPCR

Primer name	Sequence
BAK1 Fw	GACGACATCAACCGACGCTA
BAK1 Rv	AGCATGAAGTCGACCACGAA
BCL2L1 Fw	CGGCTGGGATACTTTTGTGG
BCL2L1 Rv	TGGGAGGGTAGAGTGGATGG
CASP3 Fw	AGCGAATCAATGGACTCTGGAA
CASP3 Rv	TGAGGTTTGCTGCATCGACA
CDH1 Fw	ACACCCGGGACAACGTTTA
CDH1 Rv	TGTGCAGCTGGCTCAAGT
CDKN1A Fw	CTGGAGACTCTCAGGGTCGAAA
CDKN1A Rv	GATTAGGGCTTCCTCTTGGAGAA
COX1 Fw	CTATACCTATTATTCGGCGCATGA
COX1 Rv	CAGCTCGGCTCGAATAAGGA
EIF5A Fw	GGGGGACTGATTCCAAGACA
EIF5A Rv	TACCAACCAGATGGACCTTGG
GADD45B Fw	CCTGCAAATCCACTTCACGC
GADD45B Rv	GTGTGAGGGTTCGTGACCAG

Primer name	Sequence
GAPDH Fw	TCCAAAATCAAGTGGGGCGA
GAPDH Rv	AGTAGAGGCAGGGATGATGT
LUC2 Fw	TTGATCAGGCTCTTCAGCCG
LUC2 Rv	GGCTACGTTAACAACCCCGA
MARK4 Fw	CAACGATCGGAACTCGGACAC
MARK4 RV	CAATGGTCCTCAGCAGGCG
RPS20 Fw	CCAGTTCGAATGCCTACCAAGACTT
RPS20 Rv	ACTTCCACCTCAACTCCTGGCTCA
TWIST1 Fw	TTCAAAGAAACAGGGCGTGG
TWIST1 Rv	GCACGACCTCTTGAGAATGC
VIM Rv	GCTTCCTGTAGGTGGCAATC
VIMFw	GAGAACTTTGCCGTTGAAGC
YWHAZ Fw	CAACACATCCTATCAGACTGGG
YWHAZ Rv	AATGTATCAAGTTCAGCAATGGC
ZEB1 Fw	GAAAATGAGCAAAACCATGATCCTA
ZEB1 Rv	CAGGTGCCTCAGGAAAAATGA

11.3 – List of primers used for DHX30 cloning

Primer name	Sequence
pCMV DHX30 Fw	GAGGCGATCGCATGTTCAGCCTGGACTCA
pCMV DHX30 Rv	GCGACGCGTGTCGTCAGCTGTCTTG
pCW57.1 DHX30 Fw	TTTCTAGACCACCATGTTCAGCCTGGACTCA
pCW57.1 DHX30 Rv	TAACCGGTTCAGTCGTCAGCTGTCTTGCG

11.4 – List of shRNA vectors

Gene	Vector ID	Sequence
DHX30	TRCN0000052028	GCACACAAATGGACCGAAGAA
DHX30	TRCN0000052031	CCGATGGCTGACGTATTTCAT
DHX30	TRCN0000052032	GAGTTGTTTGACGCAGCCAAA
PCBP2	TRCN0000074685	GCCATCACTATTGCTGGCATT
PCBP2	TRCN0000074687	CCTGGCTCAATATCTAATCAA
MYH9	TRCN0000029465	CGCATCAACTTTGATGTCAAT
MYH9	TRCN0000029468	GCCAAGCTCAAGAACAAGCAT
None – Scramble control	SHC206	

11.5 – List of antibodies used for western blot

Antibody	Clone	Brand	Specie	Dilution
a-Actinin	H-2	Santa Cruz	Mouse	1:6000
BAK	G23	Santa Cruz	Rabbit	1:1000
b-Tubulin	3F3-G2	Santa Cruz	Mouse	1:10000
Caspase 3	H-277	Santa Cruz	Rabbit	1:1000
cCASP3	9661S	Cell Signaling	Rabbit	1:500
DHX30	ab85687	abcam	Rabbit	1:1000
DHX30	A302-218A	Bethyl	Rabbit	1:1000
EIF5A	H8	Santa Cruz	Mouse	1:1000
Myc	9-E10	Santa Cruz	Mouse	1:1000
MYH9	A304-490a	Bethyl	Rabbit	1:2000
p53	DO-1	Santa Cruz	Mouse	1:4000
PABPC1	ab21060	abcam	Rabbit	1:1000
PARP	9542T	Cell Signaling	Rabbit	1:1000
PCBP1	ab154252	abcam	Rabbit	1:1000
PCBP2	23G	Santa Cruz	Mouse	1:1000
PCBP4	ab59534	abcam	Rabbit	1:500

11.6 – RNA probes sequences

Probe name	Sequence
WT	5' - GAAGGGCCCUCCCCAUGGCCCUGGAGAGUGGG - 3'
Mut	5' - GAAGGGCCCUCCCCAUGGAGAUGGAGAGUGGG - 3'
AU-rich	5' - AUGUAUUGUUUAUACAU - 3'
Negative	5' - AAAAAAAGAAAAAAAGAAAAAA - 3'

11.7 – Paper

LncRNA EPR controls epithelial proliferation by coordinating Cdkn1a transcription and mRNA decay response to TGF-β

https://doi.org/10.1038/s41467-019-09754-1

nature COMMUNICATIONS

ARTICLE

https://doi.org/10.1038/s41467-019-09754-1 OPEN

LncRNA EPR controls epithelial proliferation by coordinating *Cdkn1a* transcription and mRNA decay response to TGF-β

Martina Rossi[1,2], Gabriele Bucci[3], Dario Rizzotto[4], Domenico Bordo[1], Matteo J. Marzi[5], Margherita Puppo[1,2], Arielle Flinois[6], Domenica Spadaro[6], Sandra Citi[6], Laura Emionite[7], Michele Cilli[7], Francesco Nicassio[5], Alberto Inga[4], Paola Briata[1] & Roberto Gherzi[1]

Long noncoding RNAs (lncRNAs) are emerging as regulators of fundamental biological processes. Here we report on the characterization of an intergenic lncRNA expressed in epithelial tissues which we termed EPR (Epithelial cell Program Regulator). EPR is rapidly downregulated by TGF-β and its sustained expression largely reshapes the transcriptome, favors the acquisition of epithelial traits, and reduces cell proliferation in cultured mammary gland cells as well as in an animal model of orthotopic transplantation. EPR generates a small peptide that localizes at epithelial cell junctions but the RNA molecule per se accounts for the vast majority of EPR-induced gene expression changes. Mechanistically, EPR interacts with chromatin and regulates *Cdkn1a* gene expression by affecting both its transcription and mRNA decay through its association with SMAD3 and the mRNA decay-promoting factor KHSRP, respectively. We propose that EPR enables epithelial cells to control proliferation by modulating waves of gene expression in response to TGF-β.

[1] Gene Expression Regulation Laboratory, IRCCS Ospedale Policlinico San Martino, 16132 Genova, Italy. [2] DIMES Sezione Biochimica-Università di Genova, 16132 Genova, Italy. [3] Center of Translational Genomics and Bioinformatics, IRCCS Ospedale San Raffaele, 20132 Milano, Italy. [4] Laboratory of Transcriptional Networks, Center for Integrative Biology, CIBIO, University of Trento, 38123 Trento, Italy. [5] Center for Genomic Science of IIT@SEMM, Istituto Italiano di Tecnologia (IIT), 20139 Milano, Italy. [6] Department of Cell Biology, University of Geneve, 1211 Geneve, Switzerland. [7] Animal Facility, IRCCS Policlinico San Martino, 16132 Genova, Italy. These authors jointly supervised this work: Paola Briata, Roberto Gherzi. Correspondence and requests for materials should be addressed to A.I. (email: alberto.inga@unitn.it) or to P.B. (email: paola.briata@hsanmartino.it) or to R.G. (email: rgherzi@ucsd.edu)

Human transcriptome analysis has revealed the existence of a surprisingly high number of noncoding RNAs that have been classified in multiple families based on their size and biogenesis. Long noncoding RNAs (lncRNAs) are defined as transcripts longer than 200 nucleotides transcribed by RNA polymerase II and commonly originated from intergenic regions. LncRNAs can be capped, spliced, and polyadenylated and usually show limited protein coding potential (refs. [1,2], and literature cited therein).

LncRNAs are emerging as a fundamental aspect of biology due to their ability to reprogram gene expression and influence distinct cellular functions including cell fate determination, cell cycle progression, apoptosis, and aging[1,2]. Their expression is usually tissue restricted, developmentally regulated, and can change under specific pathological conditions. Many lncRNAs influence hallmarks of cancer such as uncontrolled proliferation, evasion of cell death, as well as metastasis formation and it has been suggested that lncRNAs can act as oncogenes or tumor suppressors —either directly or indirectly— by interfering with different pathways[3,4]. From a mechanistic point of view, lncRNAs may influence the function of transcriptional complexes, modulate chromatin structures, serve as scaffolds to form ribonucleoprotein (RNP) complexes or as decoys for proteins and micro-RNAs (miRNAs)[2,5]. Thus, lncRNA-mediated control of gene expression may take place at transcriptional and/or posttranscriptional levels[5–9].

Recently, lncRNAs have been described as important components of the transforming growth factor β (TGF-β) signaling pathway[10,11]. TGF-β belongs to a large family of structurally related cytokines that regulate growth, survival, differentiation, and migration of many cell types including mammary gland epithelial cells (ref. [12]. for a review). TGF-β activates membrane kinase receptors and induces phosphorylation of cell-specific SMAD proteins that, in complex with the common SMAD4, accumulate into the nucleus to regulate gene expression at different levels (ref. [13]. for a recent review).

In our previous studies, we showed that the multifunctional RNA-binding protein KHSRP acts as a regulatory hub that conveys extracellular stimuli into gene expression changes due to its ability to interact with several molecular partners[14]. KHSRP is able to posttranscriptionally regulate gene expression by promoting decay of unstable mRNAs, favoring maturation of select miRNAs from precursors, and controlling alternative splicing events[14]. Recently, we reported that KHSRP affects the alternative splicing of a cohort of pre-mRNAs that encode regulators of cell adhesion and motility—such as CD44 and FGFR2—favoring their epithelial type exon usage and that miRNA-mediated KHSRP silencing is required for TGF-β-induced epithelial-to-mesenchymal transition (EMT) in immortalized NMuMG mammary gland cells[15]. Further, we found that Resveratrol—a natural polyphenolic compound endowed with anti-inflammatory, antiproliferative, as well as proapoptotic activities—prevents TGF-β-dependent KHSRP downregulation. thus shifting *Cd44* and *Fgfr2* pre-mRNA alternative splicing from the mesenchymal-specific to the epithelial-specific isoforms[16]. Our previous observation that the lncRNA H19 interacts with KHSRP and affects its mRNA decay-promoting function[17] prompted us to identify additional KHSRP/lncRNAs interactions endowed with regulatory potential.

Here we describe a previously uncharacterized mammalian lncRNA expressed in epithelial tissues that we termed EPR (after Epithelial Program Regulator). EPR came to our attention due to its ability to interact with KHSRP and to counteract TGF-β-induced EMT. EPR contains an open reading frame (ORF) that is translated into a small peptide localized at epithelial cell junctions. However, we found that EPR regulates the expression of a large set of target transcripts independently of the peptide biogenesis. Our studies have revealed that EPR interacts with chromatin, regulates *Cdkn1a* gene expression by affecting both its transcription and mRNA decay, and controls cell proliferation in both immortalized and transformed mammary gland cells as well as in a mouse model of orthotopic transplantation.

Results

Identification of EPR, an epithelial cell-enriched lncRNA. This study was initiated in an attempt to identify lncRNAs which are able to interact with KHSRP and whose expression is regulated by TGF-β in immortalized murine mammary gland NMuMG cells. To this end, we leveraged RNA-sequencing (RNA-Seq) and anti-KHSRP RNP complexes Immunoprecipitation followed by RNA-sequencing (RIP-Seq) analyses performed in untreated or TGF-β-treated NMuMG cells. TGF-β treatment significantly reduced or increased the levels of 110 and 194 lncRNAs, respectively (|log2 fold changes| > 2.0, $p < 0.01$ (Student's t test); Supplementary Table 1a) while RIP-Seq analysis showed that TGF-β modulates the interaction of KHSRP with 67 lncRNAs (|log2 fold changes| > 2.0, $p < 0.01$ (Student's t test); Supplementary Table 1b). Among a set of lncRNA candidates of potential interest in EMT, we focused on the previously uncharacterized BC030870 (ENSMUSG00000074300, located on mouse chromosome 8 and transcribed in reverse orientation) that we renamed EPR (highlighted in yellow in Supplementary Table 1a and 1b). RIP analysis followed by quantitative RT- PCR (qRT-PCR) as well as band-shift analysis confirmed that EPR directly interacts with KHSRP (Supplementary Fig. 1a, b). TGF-β induced a small increase in EPR levels followed by rapid downregulation (Fig. 1a) that accounts for the reduced interaction between KHSRP and EPR upon a 6-h treatment (Supplementary Table 1b). TGF-β-dependent modulation of EPR expression requires TGF-β type I receptor signaling as shown by the ability of SB431542 (a selective inhibitor of ALK5, 4, and 7 [18]) to abrogate the effect of the cytokine on EPR expression (Supplementary Fig. 1c). SMAD complexes are major effectors of TGF-β-dependent transcriptional regulation[13] and our ChIP-qPCR showed that SMAD3 interacts with EPR promoter in a TGF-β-modulated way (Supplementary Fig. 1d, upper panel). Positive (*Serpine1*) and negative (*Mettl9*) controls for ChIP experiments are provided in Supplementary Fig. 1d (lower panel) and Supplementary Fig. 1e, respectively. Our data are consistent with the hypothesis that SMAD3 interacts with a corepressor complex on EPR promoter region to modulate its transcription[19]. De novo protein synthesis is not required for TGF-β-induced downregulation of EPR expression as revealed by the use of cycloheximide (Supplementary Fig. 1f, upper panel; *Zeb2* (also known as SIP1) represents the control for cycloheximide activity[20]).

EPR is expressed during embryonic development (Supplementary Fig. 1g) and in epithelial tissues of adult mice with a prevalence in the gastrointestinal tract, lung, kidney and mammary gland (Fig. 1b). EPR is polyadenylated and spliced (Supplementary Fig. 1h) and it is almost equally distributed in the cytoplasm, nucleoplasm, and chromatin of NMuMG cells (Fig. 1c; see Supplementary Fig. 2a for an immunoblot-based validation of cell fractionation). LINC01207 (a.k.a. SMIM31, located on chromosome 4 and transcribed in forward orientation; hereafter indicated as h.EPR) is the human ortholog of EPR and displays superimposable epithelial tissue-enriched expression (as evaluated through the Human BodyMap 2.0 data from Illumina; Supplementary Fig. 2b). Bioinformatics analysis performed on RNA-Seq data derived from different subpopulations of normal breast cells isolated by FACS analysis from reduction mammoplasty specimens[21] revealed that h.EPR is expressed exclusively in differentiated luminal cells of the mammary gland (Fig. 1d).

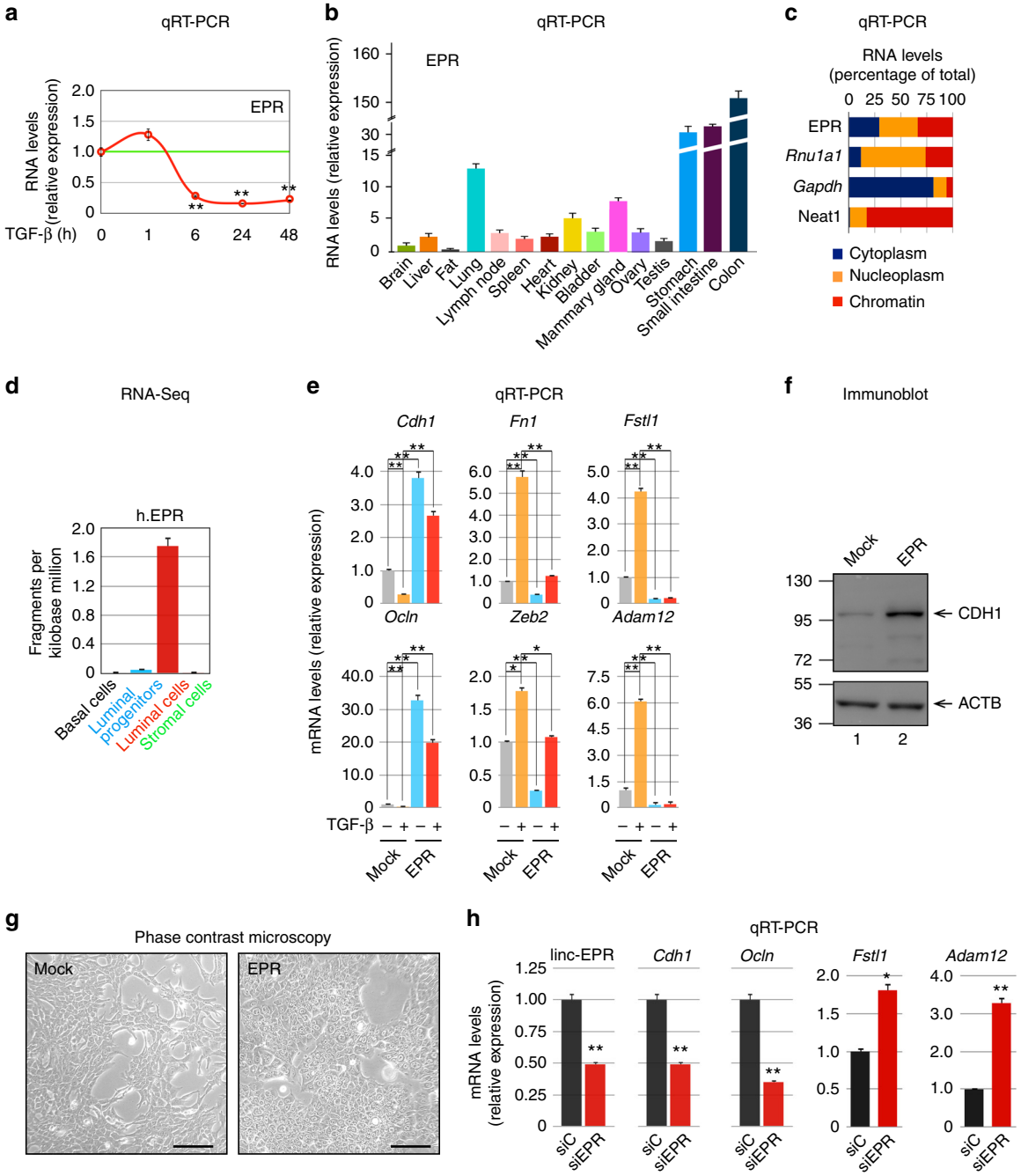


Fig. 1 EPR displays epithelial expression and antagonizes TGF-β-induced EMT in mammary gland cells. **a** Quantitative RT-PCR (qRT-PCR) analysis of EPR in NMuMG cells serum-starved (2% FBS, 16 h) and either treated with TGF-β (10 ng ml^{-1}) for the indicated times or untreated (time 0). **b** qRT-PCR analysis of EPR in the indicated mouse tissues. **c** NMuMG cells were fractionated and RNA was prepared from cytoplasm, nucleoplasm, and chromatin and analyzed by qRT-PCR to quantify the indicated RNAs. *Rnu1a1* is also known as U1 small nuclear RNA, *Gapdh* mRNA encodes the glyceraldehyde-3-phosphate dehydrogenase. **d** qRT-PCR analysis of h.EPR in normal human breast cells isolated from reduction mammoplasty specimens[21]. **e** qRT-PCR analysis of the indicated transcripts in either mock or EPR-overexpressing (EPR) NMuMG cells serum-starved and either treated with TGF-β (+) for 24 h or untreated (−). **f** Immunoblot analysis of total cell extracts from either mock or EPR-overexpressing (EPR) NMuMG cells. The indicated antibodies were used. The position of molecular mass markers is indicated on the left. Representative gels are shown. ACTB is also known as Actin Beta. **g** Phase contrast microscopy of either mock or EPR-overexpressing (EPR) NMuMG cells. Scale bars: 100 µm. **h** qRT-PCR analysis of the indicated transcripts in NMuMG cells transiently transfected with either control siRNA (siC) or siRNA designed to silence EPR expression (siEPR). The values of qRT-PCR experiments shown are averages (±SEM) of three independent experiments performed in triplicate. Statistical significance: $*p < 0.01$, $**p < 0.001$ (Student's *t* test)

In order to investigate the potential role of EPR in TGF-β-induced EMT, we decided to counteract TGF-β-dependent EPR downregulation by stably overexpressing the lncRNA in NMuMG cells (overexpression was 3- to 12-fold compared to the respective mock cells (empty vector-transfected), in different transfectant pools). EPR overexpression prevented TGF-β-induced down-regulation of epithelial factors (*Cdh1*, *Ocln*) and induction of mesenchymal markers (*Fn1*, *Fstl1*, *Zeb2*, *Adam12*) as well morphological changes (Fig. 1e, Supplementary Fig. 2c). Strikingly, we observed that EPR overexpression affects the levels of

epithelial and mesenchymal markers (Fig. 1f, Supplementary Fig. 2d), and induced a cobblestone-like cell morphology in untreated cells (Fig. 1g). Further, EPR overexpression significantly limited the migratory potential of NMuMG cells (Supplementary Fig. 2e). Conversely, transient silencing of EPR downregulated the mRNA levels of epithelial markers, enhanced the levels of mesenchymal markers (Fig. 1h), and rescued the gene expression changes induced by stable overexpression of the lncRNA (Supplementary Fig. 2f). Interestingly, bioinformatics analysis of RNA-Seq data derived from human normal breast samples revealed a statistically significant positive correlation between the expression of h.EPR and epithelial markers such as CDH1 and OCLN and a negative correlation with mesenchymal markers such as VIM and SNAI1 (Supplementary Fig. 2g). This observation is in agreement with the evidence that EPR expression is mutually exclusive with the expression of the EMT factor *Cdh2* as revealed by bioinformatics analysis of datasets derived from single-cell RNA-Seq analysis performed in mice (Supplementary Fig. 2h).

In conclusion, the name EPR that we assigned to lncRNA BC030870 (after Epithelial Program Regulator) is consistent with its enriched expression in epithelial cells and with the upregulation of epithelial markers and downregulation of mesenchymal markers induced by its overexpression.

EPR encodes a small polypeptide. A few recent reports show that certain lncRNAs contain short ORFs that can be translated into peptides endowed with regulatory functions[22–25]. The analysis of EPR sequence revealed the presence of a 213 nucleotide-long ORF potentially encoding a 71-amino acid polypeptide that, interestingly, corresponds to the lncRNA region that displays the highest identity with the human ortholog (Fig. 2a). The putative polypeptide sequence is well conserved among mammalian species and in silico methods identified a conserved α-helical transmembrane domain while a predicted second α-helix was found in the putative cytosolic domain (Fig. 2b). Importantly, polysome fractionation followed by qRT-PCR analysis revealed that EPR localizes to actively translating polysomes (Supplementary Fig. 3a).

To investigate whether EPR ORF is translated, we inserted a FLAG tag at its 3′ end and transiently transfected the resulting construct into HEK-293 cells (Fig. 2c, left). As shown in Fig. 2c (right), the ORF was translated into a short polypeptide of the expected molecular mass. To unambiguously prove the existence of the endogenous small EPR-encoded peptide (EPRp), the ORF was expressed in bacteria and the resulting peptide was purified and utilized as immunogen to generate a rabbit polyclonal antibody. Polyclonal anti-EPRp recognized a recombinant polypeptide transiently expressed in HEK-293 cells (Supplementary Fig. 3b) and, most importantly, a ~8 KDa polypeptide in mouse gastrointestinal tract organs and breast (Supplementary Fig. 3c). In keeping with EPR downregulation upon TGF-β treatment, the expression of the EPRp was downregulated in response to treatment with TGF-β for 24 h (Fig. 2d).

In order to identify the molecular partners of EPRp, we performed immunoaffinity purification of proteins interacting with EPRp in NMuMG cells. Mass spectrometry (MS) analysis of coimmunoprecipitating proteins separated by SDS-PAGE (Fig. 2e) revealed an enrichment in junctional and cytoskeletal proteins (Supplementary Data 1). Coimmunoprecipitation experiments confirmed that EPRp interacts with the tight junction proteins TJP1 (ZO-1) and CGN (Cingulin), with the tight and adherens junction protein CGNL1 (Paracingulin) as well as with the actin-associated proteins CTTN (Cortactin) and MYH9 (epithelial myosin-II) (Fig. 2f, Supplementary Fig. 3d).

To investigate EPRp subcellular localization, we performed immunofluorescence experiments in NMuMG cells stably transfected with either EPRp-FLAG or with a construct in which the second codon of the ORF—encoding glutamic acid, E, of EPRp—was mutagenized in order to obtain a STOP codon (see also below, EPRSTOPE-FLAG). Specific localization of FLAG signal at cell−cell junctions, labeled by the junctional marker CGN, was detected in cells stably expressing EPRp-FLAG (arrows in Fig. 2g) while no junctional FLAG labeling was detected in mock-transfected cells or in cells expressing the point-mutant version unable to produce the peptide. CGN labeling was wavy and discontinuous in mock-transfected cells and in cells expressing EPRSTOPE-FLAG, whereas it was linear and uninterrupted in cells expressing EPRp-FLAG, suggesting that EPRp overexpression promotes epithelial junction assembly and reorganization of the junction-associated actin cytoskeleton. A weak diffuse cytoplasmic staining observed in NMuMG cells expressing EPRp-FLAG might reflect EPRp interaction with cytoskeletal proteins (Fig. 2g).

On the basis of these results, we conclude that an ORF present in EPR is translated into a small peptide that is well conserved among species and that displays a junctional localization in mammary gland cells.

EPR regulates gene expression in NMuMG cells. We set out to investigate the function(s) of EPR in NMuMG cells. First, in order to answer the question whether the phenotypic changes that we observed by overexpressing EPR were caused by the lncRNA per se, the peptide or both, we performed transcriptome-wide RNA-Seq analyses in mock cells as well as in NMuMG cells overexpressing either EPR or a point-mutant version unable to produce the peptide (EPRSTOPE, for details see above and Fig. 3a). Bioinformatics analyses of RNA-Seq data revealed a vast rearrangement of the transcriptome as a consequence of both EPR and EPRSTOPE overexpression (Supplementary Data 2). Gene ontology (GO) analysis of RNA-Seq results revealed the enrichment of terms related to epithelial morphogenesis, cell motility, cell migration, and epithelial cell proliferation among the top regulated categories. Representative examples of transcripts either upregulated or downregulated by both EPR and EPRSTOPE are shown in Fig. 3b, c. In keeping with the sequence conservation between EPR and h.EPR, the overexpression of the human lncRNA in murine NMuMG cells yielded gene expression changes superimposable to those obtained by overexpressing the murine lncRNA (Supplementary Fig. 4a).

Interestingly, overexpression of either EPR or EPRSTOPE caused largely overlapping gene expression changes when compared to mock cells (Fig. 3d, upper panel). When we directly compared gene expression changes induced by EPR or EPRSTOPE by applying stringent statistical criteria, we noticed that only a relatively small group of genes displayed expression changes dependent on the presence of EPRp (Fig. 3d, lower panel). The analysis of three independent NMuMG transfectant pools overexpressing EPRSTOPE, followed by qRT-PCR-based validation, allowed us to further restrict the number of transcripts whose levels are affected by the peptide per se (Fig. 3e, Supplementary Fig. 4b). These include transcripts encoding a calcium-dependent cell adhesion protein (*Pcdh19*), two ion transporters (*Slc9a2*, *Scl39a4*), a cytokine receptor (*Fgfr2*) as well as a modulator of membrane transport and actin dynamics (*Anxa6*). Further, analysis of an additional EPR mutant (referred to as EPRSTOPM in which the start codon has been mutagenized to a STOP codon, see below) confirmed the restricted number of gene expression changes that can be ascribed to the peptide translation (Supplementary Fig. 4c).

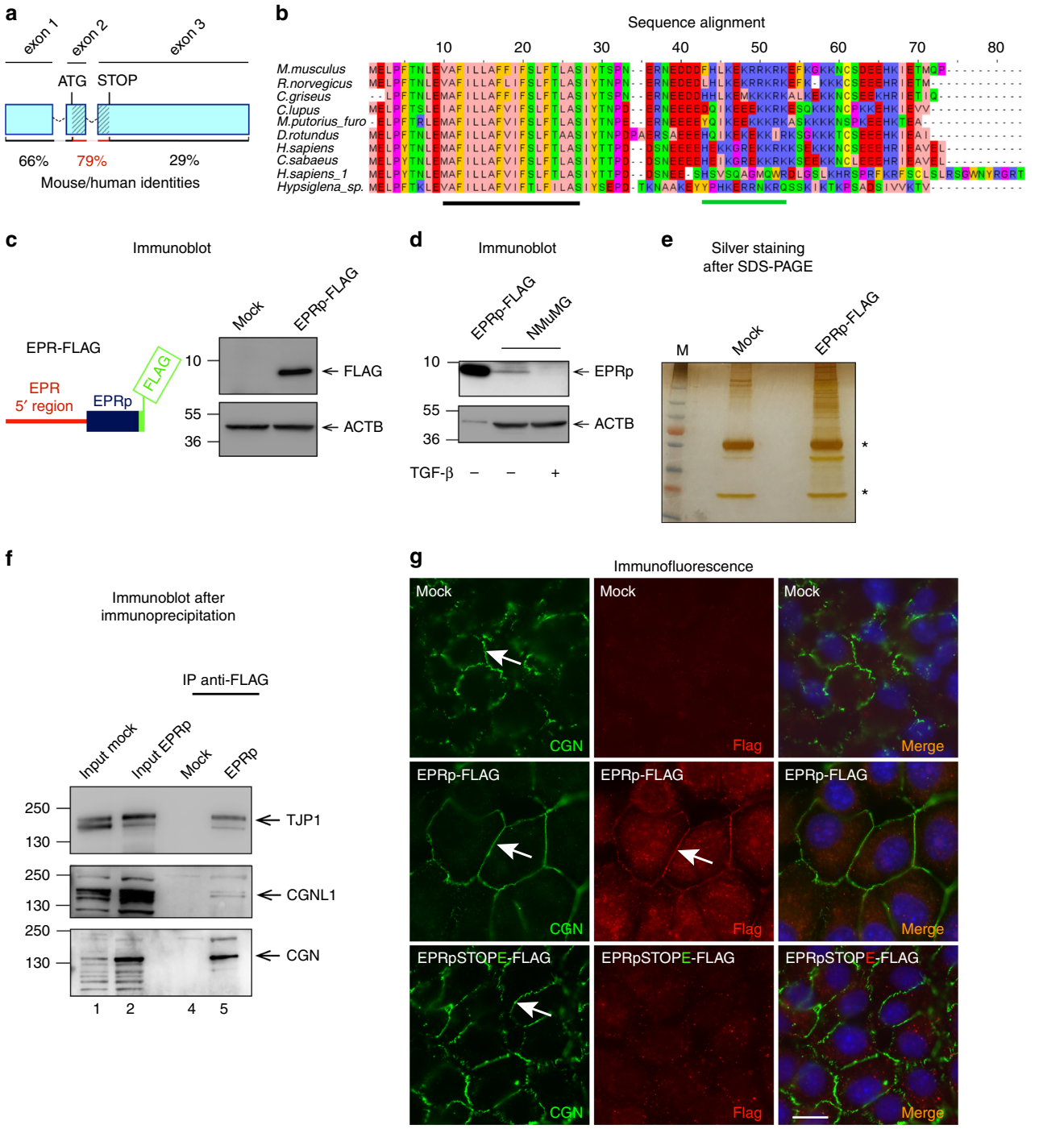


Fig. 2 A small peptide (EPRp) originates from EPR and displays junctional localization. **a** Schematic of the exon−intron structure (not in scale) of EPR, a putative ORF is dashed. Percentage of human/mouse identity in the putative ORF and in its flanking RNA regions is presented. **b** Alignment of the predicted mammalian amino acids sequence encoded by the putative ORF present in EPR and in its orthologs. The position of the predicted transmembrane α-helix is shown as a solid black bar while the position of an additional cytoplasmic α-helix is shown as a solid green bar. **c** Left, diagram of the FLAG-fusion construct used for transfection (EPRp-FLAG); right, immunoblot analysis of total cell extracts from either mock or EPRp-FLAG-transfected HEK-293 cells. **d** Immunoblot analysis of total cell extracts from NMuMG cells serum-starved and either treated with TGF-β (+) for 24 h or untreated (−); extracts from EPRp-FLAG-transfected HEK-293 cells (EPRp-FLAG) represent a positive control. **e** SDS-PAGE analysis of total cell extracts from either mock or EPRp-FLAG-overexpressing (EPRp-FLAG) NMuMG cells immuno-purified using anti-FLAG monoclonal antibody. A representative silver-stained gel is shown. Asterisks indicate the position of immunoglobulin heavy and light chains. **f** Coimmunoprecipitation of FLAG-tagged EPRp and distinct junctional proteins (as indicated) in total extracts from NMuMG cells stably transfected with EPRp-FLAG. **g** Immunofluorescence analysis of either mock or EPRp-FLAG- or EPRpSTOPE-FLAG-stably transfected NMuMG cells cultured to confluence, to allow formation of cell−cell junctions. Arrows point to CGN and EPRp-FLAG junctional localization. Scale bar: 10 μm. For immunoblots, the indicated antibodies were used; the position of molecular mass markers is presented on the left and representative gels are shown. ACTB is also known as Actin Beta

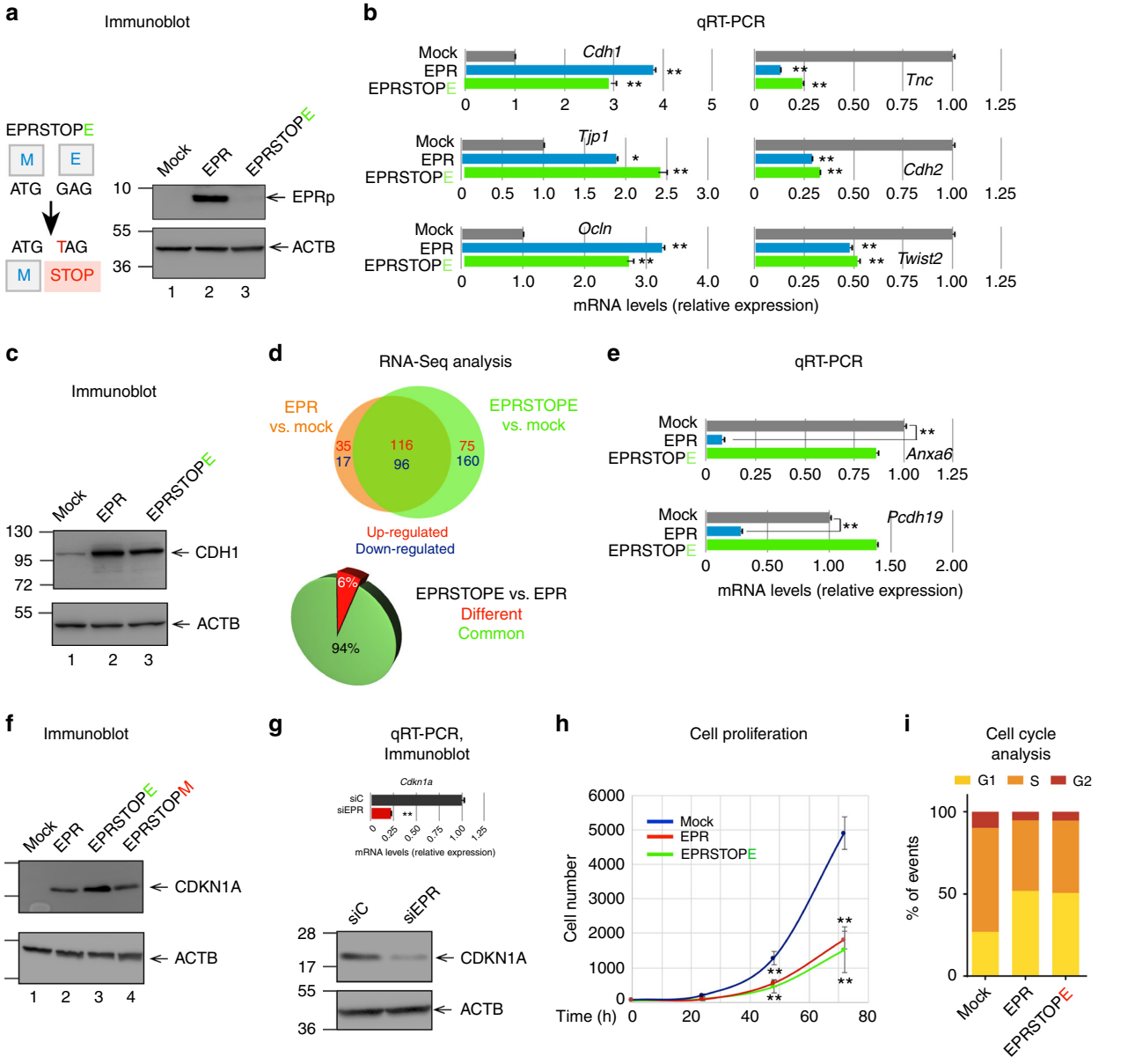


Fig. 3 EPR overexpression reshapes NMuMG cells transcriptome and reduces cell proliferation. **a** Schematic of the point mutation introduced in the second codon of the EPR ORF to generate EPRSTOPE (left panel). HEK-293 cells were transiently transfected with either EPR or EPRSTOPE and the translation of EPR was assessed by immunoblotting (right panel). **b** qRT-PCR analysis of the indicated transcripts in either mock, EPR- or EPRSTOPE-expressing NMuMG cells. **c** Immunoblot analysis of total cell extracts from either mock, EPR- or EPRSTOPE-overexpressing NMuMG cells. **d** Upper panel, TopTable Venn diagram of the limma one-factor contrast analysis showing the overlap between transcripts either upregulated (numbers in red) or downregulated (numbers in blue) in cells stably overexpressing either EPR or EPRSTOPE compared with mock cells. Only transcripts displaying |log2 fold expression difference| > 2.5 ($p < 0.0001$, Student's t test) were included in further comparisons. Lower panel, Pie diagram showing the percentage of gene expression changes common to NMuMG cells overexpressing EPR and EPRSTOPE as well as the percentage of changes that can be specifically attributed to the EPR translation. **e** qRT-PCR analysis of the indicated transcripts in either mock, EPR- or EPRSTOPE-overexpressing NMuMG cells. **f** Immunoblot analysis of total cell extracts from either mock, EPR-, EPRSTOPE- or EPRSTOPM-overexpressing NMuMG cells. **g** qRT-PCR analysis of *Cdkn1a* mRNA levels in either control siRNA (siC) or siEPR-transfected NMuMG cells (upper panel). Immunoblot analysis of total cell extracts from either control siRNA (siC) or siEPR-transfected NMuMG cells (lower panel). **h** Proliferation analysis (Operetta CLS High-Content Analysis System) of either mock, EPR- or EPRSTOPE-overexpressing NMuMG cells. **i** Cell cycle distribution in cultures grown at similar density (near 90% confluence). Cell cycle was analyzed by flow cytometry after double staining with EdU and propidium iodide of either mock, EPR- or EPRSTOPE-overexpressing NMuMG cells. Bars plot the relative proportion of cells in G1, S, and G2 phases for each cell line. The values of qRT-PCR experiments shown are averages (±SEM) of three independent experiments performed in triplicate. Statistical significance: $*p < 0.01$, $**p < 0.001$ (Student's t test). For immunoblots, the indicated antibodies were used; the position of molecular mass markers is indicated on the left and representative gels are shown

Considering the emerging evidence that some lncRNAs act locally (in *cis*) to regulate the expression of nearby genes, we investigated this possibility and RNA-Seq analysis revealed that the expression levels of genes proximal to EPR (*Palld*, *Cpe*, *Sc4mol*, *Klhl2*, *Tmem192*, *Tma16*, *Naf16*, *Nat2* and *Pssd*, localized over 8 MB of chromosome 8) are unaffected by the almost complete EPR downregulation that occurs in NMuMG cells treated with TGF-β for 24 h (R.G. and G.B., unpublished observation).

Altogether, transcriptome-wide analyses showed a EPR-dependent wide rearrangement of the transcriptome in NMuMG cells with relatively restricted effects on gene expression ascribed to the peptide. Thus, we decided to focus our further studies on the EPR functions that are independent of the peptide biogenesis.

EPR regulates *Cdkn1a* gene expression and cell proliferation. Among the GO terms significantly enriched by the overexpression of either EPR or EPRSTOPE, we identified the category Regulation of Epithelial Cell Proliferation. Indeed, both EPR and EPRSTOPE overexpression significantly affected the levels of a group of transcripts belonging to this category including the cyclin-dependent kinase inhibitor *Cdkn1a* (a.k.a. $p21^{WAF1/Cip1}$) (Supplementary Fig. 4d). Immunoblots presented in Fig. 3f show that overexpression of either EPR or EPRSTOPE or EPRSTOPM strongly enhanced CDKN1A levels. Conversely, EPR silencing strongly reduced CDKN1A expression (Fig. 3g). As expected, CDKN1A levels were enhanced by overexpression of the human ortholog of EPR (Supplementary Fig. 4e). Most important, we found that overexpression of either EPR or EPRSTOPE as well as of h.EPR strongly reduces cell proliferation rate in NMuMG cells (Fig. 3h, Supplementary Fig. 4f). Cell cycle analysis demonstrated a relevant increment of cells arrested in the G1 phase in the case of NMuMG cells transfected with either EPR or EPRSTOPE in comparison to mock cells (Fig. 3i). To exclude the possibility that gene expression changes that we observed (Fig. 3b) might be dependent on the EPR-induced G1 arrest, we sorted cells in the G1 phase and analyzed gene expression changes by qRT-PCR. Data presented in Supplementary Fig. 4g indicate that the expression changes induced in G1-enriched cells by overexpression of either EPR or EPRSTOPE are superimposable to those observed in the total cell population (Fig. 3b).

Together, these results provide evidence that modulation of EPR levels regulates *Cdkn1a* gene expression and affects cell proliferation in NMuMG cells. Given the role of CDKN1A in promoting cell cycle arrest in response to many stimuli—including TGF-β[26]—we decided to focus our further mechanistic studies on the role of EPR in TGF-β-dependent regulation of *Cdkn1a* gene expression.

EPR regulates TGF-β-dependent *Cdkn1a* gene expression. Analysis of newly synthesized transcripts revealed that overexpression of either EPR or EPRSTOPE strongly enhances *Cdkn1a* transcription (Fig. 4a) and the kinetic analysis of mRNA decay indicated that overexpression of either EPR or EPRSTOPE induces also a significant stabilization of *Cdkn1a* mRNA (Fig. 4b).

TGF-β signaling promotes tissue growth and morphogenesis during embryonic development while, as tissues mature, many cell types gain the ability to respond to TGF-β with growth arrest that is primarily due to imbalance of G1 events[27]. As similarly reported in other cell types[28,29], treatment of NMuMG cells with TGF-β for 1 h caused a rapid induction of *Cdkn1a* gene expression that was followed by return to baseline levels after 6 h (Fig. 4c). The observation that *Cdkn1a* return to baseline levels matches EPR downregulation (Fig. 4c) and that EPR overexpression strongly enhances *Cdkn1a* levels, prompted us to hypothesize a role for EPR in the TGF-β-dependent modulation of *Cdkn1a* gene expression. Our ChIP-qPCR assays showed that TGF-β treatment for 1 h stimulates the binding of SMAD3 to *Cdkn1a* promoter that returns to basal levels after 6 h (Fig. 4d, see also ref. [28]). TGF-β-dependent control of *Cdkn1a* mRNA decay was never investigated in detail but, considering that cells often achieve rapid changes of gene expression by integrating gene transcription control with regulated mRNA decay[30,31], we addressed the possibility that TGF-β could affect *Cdkn1a* mRNA decay. Figure 4e showed that *Cdkn1a* mRNA stability is unaffected by 1 h of TGF-β treatment (upper panel) but is reduced when the treatment is prolonged up to 6 h (lower panel). Thus, the TGF-β-dependent rapid fluctuations of *Cdkn1a* expression depend on the regulation of both transcription and mRNA decay in NMuMG cells.

Our hypothesis that EPR plays a role in the regulation of TGF-β-dependent CDKN1A expression was supported by the evidence that EPR silencing abrogated *Cdkn1a* mRNA induction upon TGF-β treatment for 1 h (Fig. 4f) while its overexpression enhances *Cdkn1a* levels and blunts its rapid modulation by TGF-β (Fig. 4g).

Together, our results indicate that EPR plays a dual role in TGF-β-dependent *Cdkn1a* gene expression control.

EPR affects both *Cdkn1a* gene transcription and mRNA decay. We investigated the molecular mechanism(s) by which EPR regulates *Cdkn1a* gene transcription. The evidence of enhanced RNA-Pol II occupancy and reduced presence of the H3K27me3 repressive mark at the *Cdkn1a* promoter in EPR-overexpressing cells (Supplementary Fig. 5a) together with our finding that EPR is present in the chromatin fraction (see Fig. 1c) prompted us to explore the possibility that EPR affects *Cdkn1a* transcription through direct interaction with its promoter region. Chromatin Isolation by RNA Purification (ChIRP)-Seq experiments (P.B., G. B., E. Zapparoli et al., unpublished) as well as ChIRP-qPCR experiments revealed the direct interaction of EPR with *Cdkn1a* promoter (Fig. 5a). The interaction of EPR with *Cdkn1a* promoter is not significantly affected by a 1 h TGF-β treatment (Supplementary Fig. 5b).

RIP experiments showed that SMAD3 interacts with EPR and the interaction is enhanced by treatment with TGF-β for 1 h (Supplementary Fig. 5c). In keeping with growing evidence suggesting that the interaction between lncRNAs and specific transcription factors can affect gene expression[5,32], ChIP-qPCR experiments showed that EPR overexpression enhances SMAD3-*Cdkn1a* promoter association and abrogates its dismissal after 6 h of TGF-β treatment (Fig. 5b). These effects are reproduced by overexpression of EPRSTOPE (Fig. 5b). Cell treatment with SB431542 abrogated the TGF-β-dependent enhancement of SMAD3-*Cdkn1a* promoter association in mock as well as in NMuMG cells overexpressing either EPR or EPRSTOPE (Supplementary Fig. 5d). Notably, EPR overexpression favored SMAD3-*Cdkn1a* promoter interaction also in untreated cells and this was not modified by SB431542 treatment (Fig. 5b and Supplementary Fig. 5d). To explain the association of SMAD3 with *Cdkn1a* promoter in cells overexpressing EPR also in the absence of TGF-β treatment, we hypothesize that EPR overexpression favors the association of SMAD3 molecules present in NMuMG cell nuclei of untreated cells with *Cdkn1a* promoter. Indeed, it is known that, although SMAD proteins rapidly accumulates into nuclei upon TGF-β treatment[33], a certain amount of SMAD3 is present in the nuclei of untreated cells (ref. [34]; Supplementary Fig. 5e).

In keeping with results shown in Fig. 4b, we found that EPR overexpression prevents *Cdkn1a* mRNA destabilization induced by a treatment with TGF-β for 6 h (Fig. 5c). Our initial observation that EPR interacts with KHSRP, a factor able to

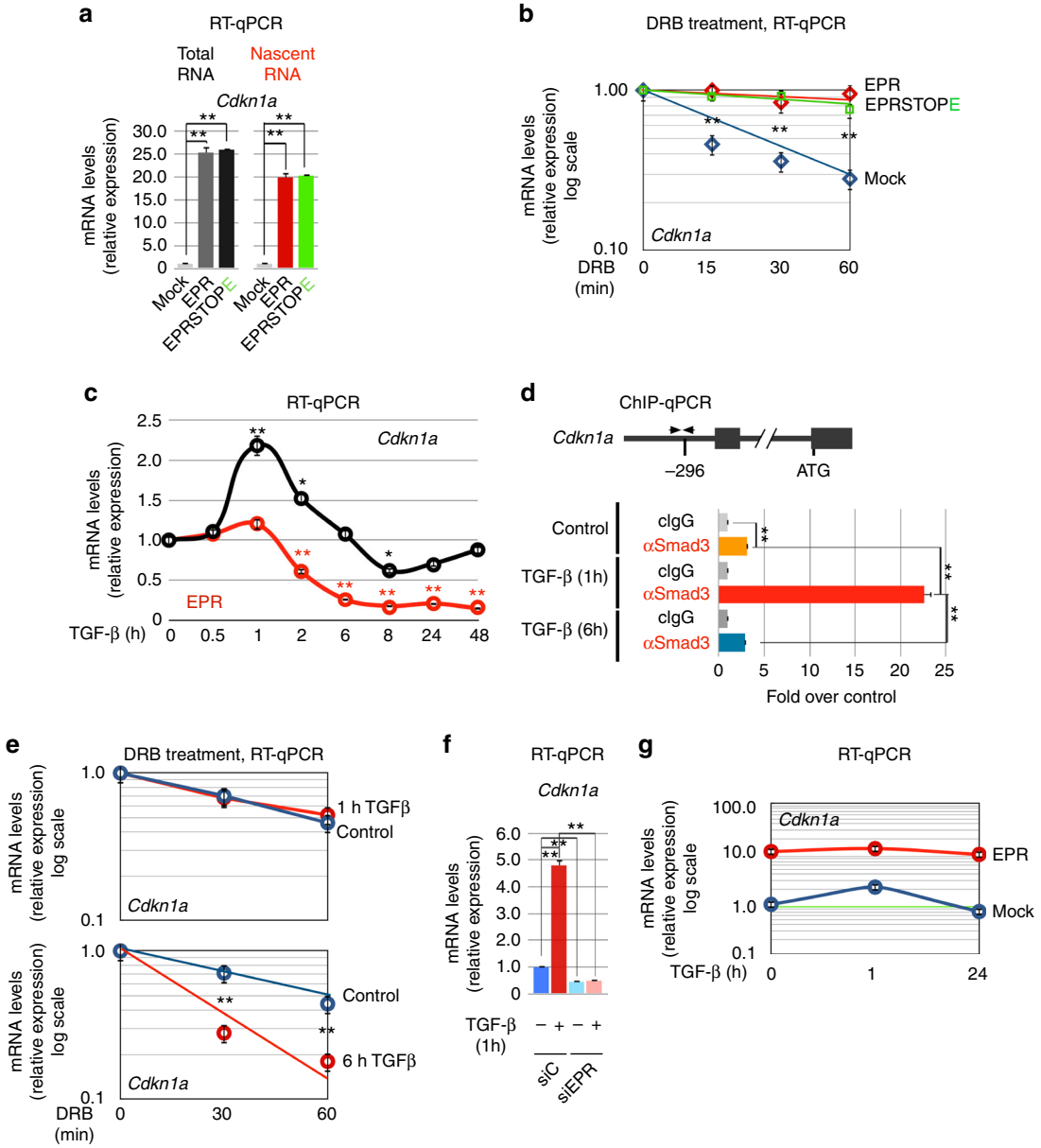
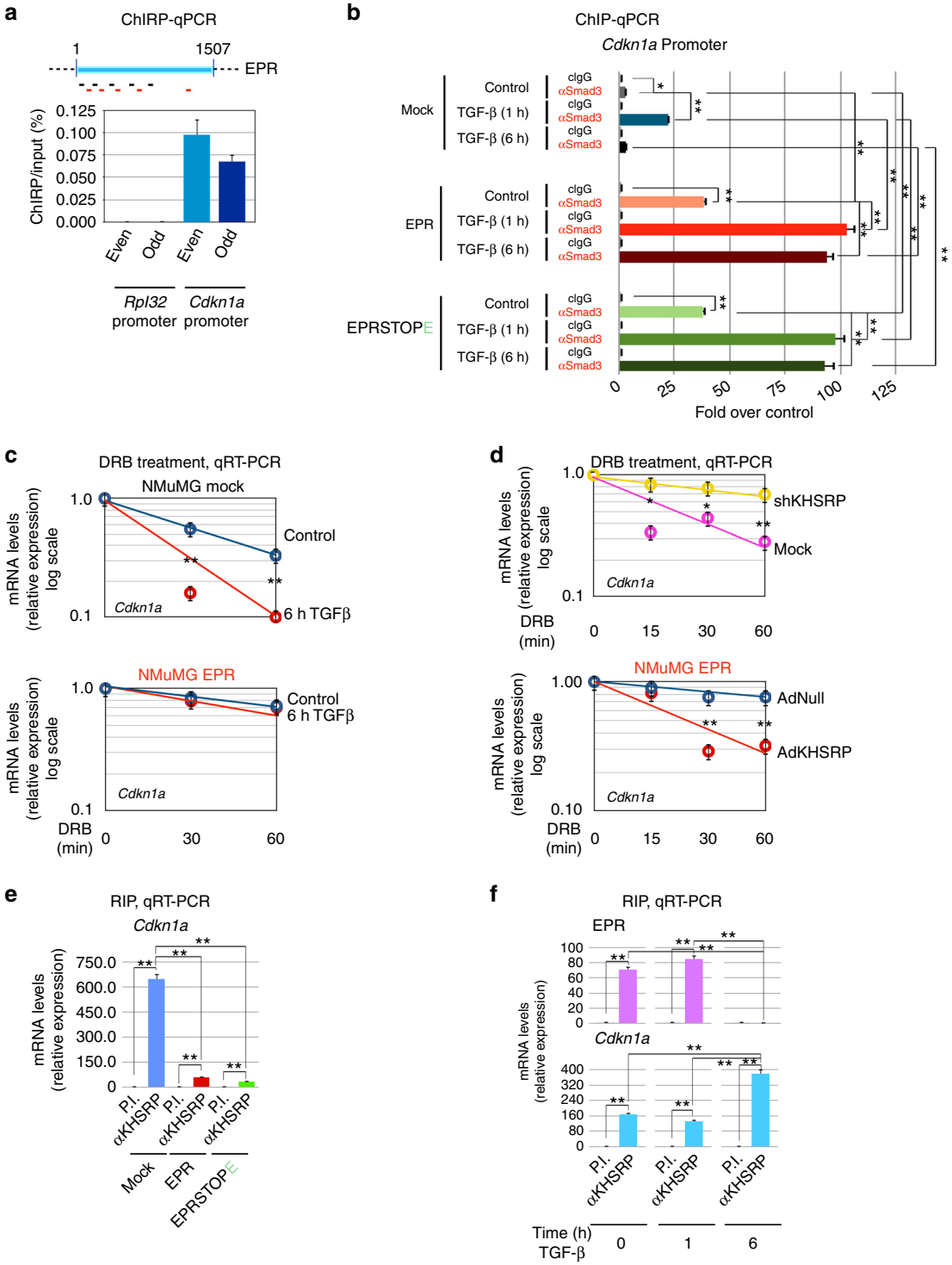


Fig. 4 Dual regulation of *Cdkn1a* gene expression by EPR. **a** qRT-PCR analysis of either total (left) or nascent (right) *Cdkn1a* transcript in either mock, EPR- or EPRSTOPE- overexpressing NMuMG cells. **b** Either mock, EPR- or EPRSTOPE-overexpressing NMuMG cells were treated with 100 μM 5,6-dichlorobenzimidazole 1-β-D-ribofuranoside (DRB) for different times (as indicated). *Cdkn1a* gene expression was analyzed by qRT-PCR. **c** qRT-PCR analysis of *Cdkn1a* (black line) and EPR (red line) in NMuMG cells serum-starved (2% FBS, 16 h) and either treated with TGF-β (10 ng ml^{-1}) for the indicated times or untreated (time 0). **d** Chromatin prepared from NMuMG cells serum-starved and either treated with TGF-β for the indicated times or untreated (control) was immunoprecipitated using either control IgG or affinity-purified anti-SMAD3 rabbit polyclonal antibody. The association of SMAD3 with *Cdkn1a* promoter (schematic on the top) was quantified by qPCR (primers indicated as arrowheads in the schematic above). **e** NMuMG cells were serum-starved (2% FBS, 16 h) and either treated with TGF-β (10 ng ml^{-1}) for either 1 h (top panel) or 6 h (bottom panel) or left untreated (control in both panels). Subsequently, cells were treated with 100 μM DRB for the indicated times and total RNA was isolated and analyzed by qRT-PCR to quantify *Cdkn1a* mRNA levels. Please note the slight, reproducible difference in the decay kinetic of *Cdkn1a* mRNA between nontransfected and mock-transfected NMuMG cells. **f** qRT-PCR analysis of *Cdkn1a* in NMuMG cells transiently transfected with either control siRNA (siC) or siRNA designed to silence EPR expression (siEPR), serum-starved (2% FBS, 16 h) and then either treated with TGF-β (10 ng ml^{-1}) for 1 h (+) or left untreated (−). **g** qRT-PCR analysis of *Cdkn1a* expression in either mock (blue line) or EPR overexpressing (red line) NMuMG cells serum-starved (2% FBS, 16 h) and either treated with TGF-β (10 ng ml^{-1}) for the indicated times or left untreated (time 0). Please note the logarithmic scale of the *Y*-axis. The values of both qRT-PCR and qPCR experiments shown are averages (±SEM) of three independent experiments performed in triplicate. Statistical significance: $^{*}p < 0.01$, $^{**}p < 0.001$ (Student's *t* test)

promote rapid decay of select labile mRNAs in many cell types[35], prompted us to explore whether KHSRP regulates *Cdkn1a* mRNA decay in NMuMG cells. KHSRP silencing induced *Cdkn1a* mRNA accumulation and prevented its rapid degradation (Supplementary Fig. 6a and Fig. 5d, upper panel) while transient KHSRP overexpression in NMuMG cells stably expressing EPR promoted *Cdkn1a* mRNA destabilization (Fig. 5d, lower panel). KHSRP is predominantly nuclear in NMuMG cells[15] and we found that mature *Cdkn1a* mRNA is abundant in nuclear fractions of these cells (Supplementary Fig. 6b) where it



undergoes rapid decay and is stabilized by EPR overexpression as measured by two independent techniques (Supplementary Fig. 6c, d). These findings suggested that EPR might interfere with the ability of KHSRP to interact with *Cdkn1a* mRNA. RIP experiments presented in Fig. 5e show that KHSRP association with *Cdkn1a* mRNA was strongly reduced by EPR and EPRSTOPE overexpression. Based on these results, we investigated whether EPR downregulation, that is physiologically obtained by TGF-β-treatment, affects the interaction of KHSRP with *Cdkn1a* mRNA. RIP analyses indicated that KHSRP-EPR association is abrogated while KHSRP interaction with *Cdkn1a* mRNA is increased upon TGF-β treatment (6 h) (Fig. 5f).

Altogether, these results suggest that EPR controls *Cdkn1a* expression by sustaining its transcription and by impairing its mRNA decay in response to TGF-β.

EPR overexpression reduces breast cancer cell proliferation. We investigated EPR in murine breast cancer cell lines and

Fig. 5 EPR associates with *Cdkn1a* promoter affecting its transcription as well as its mRNA decay. **a** ChIRP analyses. Top panel is a schematic of EPR and shows the location of biotinylated odd (black) and even (red) tiling oligonucleotides used for ChIRP. Both input and purified DNA were analyzed by qPCR to amplify either *Rpl32* (negative control) or *Cdkn1a* promoters. Values are averages (±SEM) of three independent experiments performed in triplicate. **b** Chromatin was prepared from either mock, EPR- or EPRSTOPE-overexpressing NMuMG cells serum-starved and either treated with TGF-β (1 ng ml^{-1}) for the indicated times or left untreated (control). Chromatin was immunoprecipitated using either control IgG or affinity-purified anti-SMAD3 rabbit polyclonal antibody. The association of SMAD3 with *Cdkn1a* promoter was verified by qPCR. **c** Either mock (top panel) or EPR-overexpressing (bottom panel) NMuMG cells were serum-starved, either treated with TGF-β (10 ng ml^{-1}) for 6 h or left untreated (control) and then treated with 100 μM DRB and total RNA was isolated and analyzed by qRT-PCR to quantify *Cdkn1a* mRNA levels. **d** Top panel, either mock or shKHSRP NMuMG cells were treated with 100 μM DRB. Total RNA was isolated at different times (as indicated) and analyzed by qRT-PCR to quantify *Cdkn1a* mRNA levels. Bottom panel, NMuMG cells stably overexpressing EPR were infected with either control (AdNull) or KHSRP-expressing (AdKHSRP) adenoviral vectors for 24 h then treated with 100 μM DRB. Total RNA was isolated at different times (as indicated) and analyzed by qRT-PCR to quantify *Cdkn1a* mRNA levels. NMuMG mock cells used for the experiment depicted in the upper panel differ from those presented throughout this report and have been previously described[15]. **e** Total extracts from either mock, EPR- or EPRSTOPE-overexpressing NMuMG cells were immunoprecipitated as indicated. RNA was purified from immunocomplexes and analyzed by qRT-PCR to quantify *Cdkn1a* mRNA levels. **f** Total extracts were prepared from NMuMG cells serum-starved and either treated with TGF-β (10 ng ml^{-1}) or left untreated (time 0) and immunoprecipitated as indicated. RNA was purified from immunocomplexes and analyzed by qRT-PCR. The values of both qRT-PCR and qPCR experiments shown are averages (±SEM) of three independent experiments performed in triplicate. Statistical significance: $^{*}p < 0.01$, $^{**}p < 0.001$ (Student's *t* test)

observed that its expression is severely reduced when compared with immortalized NMuMG cells (Supplementary Fig. 7a). Similarly, the expression of h.EPR was below detection levels in highly aggressive human breast cancer cell lines (M.J.M. and F.N., unpublished observation). H.EPR could be detected in about 75% of breast cancer primary samples (780/1043 cases from The Cancer Genome Atlas (TCGA) database[36]; Supplementary Fig. 7b) and, according to PAM50 molecular subtype classification, it was more expressed in Luminal A and Her2 tumors while it was almost absent in Basal-like tumors, the most frequent subtype of triple-negative breast cancers[37] (Fig. 6a).

Based on these observations, we decided to express EPR in triple-negative mesenchymal-like breast cancer cells, such as murine 4T1 and human MDA-MB-231 cell lines, respectively (Supplementary Fig. 7c).

Overexpression of either EPR or EPRSTOPE in 4T1 cells resulted in a strong induction of *Cdkn1a* gene expression as well as in a significant reduction of clonogenic potential, cell proliferation, and anchorage-independent cell growth (Fig. 6b-e). EPR overexpression in 4T1 cells also downregulated the expression of mesenchymal factors such as *Cdh2* and *Adam12* (Supplementary Fig. 7d). Very similar results were observed by overexpressing either human or murine EPR in human MDA-MB-231 cells (Fig. 6f, g, Supplementary Fig. 7e−g).

Finally, to interrogate the activity of EPR on cell proliferation control in vivo, we orthotopically injected either mock or EPR-expressing 4T1 cells into syngenic BALB/c mice. In concordance with our observations in cultured cells, EPR expression resulted in a remarkable reduction of tumor volume after 10 days (Fig. 6h, left panel). A significant reduction of the tumor mass was still evident and statistically significant also at 2 weeks after the transplant when mice were sacrificed (Fig. 6h, right panel, Supplementary Fig. 7h).

Altogether, our results indicate that EPR overexpression modulates cell proliferation and epithelial/mesenchymal markers levels in breast cancer cells and restrains cell proliferation in transplanted mice.

Discussion

Here we report on the initial functional characterization of the long intergenic noncoding RNA EPR well conserved among mammalian species and expressed in select epithelial tissues including differentiated luminal cells of human breast. The levels of EPR are rapidly downregulated by TGF-β/SMAD signaling in immortalized mammary gland cells and its sustained expression largely reshapes the transcriptome by inducing epithelial traits while preventing the acquisition of mesenchymal markers upon TGF-β treatment. Remarkably, EPR overexpression enhances the levels of the cyclin-dependent kinase inhibitor CDKN1A and strongly reduces cell proliferation in both immortalized and transformed mammary gland cells as well as in transplanted mice.

EPR is almost equally distributed in chromatin, nucleoplasm and cytoplasm and the cytoplasmic component associates with polysomes where a small peptide (EPRp) is translated. EPRp interaction with a variety of cytoskeletal and junctional proteins accounts for its junctional localization. However, the analysis of the phenotype that we observed in cells overexpressing EPR mutants unable to originate the peptide clearly indicates that the vast majority of gene expression changes that we describe here are independent of the peptide biogenesis.

In this report, we investigated how the lncRNA molecule per se controls gene expression and we focused our studies on the EPR-dependent regulation of CDKN1A that functions as both a sensor and an effector of multiple antiproliferative signals and promotes cell cycle arrest in response to TGF-β[26]. In NMuMG cells, TGF-β induces an early wave of *Cdkn1a* expression due, in part, to an increased SMAD complex-dependent gene transcription while a prolonged treatment causes the return of *Cdkn1a* levels to the baseline. Our data suggest that *Cdkn1a* promoter-bound EPR recruits SMAD3 molecules—that accumulate into the nucleus upon TGF-β treatment for 1 h—to induce rapid gene transcription. In parallel, EPR interacts with KHSRP limiting its association with *Cdkn1a* mRNA and this results in the transcript stabilization. We propose that EPR downregulation upon 6 h of TGF-β treatment causes SMAD3 dismissal from *Cdkn1a* promoter that results in a return of *Cdkn1a* transcription to basal levels and, in parallel, enables KHSRP to destabilize the *Cdkn1a* transcript . Our data suggest that EPR-regulated molecular events shape the rapid wave of *Cdkn1a* expression in response to TGF-β. The evidence that CDKN1A is more abundant in cells over-expressing EPR than in mock cells (independently of any treatment with TGF-β) allows us to hypothesize that overexpressed EPR is able to recruit SMAD3 molecules already present in cell nuclei to the *Cdkn1a* promoter region and, possibly, to distal enhancers as well as to block KHSRP-induced *Cdkn1* mRNA degradation.

Our data indicate that EPR couples *Cdkn1a* transcriptional regulation with mRNA decay control. Indeed, the integration of transcription and mRNA decay provides a kinetic boost to a series of processes that would be otherwise slower and less efficient. This report strengthen the idea that coupling transcription to mRNA decay enables cells to rapidly modulate waves of gene

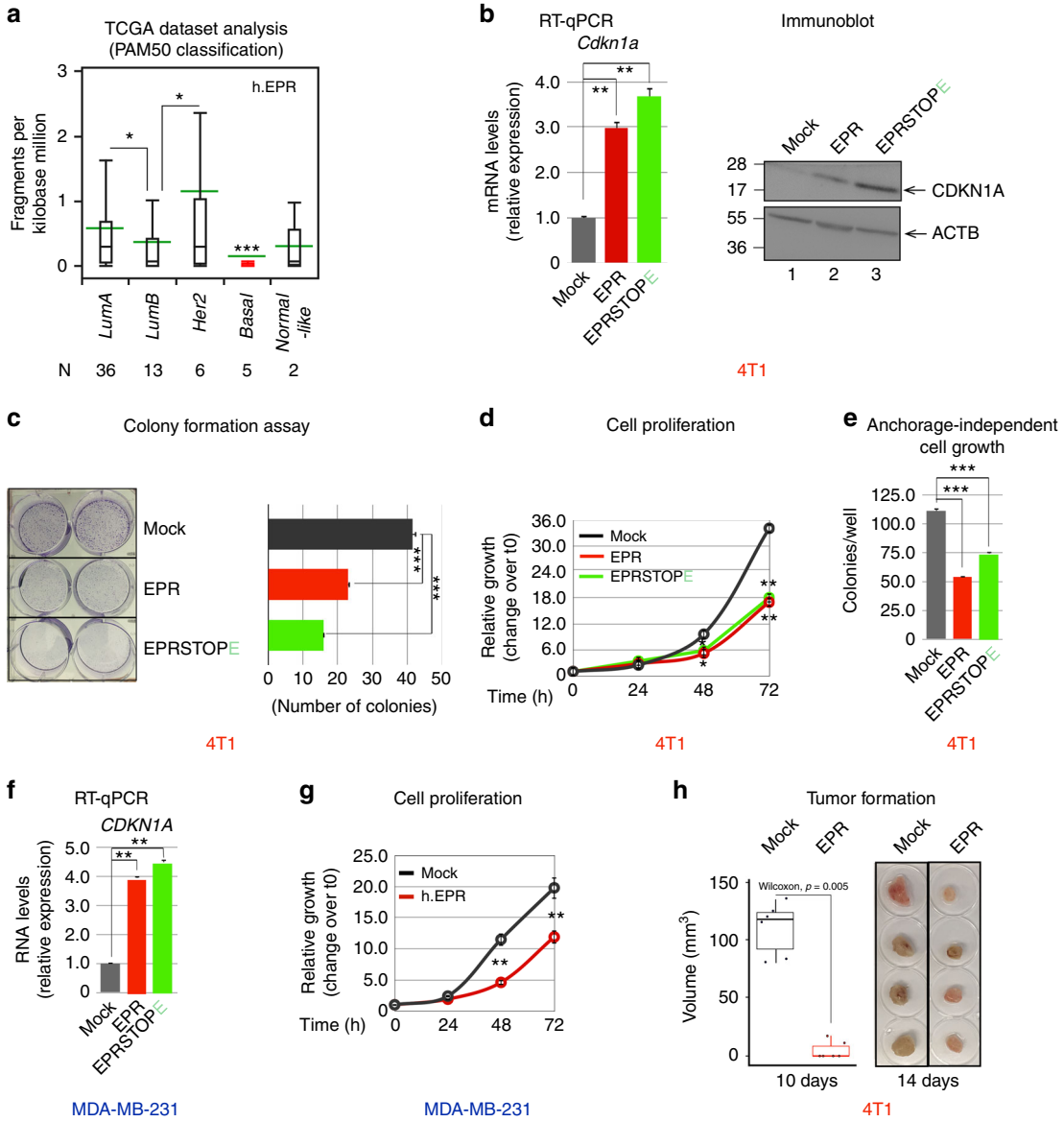


Fig. 6 Antiproliferative effect of EPR when expressed in transformed mammary gland cells. **a** Box plot shows the expression of h.EPR in the TCGA Breast Cancer (BRCA) dataset annotated according to PAM50 molecular subtype classification. The number of samples in each subtype is presented. Asterisks mark significant values (Wilcoxon's test; *p < 0.05, ***p < 0.001). **b** qRT-PCR analysis (left panel) and immunoblot analysis (right panel) of Cdkn1a expression in either mock, EPR- or EPRSTOPE-overexpressing 4T1 cells. **c** Either mock, EPR- or EPRSTOPE-overexpressing 4T1 cells were seeded at low density and colony-formation assays were performed after 4 days. A representative plate is shown. The numbers are averages (±SEM) of four independent experiments performed in duplicate. Statistical significance: ***p < 0.00001 (Student's *t* test). **d** Cell proliferation analysis of either mock, EPR- or EPRSTOPE-overexpressing 4T1 cells. **e** Cells were cultured in soft agar for 21 days and phase contrast micrographs were taken at ×10 magnification. The values are averages (±SEM) of four independent experiments performed in triplicate. Statistical significance: ***$p < 0.00001$ (Student's *t* test). **f** qRT-PCR analysis of *CDKN1A* mRNA levels in transfected MDA-MB-231 cells. **g** Cell proliferation analysis of either mock or h.EPR-overexpressing MDA-MB-231 cells. **h** Tumor volume ($n = 6$ tumors/group) was measured by digital caliper assessment 10 days after injection of either mock or EPR-expressing 4T1 cells in BALB/c mice. Box plot analysis of tumor volume is shown (left). Box plots display summaries of the data distribution: the lower whisker is the minimum, the lower box edge is first quartile, the middle line is the median, the upper box edge is the third quartile, and the upper whisker the maximum value. Data were analyzed in R version 3.4.3, using the Wilcoxon Unpaired Test as implemented in "stat_compare_means" in "ggplot 2.2.1". Images of the tumors at the end of the experiment, 2 weeks after injection (right). The values of qRT-PCR experiments shown are averages (±SEM) of three independent experiments performed in triplicate. Statistical significance: **$p < 0.001$ (Student's *t* test). The values of cell proliferation experiments (panels **d** and **g**) are averages (±SEM) of three independent experiments performed in triplicate. Statistical significance: *$p < 0.01$, **$p < 0.001$ (Student's *t* test)

expression in response to a variety of stimuli thus achieving optimal mRNA homeostasis[30,31].

Since the decay of mature mRNAs is generally thought to occur in the cytoplasm, our finding that spliced and polyadenylated *Cdkn1a* mRNA is abundant in the nucleus of NMuMG cells, where it undergoes regulated decay, might be somehow surprising. However, the possibility that mature mRNAs accumulate in the nuclei of mammalian cells has been described[38,39] and a previous report has suggested that *Cdkn1a* mRNA undergoes degradation by the nuclear exosome in mammalian cells[40]. Further, a report about mature RNA degradation controlled by nuclear histone acetyltransferases and deacetylases provocatively

suggested the possibility that mRNA degradation pathways might operate also in the nucleus[41].

The TGF-β cytostatic program in epithelial cells involves, among other molecular events, the induction of *Cdkn1a* and *Cdkn2b* but cancer cells utilize any opportunity to circumvent TGF-β ability to inhibit cell proliferation. Inactivating mutations in the TGF-β Type-II receptors and SMAD4 have been described in tumors even though cancer cells can lose the cytostatic responsiveness due to defects downstream to SMAD factors[42,43]. We hypothesize that the absence of EPR/h.EPR, that occurs in breast cancer cell lines and in certain breast cancers, may contribute to the loss of TGF-β ability to restrain cell proliferation while may enable the cytokine to sustain their carcinogenic potential.

The notion that lncRNAs are devoid of coding potential has been recently challenged by reports demonstrating that translated short peptides are responsible for the biological functions of the respective lncRNA[22,23,44,45]. However, Yu et al.[46] reported that linc-RAM, an lncRNA that per se enhances myogenic differentiation by interacting with MYOD, is the transcript encoding myoregulin, a small peptide previously reported as a mediator of muscle performance through inhibition of the pump activity of SERCA[44]. In this case, RNA and peptide functions cooperate in modulating muscle physiology. Intriguingly, under our model the lncRNA per se is responsible for most of the gene expression changes while the peptide might be responsible for specific functions related to its cytoskeletal/junctional localization. We hypothesize that the peptide could participate in multiprotein complexes serving as permeability barriers and/or it could be implicated in the establishment of apico-basal polarity as well as in the transduction of signals to the cell interior. Further studies will be needed to clarify the specific functions of the peptide but we can predict that EPR and the peptide might synergize in executing epithelial cell-specific programs.

EPR was discovered as a highly regulated lncRNA in NMuMG mammary gland cells and, by exploiting the TGF-β-responsiveness of these cells, we were able to clarify some details of its molecular functions. How EPR influences normal physiology and disease in tissues undergoing repeated rounds of proliferation/differentiation, such as the gastrointestinal tract where it is highly expressed, will represent an important area of future research.

Methods

Cell lines. Murine immortalized NMuMG cells (ATCC, no. CRL-1636) were cultured in Dulbecco's Modified Eagle Medium (DMEM) plus 10% Fetal Bovine Serum (FBS) and 10 μg ml^{-1} bovine insulin (Sigma-Aldrich), 4T1 mouse mammary gland cancer cells (obtained from ATCC, no. CRL-2539) were cultured in DMEM/F12 plus 10% FBS, human mammary gland adenocarcinoma cells MDA-MB-231 (obtained from DSMZ, Germany, through Dr. G. Fronza, authenticated by STR DNA profiling) were cultured in DMEM plus 5% FBS, and human HEK-293 cells (ATCC, no. CRL-1573) were cultured in DMEM plus 10% FBS. HEK-293 cells were used to verify transient expression of FLAG-tagged EPRp based on their highly efficient transfectability. NMuMG cells were maintained in DMEM supplemented with 2% for 16 h prior to the addition of 1–10 ng ml^{-1} human recombinant TGF-β1 purchased from R&D Systems. All cell lines were tested for mycoplasma contamination and resulted negative. SB431542 compound was purchased from Sigma-Aldrich, dissolved in Dimethyl sulfoxide (DMSO), and used at a 1 μM concentration. Cycloheximide, dissolved in DMSO, was purchased from Sigma-Aldrich and used at a 5 μM concentration.

Antibodies and immunoblots. Anti-EPRp polyclonal rabbit antibody was generated by injecting rabbits with recombinant purified EPR expressed in *E. Coli* using the pQE-EPR at Cambridge Research Biochemicals (Billingham, Cleveland, UK). Anti-CDH1 goat polyclonal antibody (sc-31020, used at 1:500 final dilution), anti-CDKN1A mouse monoclonal antibody (sc-6246, used at 1:200 final dilution), and anti-HDAC1 rabbit polyclonal antibody (sc-7872, used at 1:500 final dilution) were from Santa Cruz; anti-TJP1 rabbit polyclonal antibody (ab96587, used at 1:100 final dilution), anti-SMAD3 rabbit polyclonal antibody (ChIP grade ab28379), anti-GFP rabbit polyclonal antibody (ChIP grade ab90, used at 1:200 final dilution) were from Abcam; mouse monoclonal anti-FLAG (F1804, used at 1:500 final dilution), mouse monoclonal anti-TUBA (DM1, used at 1:1000 final dilution) and mouse monoclonal anti-ACTB (AC-74, used at 1:30,000 final dilution) were from Sigma-Aldrich. Mouse monoclonal anti-RNA Polymerase II (clone CTD4H8) and rabbit polyclonal antibody to H3K27me3 (CS200603) were from Millipore. Rabbit polyclonal anti-CGN serum (C532, used at 1:5000 final dilution) against a purified recombinant 50 kDa C-terminal fragment of chicken cingulin as well as anti-CGNL1 rabbit polyclonal antibody (20893, used at 1:100 final dilution) were raised at the University of Geneve. Images of the uncropped and unprocessed scans of the most important Immunoblots are presented in Supplementary Fig. 8.

Plasmids. Plasmid EPR was obtained by inserting the sequence from nucleotide 1 to 1487 of murine BC030870 into pBICEP-CMV-2 vector (Sigma-Aldrich); plasmid h.EPR was obtained by inserting the sequence from nucleotide 4 to 1126 of human LINC01207 into pBICEP-CMV-2 vector; plasmids EPRSTOPE and EPRSTOPM were obtained by Site-Directed Mutagenesis of plasmid EPR using the QuikChange II mutagenesis kit (Agilent Technologies) and the oligonucleotides 5′—CACCGTTAGTCTTCCATGTAGCTACCATTC—3′ and 5′—CACCGTTAGTCTTCCTAGTAGCTACCATTC—3′, respectively. Plasmids EPR-FLAG and EPRSTOPE-FLAG were generated by inserting the sequence from nucleotide 1 to 560 of murine BC030870 obtained by PCR and Flagged at its 3′ (either wild-type or mutagenized as above) into pIRES1*hyg* vector. Plasmids GFP-mouse cortactin (#26722, CTTN-GFP) and CMV-GFP-human NMHC II-A (#11347, MYH9-GFP) were obtained from Addgene. Plasmid pQE-EPR was obtained by inserting the sequence from nucleotide 345 to 560 of murine BC030870 into pQE-30 vector (Qiagen).

For inserts obtained by RT-PCR, the Pfu DNA Polymerase (Promega) was used. The inserts cloned in all constructs were sequenced on both strands (BMR Genomics, Padova, Italy).

Cell transfections. NMuMG, 4T1, and MDA-MB-231 cells were transfected with Lipofectamine 2000 (ThermoFischer) while HEK-293 cells were transfected with Attractene transfection Reagent (Qiagen). NMuMG, 4T1, and MDA-MB-231 cells stably transfected with recombinant pBICEP-CMV-2-based vectors were maintained in selective medium containing 800, 350, and 750 μg ml^{-1} G418 (Sigma-Aldrich), respectively. NMuMG cells stably transfected with recombinant pIRES1*hyg*-based vectors were maintained in selective medium containing 600 μg ml^{-1} Hygromycin B (Sigma-Aldrich). Specific mock cells were generated (for every cell line and every plasmid backbone) by transfecting the corresponding empty vector in each cell type. Mock cells were subjected to a selection procedure identical to the other transfectants. siRNAs utilized to knockdown murine EPR (5′—GAGCAAAAGAGAAUGCUUA—3′) were purchased from Thermo Fisher. Stable KHSRP knockdown in NMuMG cells was obtained using previously described silencing sequences and pSuper-Neo (Oligoengine) according to the manufacturer's instructions[15]. The adenoviral vectors pAdCMVnull (AdNull) and pAdKHSRP (full-length human KHSRP cDNA cloned into an Adenoviral-Type 5 backbone) were purchased from Vector Biolabs[15].

Scratch wound closure assay. Either mock or EPR-overexpressing NMuMG cells were cultured in six-wells plates up to confluence and pretreated for 2 h with 5 μg ml^{-1} Mitomycin C (Sigma-Aldrich). A wound was scratched into monolayers and cells were cultured for up to 48 h in the presence of 5 μg ml^{-1} Mitomycin C. Images were taken using an Olympus CKX41 microscope and analyzed using the ImageJ 1.49r package (http://imagej.nih.gov/ij/index.html). Average distance of wound obtained from six microscopic fields was used for the calculation of percent wound healed. Experiments were performed three times in triplicate.

Immunofluorescence. Either mock NMuMG cells or stable transfectants overexpressing either EPR-FLAG or EPRSTOPE-FLAG were plated on glass coverslips in 24-well plates (60,000 cells/well). Immunofluorescence was carried out 2 days after plating essentially as reported in ref. [47]. Rabbit polyclonal anti-cingulin antiserum (C532) was used at a 1:5000 dilution while anti-FLAG antibody (F1804, Sigma) was used at a 1:500 dilution. Secondary antibodies were diluted in IF buffer and incubated for 30 min at 37 °C, Alexa488 anti-rabbit (711-545-152, Jackson Laboratory) dilution 1:400, Cy3 anti-mouse (715-165-151, Jackson Laboratory) dilution 1/400. Pictures were taken using a Zeiss Axiophot widefield fluorescent microscope (X-Cite 120Q mercury lamp light source, Excelitas Technologies; retiga EXi, cooled mono 12-bit, Qimaging camera; ×63 oil objective; Openlab software). Images were imported into ImageJ to split and merge channels, cropped and adjusted for resolution and for intensity level range using Photoshop (scale bar = 10 μm).

Orthotopic 4T1 injection in BALB/c mice. BALB/c 8–10-week-old female mice (Envigo) were anesthetized using 100 mg kg^{-1} ketamine and 10 mg kg^{-1} xylazine intra peritoneal. Eye lubricant was applied, hair around the abdominal and inguinal fat pads were trimmed and the skin was sterilized. With the aid of magnifying surgical loupes, a small incision of less than 3 mm was made externally and caudally to the fourth nipple with the tip of micro-dissecting scissors. The fourth mammary gland fat pad below was located and 100 μl of a suspension of either

mock or EPR-expressing 4T1 cell were injected. Successful injection is confirmed by the swelling of the tissue. The incision was then sutured. All procedures involving animals have been approved by the Institutional Animal Welfare Body (O.P.B.A.) and complied with the national current ethical regulations regarding the protection of animals used for scientific purpose (D. Lvo, March 4, 2014, n. 26, legislative transposition of Directive 2010/63/EU of the European Parliament and of the Council of September 22, 2010 on the protection of animals used for scientific purposes). Tumor length and width were measured using a digital caliper at day 10 post injection and tumor volume was calculated using the formula: volume = (length × (width)^2/2). Mice were euthanized after 2 weeks and tumor masses were removed, weighted and photographed.

Electrophoretic mobility shift assay. Electrophoretic mobility shift assays (EMSA) were performed utilizing purified recombinant proteins that were incubated at room temperature for 20 min in a RNA-binding buffer (20 µl) containing 10 mM 4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) (pH 7.6), 3 mM $MgCl_2$, 100 mM KCl, 2 mM Dithiothreitol (DTT), 5% glycerol, 0.5% NP-40, yeast RNA (1 µg), and heparin (1 µg). The labeled RNA was transcribed using Sp6 polymerase from a template generated by inserting into pCY vector[48–50], a PCR product corresponding to nucleotides from 276 to 407 of murine BC030870.

RNA isolation from cytoplasm, nucleoplasm, and chromatin. We followed the protocol recently published by Corey and coworkers[51] starting from 10×10^7 cells. Both cytoplasmic and nucleoplasmic RNAs were precipitated and washed with ice-cold 70% (vol/vol) ethanol prior to be dissolved in QIAzol Lysis Reagent (Qiagen) while the chromatin pellets were immediately dissolved in QIAzol. Ten microliters of 0.5 M EDTA was added to all the samples in QIAzol that were heated to 65 °C with vortexing until dissolved (~10 min). The preparation of RNA was continued as described below. In parallel to RNA, protein extracts were prepared as described by Corey and coworkers[51].

qRT-PCR, analysis of nascent transcripts, and mRNA decay. Total RNA was isolated using either the miRNeasy mini kit or QIAzol (Qiagen) and retro-transcribed (50–100 ng) using Transcriptor Reverse Transcriptase (Roche) and random hexamers in most experiments according to the manufacturers' instructions. In order to verify if EPR is polyadenylated, qPCR amplification was performed using as template the product of reverse transcription reactions performed with oligo-dT (that pairs with the poly-A tail). Quantitative PCR was performed using the Precision 2× QPCR master mix (Primer Design), and the Realplex II Mastercycler (Eppendorf) according to the manufacturers' instructions. The sequence-specific primers utilized for PCR reactions are listed in Supplementary Data 3. In order to analyze gene expression changes among the pool of nascent mRNAs, we adopted the Click-iT Nascent RNA Capture kit (ThermoFischer) and performed experiments according to the manufacturer's instructions. NMuMg cells were pulsed with 0.5 mM 5-ethynyl Uridine (EU) for 1 h. In order to analyze mRNA decay we either blocked transcription by treating cells with 100 µM 5,6-Dichlorobenzimidazole 1-β-D-ribofuranoside (DRB, Sigma-Aldrich) and isolating total RNA at different intervals of times or performing EU labeling-based pulse chase experiments labeling cells with 0.2 mM EU for 16 h, removing the culture medium, and chasing cells for 1 h. RNA was prepared, clicked, retrotranscribed, and analyzed by qRT-PCR according to Click-iT Nascent RNA Capture kit instructions.

Ribonucleoprotein complexes immunoprecipitation (RIP) assays. Briefly, total cell lysates were immunoprecipitated with Dynabeads (Thermo Fisher) coated with protein A/protein G and precoupled to specific antibodies at 4 °C overnight. Pellets were washed three times with a buffer containing 50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 0.5% Triton X-100, 1× Complete (Roche)[17]. Total RNA was prepared from immunocomplexes using the QIAzol Lysis Reagent, retrotranscribed, and amplified by qPCR as described above. The primer sequences are detailed in Supplementary Data 3.

Protein identification by mass spectrometry (MS) analysis. Total extract from either mock or EPR-FLAG (10×10^7 cells) were immunoprecipitated using anti-FLAG antibody-coupled Dynabeads. Immunoprecipitated material was analyzed by SDS-PAGE followed by silver staining. Protein identification was performed as a service at the Functional Proteomic Unit of IFOM (Milano, Italy; Drs. Angela Cattaneo and Angela Bachi). Bands of interest from SDS-PAGE were excised from gels, reduced, alkylated and digested overnight with bovine trypsin (Roche, Milan, Italy), as described[52]. One microliter aliquots of the supernatant were used for mass analysis using the dried droplet technique and α-cyano-4-hydroxycinnamic acid as matrix. Mass spectra were obtained on a MALDI–TOF Voyager-DE STR mass spectrometer (Applied Biosystem). Alternatively, acidic and basic peptide extraction from gel pieces after tryptic digestion was performed and the resulting peptide mixtures subjected to a single desalting/concentration step before MS analysis over Zip-TipC18 (Millipore Corporation). Spectra were internally calibrated using trypsin autolysis products and processed via Data Explorer software. Proteins were unambiguously identified by searching a comprehensive nonredundant protein database of the National Center for Biotechnology Information (NCBI, https://www.ncbi.nlm.nih.gov/) and the Mass Spectrometry protein sequence DataBase (MSDB, http://msdn.microsoft.com/en-us/library/ms187112.aspx), selected by default using in-house software programs ProFound v4.10.5 and Mascot v1.9.00, respectively. Protein identifications were accepted if they could be established at greater than 99.0% probability and contained at least three identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm[53]. Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony.

RNA deep-sequencing (RNA-Seq). High-quality RNA was extracted from either mock, EPR-, or EPRSTOPE- overexpressing NMuMG cells (biological triplicates for each experimental condition), and a total of nine libraries were prepared using standard Illumina TrueSeq SBS PE 200 cycles protocol and sequenced on HiSeq2500. Image analysis and base calling were performed using the HiSeq Control Software and RTA component from Illumina. This approach yielded between 68 and 77 millions of reads that were further processed.

Analysis of h.EPR (LINC01207) expression in human samples. Meta-analysis of RNA-Seq data of h.EPR in normal samples was performed by searching for h.EPR expression in different subpopulations of FACS-sorted normal breast cells[21] and in different human organs through either the Expression Atlas (https://www.ebi.ac.uk/gxa/home) or the GEPIA web server (http://gepia.cancer-pku.cn). h.EPR expression in breast cancer samples was analyzed using TCGA data, deriving PAM50 [37].

Chromatin isolation by RNA purification (ChIRP). Chromatin isolation by RNA purification (ChIRP) was performed according to the protocol published by Chu et al.[54] with minor modifications. Briefly, 2.5×10^7 NMuMG cells were crosslinked in 20 ml of 1% glutaraldehyde in PBS at room temperature for 10 min on an end-to-end rotator. After glutaraldehyde quenching and repeated washes, cell pellets were weighted and resuspended in 1.0 ml of complete Lysis Buffer (50 mM Tris-Cl pH 7.0, 10 mM EDTA, 1% SDS, 1× Complete, 500U RNAse inhibitor) per each 100 mg of cell pellet. Cell suspensions were sonicated for 90 min (power set to 70%) and the sonicated cell lysate was centrifuged at 16,100 × *g* at 4 °C for 10 min. Lysates were divided into two 1 ml aliquots, transferred into polypropylene tubes, mixed with 2 ml Complete Hybridization Buffer (750 mM NaCl, 1% SDS, 50 mM Tris-Cl, pH 7.0, 1 mM EDTA, 15% formamide, 1× Complete, 1000 U RNAse Inhibitor) and hybridized with 1 µl (100 pmol) of either EVEN or ODD pools of 20-mer 3′ Bio-TEG DNA oligonucleotides designed with single-molecule FISH online designer (Stellaris) (see Supplementary Data 3), respectively. Hybridization was carried out at 37 °C for 4 h under continuous shaking. Seventy microliters of prewashed C-1 magnetic beads (ThermoFisher) were added to each hybridization mixture for 30 min at 37 °C under continuous shaking. Beads were immobilized and washed four times for 5 min at 37 °C with shaking (wash buffer: 2× NaCl and Sodium citrate (SSC), 0.5% SDS, 1× Complete). While one aliquot (10% of the material) was utilized for RNA extraction, the remaining 90% was subject to DNA purification by incubating two times each bead pellet with 150 µl Complete DNA Elution Buffer (50 mM $NaHCO_3$, 1% SDS, 25 µg ml^{-1} RNAseA, 100 U ml^{-1} RNAseH) for 30 min at 37 °C with shaking. Eluted DNA was incubated with Proteinase K (1 mg ml^{-1} final dilution) for 45 min at 50 °C with shaking, extracted with Phenol/Chlorophorm/Isoamylalchool, ethanol-precipitated, and aliquots were analyzed by qPCR.

ChIP-qPCR. ChIP experiments were performed according to the protocol published by Ghisletti et al.[55]. Briefly, ChIP lysates were generated from 40×10^6 cells. Each lysate was immunoprecipitated with 10 µg of anti-Pol II, anti-H3K27me3, and anti-SMAD3 antibodies (and the corresponding control IgG).

Antibodies were prebound overnight to 100 µl of A/G protein-coupled paramagnetic beads (ThermoFisher) in PBS/BSA 0.5%. Beads were then added to lysates (the preclearing step was omitted), and incubation was allowed to proceed overnight. Beads were washed six times in a modified RIPA buffer (50 mM HEPES (pH 7.6), 500 mM LiCl, 1 mM EDTA, 1% NP-40, and 0.7% Na-deoxycholate) and once in TE containing 50 mM NaCl. DNA was eluted in TE containing 2% SDS and crosslinks reversed by incubation overnight at 65 °C. DNA was then purified by Qiaquick columns (Qiagen) and quantified with PicoGreen (ThermoFisher).

Sucrose-gradient fractionation and polysome profiling. Experiments were performed as described[56]. NMuMG cells (~70% confluence) were treated with cycloheximide (0.1 mg ml^{-1}) for 5 min at 37 °C, washed twice with PBS supplemented by 0.01 mg ml^{-1} cycloheximide, scraped in PBS 1× with 0.01 mg ml^{-1} cycloheximide, pelleted by centrifugation, lysed in 500 µl of ice-cold Lysis Buffer (Salt Solution 1×, 1% Triton-X100, 1% NaDeoxycholate, 0.2 U µl^{-1} RNase Inhibitor, 1 mM DTT, 0.01 mg ml^{-1} cycloheximide), centrifuged for 5 min at 16,000 × *g* at 4 °C, and supernatants were loaded onto sucrose gradients. One milliliter fractions were collected monitoring the absorbance at 260 nm using a Density Gradient Fractionation System by Teledyne ISCO with sensitivity set to 0.2. Using the profile of the 260 nm absorbance, fractions corresponding to free ribosomal subunits (40S and 60S) and monosomes (80S, considered as not translating), separately from fractions corresponding to light polysomes (2–5 ribosomes) and

heavy polysomes (>6 ribosomes) were pooled together and processed for RNA extraction and RNA was quantified by Nanodrop (ThermoFisher).

Cell cycle analyses by flow cytometer. NMuMg cells (either mock or EPR- or EPRSTOPE-transfected) were seeded in six-well plates. For the analysis by the Cycletest^{TM} Plus DNA Kit (BD Medical Technology), cells are detached by trypsinization and centrifuged in Eppendorf tubes at 300 × *g* for 5 min at room temperature. Supernatant is removed and the pellet is resuspended in 1 ml 1× PBS followed by centrifugation. Cells are then resuspended in PBS and counted using Countess® Automated Cell Counter and the cell concentration is adjusted to 7×10^5 cells ml^{-1} using the same buffer. The DNA staining procedure is performed using 0.5 ml of cell suspension (7×10^5 cells). Cells are pelleted by centrifugation (400 × *g* for 5 min at RT). After carefully removing the supernatant, cells are mixed in Solution A (provided by the kit, containing trypsin in a spermine tetrahydrochloride buffer for digestion of cell membranes and cytoskeleton), without using a vortex. Two hundred microliters of solution B (provided by the kit, containing trypsin inhibitor and ribonuclease A in citrate-stabilizing buffer, to inhibit the trypsin activity and to digest the RNA) is gently added and the sample is incubated for 10 min at RT, followed by the addition of 200 µl of cold solution C (provided by the kit, containing Propidium Iodide and spermine tetrahydrochloride in citrate-stabilizing buffer). The sample is incubated in the dark and on ice for 10 min and then filtered by cell strainer caps and analyzed by flow cytometer (BD FACS Canto^{TM}). Data on at least 10,000 events for sample were processed using ModFit LT 4 software. The experiment was repeated two times.

To estimate more precisely the fraction of cells in S phase, the Click-iT^{TM} Plus EdU Flow Cytometry Assay (Invitrogen) was used. EdU (10 mM stock in DMSO) was added directly to the culture medium at the 20 µM final concentration and incubated for 40 min. Cells were then harvested by trypsinization and washed using 3 ml of PBS containing 1% BSA. Pellets are resuspended in PBS+1% BSA, counted using Countess® Automated Cell Counter and 1.5×10^6 cells are transferred to flow tubes, washed again with 3 ml of PBS containing 1% BSA, pelleted by centrifugation followed by removal of the supernatant. Cells are resuspended in 100 µl of Click-iT^{TM} fixative mixing well with a pipette and incubated for 15 min at RT in the dark. Cells are then washed as performed in the previous step, resuspended and incubated for 10 min in 100 µl of 1× Click-iT^{TM} saponin-based reagent. Samples are then processed for the Click-iT^{TM} reaction, preparing the Click-iT^{TM} Plus reaction cocktail according to the manufacturer's guidelines, adding 0.5 ml of it to each sample, to reach a final volume of 600 µl containing 1.5×10^6 cells and incubating for 30 min at room temperature, in the dark. Cells are then washed once using 3 ml of 1× Click-iT^{TM} saponin-based reagents, pelleted and resuspended in 600 µl of the same solution to which the propidium iodide staining solution is added to stain DNA. Propidium iodide solution contains 50 µg ml^{-1} PI and 100 µg ml^{-1} RNAse. Samples are then analyzed by flow cytometer (counting 20,000 events, BD FACS Canto^{TM}). As controls, cell aliquots incubated with EdU and processed by the same protocol, but skipping the Click-iT^{TM} reaction or the PI staining, or both.

G1 phase cell sorting. Cells (mock, EPR, EPRSTOPE) were harvested by trypsinization when they reached ~90% confluence, washed once in 1× PBS and resuspended in DMEM without serum at a concentration of 10^6 cells ml^{-1}. Hoechst 33342 (ThermoFisher Scientific) was added to the media at the concentration of 10 µg ml^{-1} and cells were incubated for 1 h at 37 °C. Cells were then centrifuged to remove Hoechst-containing media and resuspended in 1× PBS.

Sorting was performed by BD Aria II^{TM} cytometer (BD Bioscience) using a 100 µm nozzle and setting a gate on the population of cells in G1. At least 90 K events for every sample were sorted in 1× PBS at room temperature. After sorting, purity was assessed by re-running the samples. Sorted cells were pelleted and immediately stored at −80 °C. RNA was extracted and analyzed by qRT-PCR as described above.

Cell proliferation analysis by high-content image analysis. The proliferation of NMuMG cells (either mock or EPR- or EPRSTOPE-transfected) was quantified using Operetta High-Content Imaging System, acquiring images at different time points by digital phase contrast with a ×40 objective. Images were analyzed using Harmony® High Content Imaging and Analysis Software. Five hundred cells were seeded in 96-well plates in triplicates. Pictures were taken at different time points, by automatically acquiring eight fields for each well. Data were analyzed in Excel and plotted as average and standard deviations of replicates.

Quantification of cell proliferation by crystal violet. For some experiments cell proliferation was assessed by crystal violet staining. At the indicated time after plating, cells were fixed (10% formalin) and stained (0.1% crystal violet) with crystal violet solution. After two washes with water, crystal violet staining was measured by spectrophotometer at a wavelength of 590 nm.

Clonogenic and anchorage-independent cell growth assays. For the clonogenic assays, cells were plated in triplicate on six-well plates at 500 cells per well and left to grow for 4–6 days. Cells were fixed and stained with crystal violet solution. Anchorage-independent cell growth assays were assessed according to the protocol published by Borowicz et al.[57] with minor modifications. Briefly, 2500 cells were seeded in 0.3% top agar in complete medium and placed on a layer of 0.5% of bottom agar in 12-well plates. Each cell line was seeded in sextuplicates and fed every 3 days. After 21 days cells were colored with crystal violet and photographs were taken.

Quantification and statistical analysis of RNA-Seq. Raw FASTQ reads were trimmed at the ends to remove low-quality calls with FASTX (http://hannonlab.cshl.edu/fastx_toolkit). Paired-end reads were aligned to indexed mm10 genome with STAR (v 2.3.0e_r291).

To quantify expression levels mapped reads were counted from BAM files with HTSeq counts version 1.2.1, in intersection-strict mode, feature type exon and id attribute gene_name against reference annotation Ensembl GRCm38.74.

Quantitation of transcript differential expression analysis. In addition to gene-level analysis with STAR-HTSeq, the transcript abundance was further re-estimated using an alignment-free approach based on Kallisto 0.43.1 software, using Gencode Mouse vM15 transcripts as reference.

Abundance files were imported in R.3.1.1 with TxImport.1.2.0 with option txOut =TRUE to quantify alternatively spliced transcripts. edgeR_3.16.5 and limma_3.30.13 were used to log2 transform transcripts count in Count Per Million (cpm). Only transcript with ≥1 cpm in at least three samples were retained. Cpm were transformed by library size and normalized by mean variance with limma-voom. Statistics and log-ratio were calculated with limma-eBayes, by fitting data to a single-factor linear-model with three different levels (mock, EPR, EPRSTOPE).

Venn diagram and box plots. We kept differentially expressed transcripts when the observed Bayesian statistic was significant (Benjamini and Hochberg corrected *p* value <0.01; logFC > | 0.5 |). The functions *limma-vennDiagram* and *pheatmap* were used to cluster and visualize the significant genes. For box plots, summary statistics and plots were calculated and rendered with R software (version 3.5.0, https://www.R-project.org) through package ggpubr (https://CRAN.R-project.org/package=ggpubr) and ggplot2 (https://ggplot2.tidyverse.org). Data distributions and normality have been evaluated using the Shapiro−Wilk and Mann−Whitney tests for unpaired nonparametric data.

Gene ontology and pathway enrichment. Significant transcripts were summarized at gene level (*tximport-summarizeToGene*), annotated by Gene Ontology and enriched by statistically over-represented term with the *EnrichR* web-application using a nominal *p* value (Student's *t* test) threshold of $p < 0.01$. The *EnrichR p* values refer to the Fisher Exact Test statistics, which is a proportion test that assumes a binomial distribution and independence for probability of any gene belonging to any set.

Protein alignment. Multiple alignment of mammalian EPR sequences was conducted by using the ClustalW2 package (https://www.ebi.ac.uk).

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Raw data from RNA deep-sequencing analyses have been published on the GEO archive under the accession GSE113178. Human EPR expression in different subpopulations of FACS-sorted normal breast cells[21] and in different human organs was inferred through either the Expression Atlas (https://www.ebi.ac.uk/gxa/home) or the GEPIA web server (http://gepia.cancer-pku.cn). Proteins interacting with EPRp were unambiguously identified by searching a comprehensive nonredundant protein database of the National Center for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov/) and the Mass Spectrometry protein sequence DataBase (MSDB, http://msdn.microsoft.com/en-us/library/ms187112.aspx).

Received: 19 May 2018 Accepted: 27 March 2019
Published online: 29 April 2019

References

1. Jandura, A. & Krause, H. M. The new RNA world: growing evidence for long noncoding RNA functionality. *Trends Genet.* **33**, 665–676 (2017).
2. Kopp, F. & Mendell, J. T. Functional classification and experimental dissection of long noncoding RNAs. *Cell* **172**, 393–407 (2018).
3. Schmitt, A. M. & Chang, H. Y. Long noncoding RNAs in cancer pathways. *Cancer Cell* **29**, 452–463 (2016).
4. Arun, G., Diermeier, S. D. & Spector, D. L. Therapeutic targeting of long non-coding RNAs in cancer. *Trends Mol. Med.* **24**, 257–277 (2018).

5. Long, Y., Wang, X., Youmans, D. T. & Cech, T. R. How do lncRNAs regulate transcription? *Sci. Adv.* **3**, eaao2110 (2017).
6. Gupta, R. A. et al. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* **464**, 1071–1076 (2010).
7. Yoon, J. H. et al. LincRNA-p21 suppresses target mRNA translation. *Mol. Cell* **47**, 648–655 (2012).
8. Cao, L., Zhang, P., Li, J. & Wu, M. LAST, a c-Myc-inducible long noncoding RNA, cooperates with CNBP to promote CCND1 mRNA stability in human cells. *Elife* **6**, e30433 (2017).
9. Geisler, S. & Coller, J. RNA in unexpected places: long non-coding RNA functions in diverse cellular contexts. *Nat. Rev. Mol. Cell. Biol.* **14**, 699–712 (2013).
10. Yuan, J. H. et al. A long noncoding RNA activated by TGF-β promotes the invasion-metastasis cascade in hepatocellular carcinoma. *Cancer Cell* **25**, 666–681 (2014).
11. Dhamija, S. & Diederichs, S. From junk to master regulators of invasion: lncRNA functions in migration, EMT and metastasis. *Int. J. Cancer* **139**, 269–280 (2016).
12. Shi, Y. & Massague, J. Mechanisms of TGF-beta signaling from cell membrane to the nucleus. *Cell* **113**, 685–700 (2003).
13. Budi, E. H., Duan, D. & Derynck, R. Transforming growth factor-β receptors and Smads: regulatory complexity and functional versatility. *Trends Cell Biol.* **7**, 658–672 (2017).
14. Briata, P. et al. Diverse roles of the nucleic acid-binding protein KHSRP in cell differentiation and disease. *Wiley Interdiscip. Rev. RNA* **7**, 227–240 (2016).
15. Puppo, M. et al. miRNA-mediated KHSRP silencing rewires distinct post-transcriptional programs during TGF-β-induced epithelial-to-mesenchymal transition. *Cell Rep.* **16**, 967–978 (2016).
16. Moshiri, A., Puppo, M., Rossi, M., Gherzi, R. & Briata, P. Resveratrol limits epithelial to mesenchymal transition through modulation of KHSRP/hnRNPA1-dependent alternative splicing in mammary gland cells. *Biochim. Biophys. Acta* **1860**, 291–298 (2017).
17. Giovarelli, M. et al. H19 long noncoding RNA controls the mRNA decay promoting function of KSRP. *Proc. Natl. Acad. Sci. USA* **111**, E5023–E5502 (2014).
18. Inman, G. J. et al. SB-431542 is a potent and specific inhibitor of transforming growth factor-beta superfamily type I activin receptor-like kinase (ALK) receptors ALK4, ALK5, and ALK7. *Mol. Pharmacol.* **62**, 65–74 (2002).
19. Hill, C. S. Transcriptional control by the SMADs. *Cold Spring Harb. Perspect. Biol.* **8**, a022079 (2016).
20. Shirakihara, T., Saitoh, M. & Miyazono, K. Differential regulation of epithelial and mesenchymal markers by deltaEF1 proteins in epithelial mesenchymal transition induced by TGF-beta. *Mol. Biol. Cell* **18**, 3533–3544 (2007).
21. Pellacani, D. et al. Analysis of normal human mammary epigenomes reveals cell-specific active enhancer states and associated transcription factor networks. *Cell Rep.* **17**, 2060–2074 (2016).
22. Matsumoto, A. et al. (2017). mTORC1 and muscle regeneration are regulated by the LINC00961-encoded SPAR polypeptide. *Nature* **541**, 228–232 (2017).
23. Huang, J. Z. et al. A peptide encoded by a putative lncRNA HOXB-AS3 suppresses colon cancer growth. *Mol. Cell* **68**, 171–184.e6 (2017).
24. Andrews, S. J. & Rothnagel, J. A. Emerging evidence for functional peptides encoded by short open reading frames. *Nat. Rev. Genet.* **15**, 193–204 (2014).
25. Makarewich, C. A. & Olson, E. N. Mining for micropeptides. *Trends Cell Biol.* **27**, 685–696 (2017).
26. Abbas, T. & Dutta, A. p21 in cancer: intricate networks and multiple activities. *Nat. Rev. Cancer* **9**, 400–414 (2009).
27. Siegel, P. M. & Massagué, J. Cytostatic and apoptotic actions of TGF-beta in homeostasis and cancer. *Nat. Rev. Cancer* **3**, 807–821 (2003).
28. Seoane, J., Le, H. V., Shen, L., Anderson, S. A. & Massagué, J. Integration of Smad and forkhead pathways in the control of neuroepithelial and glioblastoma cell proliferation. *Cell* **117**, 211–223 (2004).
29. Pardali, K., Kowanetz, M., Heldin, C. H. & Moustakas, A. Smad pathway-specific transcriptional regulation of the cell cycle inhibitor p21(WAF1/Cip1). *J. Cell. Physiol.* **204**, 260–272 (2005).
30. Komili, S. & Silver, P. A. Coupling and coordination in gene expression processes: a systems biology view. *Nat. Rev. Genet.* **9**, 38–48 (2008).
31. Braun, K. A. & Young, E. T. Coupling mRNA synthesis and decay. *Mol. Cell. Biol.* **34**, 4078–4087 (2014).
32. Li, X. L. et al. Long noncoding RNA PURPL suppresses basal p53 levels and promotes tumorigenicity in colorectal cancer. *Cell Rep.* **20**, 2408–2423 (2017).
33. Reguly, T. & Wrana, J. L. In or out? The dynamics of Smad nucleocytoplasmic shuttling. *Trends Cell Biol.* **13**, 216–220 (2003).
34. Schmierer, B. & Hill, C. S. Kinetic analysis of Smad nucleocytoplasmic shuttling reveals a mechanism for transforming growth factor beta-dependent nuclear accumulation of Smads. *Mol. Cell. Biol.* **25**, 9845–9858 (2005).
35. Briata, P., Chen, C.-Y., Ramos, A. & Gherzi, R. Functional and molecular insights into KSRP function in mRNA decay. *Biochim. Biophys. Acta* **1829**, 689–694 (2013).
36. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70 (2012).
37. Bianchini, G., Balko, J. M., Mayer, I. A., Sanders, M. E. & Gianni, L. Triple-negative breast cancer: challenges and opportunities of a heterogeneous disease. *Nat. Rev. Clin. Oncol.* **13**, 674–690 (2016).
38. Gondran, P., Amiot, F., Weil, D. & Dautry, F. Accumulation of mature mRNA in the nuclear fraction of mammalian cells. *FEBS Lett.* **458**, 324–328 (1999).
39. Bahar Halpern, K. et al. Nuclear retention of mRNA in mammalian tissues. *Cell Rep.* **13**, 2653–2662 (2015).
40. Singer, S. et al. Nuclear pore component Nup98 is a potential tumor suppressor and regulates posttranscriptional expression of select p53 target genes. *Mol. Cell* **48**, 799–810 (2012).
41. Sharma, S. et al. Acetylation-dependent control of global Poly(A) RNA degradation by CBP/p300 and HDAC1/2. *Mol. Cell* **63**, 927–938 (2016).
42. Bierie, B. & Moses, H. L. Tumour microenvironment: TGFbeta: the molecular Jekyll and Hyde of cancer. *Nat. Rev. Cancer* **6**, 506–520 (2006).
43. Morikawa, M., Derynck, R. & Miyazono, K. TGF-β and the TGF-β family: context-dependent roles in cell and tissue physiology. *Cold Spring Harb. Perspect. Biol.* **8**, a021873 (2016).
44. Anderson, D. M. et al. A micropeptide encoded by a putative long noncoding RNA regulates muscle performance. *Cell* **160**, 595–606 (2015).
45. Nelson, B. R. et al. A peptide encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle. *Science* **351**, 271–275 (2016).
46. Yu, X. et al. Long non-coding RNA Linc-RAM enhances myogenic differentiation by interacting with MyoD. *Nat. Commun.* **8**, 14016 (2017).
47. Spadaro, D. et al. Tension-dependent stretching activates ZO-1 to control the junctional localization of its interactors. *Curr. Biol.* **27**, 3783–3795 (2017).
48. Trabucchi, M. et al. The RNA-binding protein KSRP promotes the biogenesis of a subset of microRNAs. *Nature* **459**, 1010–1014 (2009).
49. Briata, P. et al. p38-dependent phosphorylation of the mRNA decay-promoting factor KSRP controls the stability of select myogenic transcripts. *Mol. Cell* **20**, 891–903 (2005).
50. 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**, 1236–1248 (2000).
51. Gagnon, K. T., Li, L., Janowski, B. A. & Corey, D. R. Analysis of nuclear RNA interference in human cells by subcellular fractionation and Argonaute loading. *Nat. Protoc.* **9**, 2045–2060 (2014).
52. Shevchenko, A., Wilm, M., Vorm, O. & Mann, M. Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. *Anal. Chem.* **68**, 850–858 (1996).
53. Nesvizhskii, A. I., Keller, A., Kolker, E. & Aebersold, R. A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* **75**, 4646–4658 (2003).
54. Chu, C., Quinn, J. & Chang, H. Y. Chromatin isolation by RNA purification (ChIRP). *J. Vis. Exp.* **25**, 3912 (2012).
55. Ghisletti, S. et al. Identification and characterization of enhancers controlling the inflammatory gene expression program in macrophages. *Immunity* **32**, 317–328 (2010).
56. Zaccara, S. et al. p53-directed translational control can shape and expand the universe of p53 target genes. *Cell Death Differ.* **21**, 1522–1534 (2014).
57. Borowicz, S. et al. The soft agar colony formation assay. *J. Vis. Exp.* **92**, e51998 (2014).

Acknowledgements

We thank Pier Lorenzo Puri (SBP Medical Discovery Institute, La Jolla, CA) for discussions, Angela Cattaneo and Angela Bachi for discussions related to the MS analysis of EPRp interactors performed at the Proteomic Unit of IFOM (Milano), Sandro Poggi, Francesco Alessandrini, Federico Villa and Aldo Pagano (IRCCS San Martino) for sharing reagents and expertise, Claudia Bagni (University of Leuven), Paola Menichini and Gilberto Fronza (IRCCS San Martino) for sharing cells and reagents. This project has been supported, in part, by grants from the Associazione Italiana per la Ricerca sul Cancro (AIRC I.G. grant 10090 and 21541) to R.G., the Italian Ministry of Health with 5 × 1000 funds 2013 to P.B., and the Italian Ministry of Health with 5 × 1000 funds 2014 and 2015 to R.G. The work of G.B. was supported by the Italian Ministry of Health with 5 × 1000 funds. We wish to thank Michael Pancher of the High Throughput Screening facility, CIBIO, for assistance with high-content proliferation assay and Isabella Pesce of the Cell Analysis & Separation facility for assistance with the cell cycle analyses. We are grateful to people at TIB Mol.Biol. (Genova branch) for their professional assistance.

Author contributions

P.B. and R.G. conceived the study and performed some experiments. M.R. performed most experiments. G.B. performed the bioinformatics analysis of the RNA-Seq data. D.R. performed polysome profiling, analyses of cell cycle distribution, and high-content image analysis. M.J.M. analyzed h.EPR expression in human samples. D.B. performed

bioinformatics analyses related to EPR peptide. A.F. and D.S. performed immunofluorescence experiments. M.P. produced the recombinant peptide EPRp and some recombinant plasmids. L.E. and M.C. performed the orthotopic transplants in mice. F.N., S. C. and A.I. analyzed and discussed the data obtained in their laboratories with P.B. and R.G. P.B. and R.G. wrote the manuscript. A.I. discussed the revision strategy with P.B. and R.G.

Additional information

Supplementary Information accompanies this paper at https://doi.org/10.1038/s41467-019-09754-1.

Competing interests: The authors declare no competing interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Journal peer review information: *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.