GC content shapes mRNA decay and storage in human cells

Maïté Courel[1], Yves Clément[2], Dominika Foretek[3], Olivia Vidal[4], Zhou Yi[5], Michel Kress[1], Caroline Vindry[6], Marianne Bénard[1], Clémentine Bossevain[1], Christophe Antoniewski[7], Antonin Morillon[3], Patrick Brest[4], Arnaud Hubstenberger[5], Hugues Roest Crollius[2], Nancy Standart[6], Dominique Weil[1*]

[1] Sorbonne Université, CNRS, Institut de Biologie Paris-Seine (IBPS), Laboratoire de Biologie du Développement, F-75005 Paris, France.
[2] Ecole Normale Supérieure, Institut de Biologie de l'ENS, IBENS, Paris F-75005, France.
[3] ncRNA, epigenetic and genome fluidity, Institut Curie, PSL Research University, CNRS UMR 3244, Université Pierre et Marie Curie, 75248 Paris Cedex 05, France
[4] Université Côte d'Azur, CNRS, INSERM, IRCAN, FHU-OncoAge, F-06107 Nice, France
[5] Université Côte d'Azur, CNRS, INSERM, iBV, F-06107 Nice, France.
[6] Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, UK.
[7] Sorbonne Universités, CNRS, Institut de Biologie Paris Seine (IBPS), ARTbio Bioinformatics Analysis Facility, F-75005 Paris, France.

[*] Corresponding author

Summary

The control of protein expression results from the fine tuning of mRNA synthesis, decay and translation. These processes, which are controlled by a large number of RNA binding proteins and by localization in RNP granules such as P-bodies, appear often intimately linked although the rules of this interplay are not well understood. In this study, we combined our recent P-body transcriptome with various transcriptomes after silencing broadly acting mRNA decay and repression factors. This analysis demonstrated the existence of two mRNA decay pathways in human cells, which apply to distinct mRNA subsets. It also revealed the central role of GC content in mRNA fate, in terms of mRNA decay, mRNA translation and P-body localization. Furthermore, sequence-specific RBPs and miRNAs had only modest additional effects on their bulk targets. Finally, the different GC content biases associated with the two mRNA decay pathways and with mRNA translation repression had a strong impact on codon usage and protein yield. Altogether, these results lead to an integrated view of post-transcriptional control in human cells where regulation at the level of translation is dedicated to AU-rich mRNAs, which have a limiting protein yield, whereas regulation at the level of 5' decay is for GC-rich mRNAs, whose translation is optimal.

Introduction

A large part of eukaryotic genomes encode RNA-binding proteins (RBPs). In human, it amounts to more than 8% of the proteome (Hentze et al., 2018). RBPs control all aspects of RNA life and activity, both in the nucleus and in the cytosol. In the cytosol, they regulate mRNA fate in terms of mRNA translation, storage, localization and decay, processes which are often closely coupled. These numerous RBPs have to act in a coordinated manner to produce a transcriptome both coherent with cellular physiology and responsive to new cellular needs. While this is far from being understood at the transcriptome-wide level, the expanding knowledge of RBP RNA targets provides the groundwork to progress on these questions.

Some RBPs control the fate of all transcripts harboring a specific RNA motif. For instance, CPEB1, by binding the CPE motif in the 3' untranslated region (UTR) of maternal transcripts, controls their storage in *Xenopus* oocytes and their translational activation upon hormone stimulation (Standart and Minshall, 2008). Additional examples include the proteins which bind 3'UTR AU-rich elements (ARE), such as HuR and TTP, to control translation and decay, and play key roles in the inflammation response, apoptosis and cancer (Wells et al., 2017). RNA binding motifs are generally not unique but are rather defined as a consensus binding motif. In the case of RISC, the binding specificity is given by a guide miRNA, which also hybridizes with some flexibility with complementary sequences. A variety of techniques have therefore been developed to identify the effective RNA targets of such factors, ranging from affinity purification (such as RIP or CLIP) to transcriptome and polysome profiling after RBP silencing.

Other RBPs do not have any obvious sequence specificity. They can nevertheless be recruited to a specific motif by a sequence-specific cofactor. This is the case for the RNA helicase DDX6, which binds CPE-containing mRNAs when guided by the CPEB complex in oocytes (Minshall et al., 2007). However, DDX6 also seems to have a generalist activity. Human DDX6 interacts with both translational repressors, and the decapping enzyme DCP1/2 and its activators (Ayache et al., 2015; Bish et al., 2015). Its yeast homologue dhh1 is a cofactor of DCP2, as well as a translational repressor (Coller and Parker, 2005). PAT1B has also been defined as an enhancer of decapping, as it interacts with DDX6 and with the decapping complex in mammalian cells (Vindry et al., 2017), while in yeast Pat1p activates Dcp2 directly (Nissan et al., 2010) and its deletion results in deadenylated but capped intact mRNA (Bonnerot et al., 2000; Bouveret et al., 2000). DDX6 and PAT1B also interact with the CCR4-NOT deadenylase complex and the DDX6-CNOT1 interaction is required for miRNA silencing (Chen et al., 2014; Mathys et al., 2014; Ozgur et al., 2015; Vindry et al., 2017). DDX6 also binds 4E-T directly, which in turn interacts with the cap-binding factor eIF4E and inhibits translation initiation, including that of miRNA target mRNAs (Kamenska et al., 2016). Altogether, DDX6 and PAT1B have been proposed to link deadenylation/translational repression with decapping. Finally, the 5'-3'exonuclease XRN1 decays RNAs following decapping by DCP1/2, a step triggered by deadenylation mediated by PAN2/3 and CCR4-NOT or by exosome activity (Łabno et al., 2016).

mRNA fate is also intimately linked with their localization in P-bodies (PBs). In mammalian cells, DDX6 is a key factor for PB assembly (Minshall et al., 2009). We recently identified the transcriptome and proteome of PBs purified from human cells. Their analysis showed that human PBs are broadly involved in mRNA storage rather than decay (Hubstenberger et al., 2017; Standart and Weil, 2018), as also observed using fluorescent decay reporters (Horvathova et al., 2017). While a general view of decay and translation mechanisms is progressively emerging, the connection between the above general repression/decay factors and the numerous sequence-specific repression/decay regulators (ARE-binding proteins, CPEBs, FRX proteins, proteins recognizing TOP and G4 motifs, etc.) is still not well understood.

In this study, we addressed the question of the general status of mRNA decay and storage in unstressed cells, using several high throughput experiments performed after silencing of general storage and decay factors: a polysome profiling after DDX6 silencing, a transcriptome profiling after PAT1B silencing, two transcriptome profiling after XRN1 silencing, and the transcriptome of PB-stored mRNAs. We also used datasets available from the literature, including a transcriptome profiling after DDX6 silencing and a DDX6-CLIP experiment. Their combined analysis revealed the central role of mRNA GC content in dictating mRNA fate, contributing to the coordination between two opposite processes: decay and storage. Known binding preferences of sequence-specific RBPs and codon usage in human appear consistent with such a central role of the GC content. This general feature of post-transcriptional regulation could, we propose, underpin an integrated gene expression program.

Results

DDX6 has opposite effects on mRNA decay and translation depending on their GC content

In human epithelial cells, DDX6 is associated with proteins involved in both mRNA translation repression and mRNA decapping (Ayache et al., 2015; Bish et al., 2015), but its impact on mRNA fate remains poorly characterized. To investigate the function of DDX6 on endogenous mRNAs, we conducted a polysome profiling experiment in HEK293 cells transfected with DDX6 or control β-globin siRNAs for 48h. In these conditions, DDX6 expression decreased by 90% compared to control cells (Figure S1A). The polysome profile was largely unaffected by DDX6 silencing, implying that DDX6 depletion did not grossly disturb global translation (Figure S1B). Polysomal RNA isolated from the sucrose gradient fractions (Figure S1B) and total RNA was used to generate libraries using random hexamers to allow their amplification independently of the length of their poly(A) tail. As expected, DDX6 mRNA itself was markedly decreased (by 72%) in total and polysomal mRNA (Figure S1C-E; Table S1, sheet1). A pie chart representation of mRNA fold-changes showed that the mRNAs most sensitive to DDX6 knock-down (threshold: log2 FC=+/-0.25) were affected either in polysomes (55%) or in total RNA (28%), but rarely in both (13%) (Figure 1A). As some translation changes can be masked by simultaneous changes in polysomal and total mRNA, we used in the following the polysomal to total mRNA ratio as a proxy measurement of translation rate for the whole dataset. Moreover, since DDX6 is cytoplasmic (Ernoult-Lange et al., 2009), we assumed that mRNA accumulation after DDX6 knock-down generally reflects an increased stability. Consistent with the pie chart, mRNAs with the most upregulated translation rate were the least stabilized, and reciprocally (Figure 1B).

This opposite impact of DDX6 on mRNA translation and decay prompted us to search for determinants of the DDX6 dichotomic effect. We first subdivided mRNAs into six classes depending on their size. The mRNAs under 1.5 kb tended to be more stabilized after DDX6 knock-down than the longest (Figure S2A, left panel, and S2B), whereas the mRNAs over 5 kb were more affected in terms of translation (Figure S2C, left panel, and S2B). Of note, these weak length dependencies were a combined effect of CDS and 3'UTR length for mRNA stability (Figure S2A, middle and right panels, and S2B), but mostly of 3'UTR length for translation (Figure S2C, middle and right panels, and S2B).

In contrast, we observed a striking dependency on mRNA GC content. When mRNAs were subdivided into six classes ranging from <40% to >60% GC, the extent of mRNA stabilization steadily increased with the GC content and became predominant for transcripts with >50% GC (Figures 1C, left panel, and S2D). This analysis was repeated on an independent dataset available from the ENCODE project (ENCODE Project Consortium, 2012), obtained in a human erythroid cell line, K562, following induction of a stably transfected DDX6 shRNA, and using an oligo(dT)-primed library). Despite the differences in cell type, and depletion and sequencing methods, mRNA stabilization was again restricted to those with high GC content (Figures 1C, right panel, and S2D; Table S1, sheet2). In contrast, the mRNAs

regulated by DDX6 at the level of translation in HEK293 cells were particularly AU-rich (Figure 1D). The GC content of both the CDS and the 3'UTR was determinant, rather than the 5'UTR (Figure 1E and S2E,F).

To investigate the relationships between the DDX6 effects on mRNA decay and translation and DDX6 binding to mRNA, we used the CLIP dataset of K562 cells, also available from the ENCODE project. In both HEK293 and K562 cells, the mRNAs clipped to DDX6 were particularly stabilized after DDX6 knockdown, as compared to all mRNAs (Figure S2G; Table S1, sheet3). However, this was not the case for the mRNAs regulated at the level of translation (Figure S2H). In agreement, mRNAs with a high GC-rich content were strongly enriched in the DDX6 CLIP experiment (Figures 1F and S2D).

In conclusion, the dichotomic effect of DDX6 knockdown on mRNA decay and translation was related to their GC content, with the most GC-rich mRNAs being regulated at the level of decay and the most AU-rich mRNAs at the level of translation.

XRN1 and PAT1B targets separate sub-classes of mRNAs depending on their GC content

We then investigated the relationship between DDX6-dependent mRNA decay and mRNA decay by the XRN1 5' to 3' exonuclease. To this aim, we performed two XRN1 silencing experiments in human HeLa and HCT116 cells. HeLa cells were transfected with XRN1 siRNA for 48h (Figure S3A; Table S1, sheet4), while HCT116 cells stably transfected with an inducible XRN1 shRNA were induced with doxycyclin for 48h (Figure S3B; Table S1, sheet5). Like DDX6-dependent decay, in both cell lines, XRN1-dependent decay was strongly linked to mRNA GC content and restricted to GC-rich mRNAs (Figures 2A and S2D). Accordingly, 65% and 67% of the transcripts stabilized after siDDX6 in HEK293 and K562 cells, respectively, were also stabilized after siXRN1 in HeLa cells, and this proportion increased to 72% and 76% in HCT116 cells (Figure S3C). Therefore, XRN1 tends to decay the same mRNAs as DDX6. In agreement, 62% and 65% of the transcripts enriched in the DDX6 CLIP were stabilized after XRN1 silencing in HeLa and HCT116 cells, respectively (Figure S3C).

PAT1B is a well-characterized direct DDX6 partner known for its involvement in mRNA decay (Braun et al., 2010; Marnef and Standart, 2010; Ozgur et al., 2010; Vindry et al., 2017). However, using our previous PAT1B silencing experiment in HEK293 cells (Vindry et al., 2017), we surprisingly found a negative correlation between mRNA stabilization after PAT1B and after DDX6 silencing (Figure S3D; Table S1, sheet6), suggesting that they largely target separate sets of mRNAs. Accordingly, mRNAs enriched in the DDX6 CLIP experiment were not stabilized after PAT1B silencing (Figure S3E). However, the correlation was positive between mRNA stabilization after siPAT1B and translational derepression after siDDX6 (Figure S3F), indicating that PAT1B targets mRNAs that are translationally repressed by DDX6. Indeed, in contrast to DDX6 and XRN1 decay targets, the mRNAs subject to PAT1B-dependent decay were strikingly AU-rich (Figures 2B and S2D).

PAT1B has no known enzymatic activity and could be involved in 5' or 3' mRNA decay. To gain insight into this issue, we analyzed the read coverage of PAT1B targets. These transcripts had a higher 5' coverage than non-target mRNAs, in the presence or absence of PAT1B (Figure 2C). In contrast, XRN1 targets had a lower 5' coverage than non-targets prior to silencing (Figure 2D, left panel), and this difference of coverage was suppressed following XRN1 depletion, as expected from a 5' decay pathway (Figure 2D, right panel). These results suggest that PAT1B-dependent mRNA stabilization does not result from a better protection of their 5' extremity.

Altogether these results indicate that (i) PAT1B is involved in a distinct decay pathway to that mediated by DDX6 and XRN1, (ii) these two decay pathways act on separate mRNA subsets which strongly differ

in their GC content, and (iii) the PAT1B-dependent decay pathway acts on the mRNAs that undergo DDX6-dependent translation repression and (iv) likely relies on 3' to 5' degradation.

PBs only accumulate AU-rich mRNAs

In terms of cellular distribution, DDX6 is a central component of PBs, which store untranslated mRNAs in the cytoplasm without decay, while the mRNAs stabilized after DDX6 silencing are localized outside PBs (Hubstenberger et al., 2017). Similarly, there was a negative correlation between mRNA stabilization after XRN1 depletion in HeLa cells and PB enrichment (Figure S4A). In contrast, as we reported previously, transcripts stabilized after PAT1B silencing are localized in PBs (Vindry et al., 2017). In agreement with these results and the impact of the GC content described above, mRNA accumulation in PBs was strongly biased towards AU-rich mRNAs (Figures 3A). While reminiscent of the low GC content reported for stress granule mRNAs in HEK293 cells (Khong et al., 2017), the correlation was much higher with PB localization (correlation coefficient r=0.66) than with stress granule localization (r=0.13) (Figure S4B). Indeed, comparing the GC content distribution of PB-enriched and -excluded mRNAs to all transcripts expressed in HEK293 cells, revealed that mRNA storage in PBs is confined to the AU-rich fraction of human genes (Figure 3B).

This raised the possibility that the impact of GC content on PB enrichment was due to the genomic context of the genes. To address this issue, we looked at the link between PB enrichment and meiotic recombination, which can influence GC content through GC-biased gene conversion (Duret and Galtier, 2009). The correlation between PB enrichment and meiotic recombination was weaker than between PB enrichment and mRNA GC content (Spearman r = -0.16 vs -0.66, both p-values < 0.05). Moreover, the correlation between PB enrichment and mRNA GC content was virtually unchanged when controlling for meiotic recombination (Spearman r = -0.65, p-value < 0.05). This correlation was still significant when controlling for intronic or flanking GC content (Spearman r = -0.329 and -0.449 respectively, all p-values < 0.05), showing that mRNA base composition and P-body enrichment are associated independently of meiotic recombination or the genomic context.

As transcript length is a key determinant of mRNA accumulation in stress granules (Khong et al., 2017), we also analyzed its importance for mRNA accumulation in PBs. PB-enriched mRNAs were longer than PB-excluded mRNAs, independently of their GC content (Figures 3C, S4C). However, their increased length (4.4 kb on average) was less striking than previously observed in stress granules (7.1 kb on average) (Khong et al., 2017).

PB-mRNAs are exclusively AU-rich, and they tend to be longer than average. Interestingly, compared to AU-rich transcripts in PBs, AU-rich cytosolic transcripts were more sensitive to XRN1-dependent decay (Figure 3D) but not to DDX6-dependent decay (Figure 3E). This suggested that XRN1 preference for GC-rich mRNAs is at least in part related to their exclusion from PBs, whereas DDX6 has a true preference for GC-rich mRNAs. To recapitulate, GC-rich mRNAs are regulated by DDX6 and XRN1 in terms of decay and excluded from PBs, whereas AU-rich mRNAs are regulated by PAT1B in terms of decay and by DDX6 in terms of translation and are recruited to PBs (Figure S4D).

Specific decay pathways target GC-rich mRNAs

Having shown that GC content is key to DDX6- and XRN1- versus PAT1B-dependent mRNA decay, we investigated the link between this global sequence determinant and several sequence-specific post-transcriptional regulators, for which relevant genome-wide datasets are available.

We first considered the regulation of mRNA decay. It was previously shown in yeast that mRNAs accumulate in polysomes when translation elongation is slowed-down by the presence of suboptimal codons, and recruit Dhh1/DDX6 to induce their decay (Radhakrishnan et al., 2016). In HEK293 cells, we

observed only a weak relationship between polysome engagement (defined as the fraction of total mRNA present in polysomes) and DDX6-dependent decay (Figure S5A), suggesting that, if this mechanism is conserved in human, it accounts for a minor part of DDX6-dependent decay, relative to the impact of GC content. Next, we considered the Non-sense Mediated Decay (NMD) and the m^6A-associated decay pathways. We took as NMD targets the mRNAs cleaved by SMG6 (Schmidt et al., 2015), and for m^6A the targets of the decay-associated YTHDF2 reader, as defined by CLIP (Wang et al., 2013; Yang et al., 2015). We also analyzed the effect of G4 motifs, which have been shown to be preferential substrates of murine XRN1 in vitro (Bashkirov et al., 1997). TOP mRNAs, which are strongly excluded from PBs (Hubstenberger et al., 2017), were included in the analysis for comparison.

SMG6 and YTHDF2 targets as well as mRNAs with G4 motifs tended to be dependent on DDX6 for decay in both HEK293 and K562 cells (Figure 4A), while they were independent of PAT1B for decay (Figure S5B) and of DDX6 for translation (Figure S5C). Consistent with the global analysis, they also tended to be excluded from PBs, particularly SMG6 targets (Figure 4B). However, surprisingly, within PB-excluded mRNAs, there was little or no additional effect of being a SMG6 target, an YTHDF2 target or containing a G4 motif, neither for DDX6-sensitive decay (Figure 4C) nor for XRN1-dependent decay (Figures S5D-E). While a tempting hypothesis, exclusion from PBs could not explain all DDX6- and XRN1-dependent decay. Indeed, TOP mRNAs, which are strongly excluded from PBs (Figure 4E), were unaffected by DDX6 (Figures 4A) or XRN1 silencing (Figure S5D). We therefore considered the GC content of these transcripts. In fact, SMG6 and YTHDF2 targets, and mRNAs containing G4 motifs, tended to be GC-rich, and their DDX6-dependency was in the same range as for any other mRNAs of similar GC content (Figure 4D). In terms of PB localization, only SMG6 targets were more excluded than expected from their GC content (Figure 4E). In conclusion, the dominant parameter for DDX6-dependent decay in steady-state conditions was mRNA GC content, with the regulatory proteins SMG6 and YTHDF2 and G4 motifs making subsidiary contributions.

Translation repressors target AU-rich mRNAs

We then searched for sequence-specific translation regulators associated with PB localization and PAT1B-dependent decay. As reported previously, ARE-containing mRNAs and the targets of two ARE-binding proteins (ARE-BPs), HuR and TTP, were stabilized after PAT1B silencing (Figure 5A) (Vindry et al., 2017). These mRNAs were also enriched in PBs (Figures 5B and S6A), and translationally more active after DDX6 silencing (Figure 5C), which is consistent with the presence of HuR and TTP proteins in PBs (Franks and Lykke-Andersen, 2007; Hubstenberger et al., 2017). In agreement, their stability was not increased upon DDX6 silencing (Figure S6B). We also analyzed mRNAs with a CPE motif, as DDX6 is a component of the CPEB complex that binds CPEs in *Xenopus* oocytes (Minshall et al., 2007), as well as targets of several DDX6 partners and/or PB proteins: FXR1-2, FMR1, PUM1-2, IGF2BP1-3, the helicase MOV10, ATXN2 and 4E-T (Figure S6A) (Ayache et al., 2015; Hubstenberger et al., 2017). These targets were defined in a variety of human cell lines, using approaches including RIP and CLIP (see Methods). As summarized in Figure 5D, CPE-containing mRNAs and the targets of FXR1-2, FMR1, PUM1-2, IGF2BP1-3 and MOV10 behaved like ARE-BP targets, with a PAT1B-dependent stability and a DDX6-dependent translation, associated with PB localization (Figures 5A-C and S6B).

Two major DDX6 partners (Ayache et al., 2015) showed a distinct pattern, ATXN2 and 4E-T. ATXN2 targets were excluded from PBs (Figure 5B), which is consistent with the absence of ATXN2 from PBs (Ayache et al., 2015; Nonhoff et al., 2007). They were unaffected by either PAT1B or DDX6 silencing (Figure 5A,C and S6B), leaving unresolved the function of the ATXN2/DDX6 interaction in the cytosol. The translation repressor 4E-T is required for PB assembly (Ayache et al., 2015; Kamenska et al., 2016). Interestingly, its targets were markedly enriched in PBs (Figure 5B), though poorly affected by PAT1B silencing in terms of stability, or by DDX6 silencing in terms of translation (Figure 5A,C). This

dissociation between PB localization and mRNA fate indicated that PB recruitment is not sufficient to determine PAT1B-dependent decay and DDX6-dependent translation repression. It also pointed to a particular role of 4E-T in PB targeting or scaffolding.

As observed for mRNA stability, a large part of these results was attributable to the GC content, except that the targets of all tested translation repressors (though not ATXN2) were AU-rich (Figure 5E-G). Nevertheless, in the case of HuR, TTP, FXR1-2, FMR1, PUM2, IGF2BP1-3 and MOV10, there was some additional PAT1B-sensitivity and PB enrichment. In terms of translation repression, only FXR1 targets were more sensitive to DDX6 level than expected from their GC content. Thus, as observed for decay regulators, translation regulators appeared to make only subsidiary contribution to DDX6-dependent repression, as compared to the GC content.

PB localization of miRNA targets reflects their GC content

The miRNA pathway, which leads to both translation repression and mRNA decay, has been previously associated with DDX6 activity and PB localization ((Bhattacharyya et al., 2006; Chu and Rana, 2006). To study this pathway, we first used the list of AGO1-4 targets, as identified in CLIP experiments (Yang et al., 2015) (Figure S7A). The translation of these mRNAs was DDX6-dependent (Figure 6A; note that the number of AGO4 targets was too small to reach statistical significance) and they tended to accumulate in PBs (Figure 6B). In terms of stability, only AGO2 targets were marginally DDX6-dependent (Figure S7B), and the effects were not stronger when analyzing separately the mRNAs enriched or excluded from PBs (Figure S7D). In contrast, their decay was PAT1B-dependent (Figure 6C).

As observed for translation repressors, the AGO targets were rather AU-rich, which explained a large part of their behavior after DDX6 and PAT1B silencing, as well as their enrichment in PBs (Figures 6D-G). There was nevertheless an additional effect of being an AGO1-3 target for PB localization and PAT1B–dependent decay (Figures 6D-E). Regarding DDX6 dependency, the translation repression of AGO1-4 targets was not stronger than expected from their GC content (Figure 6F), but their decay was slightly higher (Figure 6G), which would have been missed without taking into account mRNA GC content.

We then considered the experimentally documented targets of the 22 most abundant miRNAs in HEK293 cells (19 from (Hafner et al., 2010), and 3 additional ones from our own quantitation, Figure S7C), as described in miRTaBase (Hsu et al., 2014). Altogether, they were less affected in all datasets than AGO targets (Figures 6A-C and S7B,C). However, the separate analysis of the targets of the 22 miRNAs revealed that the extent of PB storage depended on the miRNA (Figure S7E). Again, this was associated with various GC content (Figure S7F). At the two extremes, miR21-5p targets were particularly AU-rich and strongly enriched in PBs, while the targets of miR-99b-5p, the most GC-rich in this set, were not. The results were similar when analyzing PAT1B- and DDX6-dependency (Figures 6H-K). There were nevertheless additional and specific effects of some miRNAs, as illustrated by miR-101-3p and miR-21-5p (Figure 6H,I): while their targets were particularly enriched in PBs, only miR-101-3p targets were particularly dependent on PAT1B for decay. In terms of DDX6 dependency, the translation repression followed the one expected from the GC content of all miRNA targets (Figure 6J), while decay was slightly higher than expected for the targets of several miRNAs, including in particular miR-99b-5p (Figure 6K), whose targets are not enriched in PBs (Figure S7E).

In conclusion, the behavior of miRNA targets was primarily related to their GC content. Nevertheless, individual miRNAs could also specifically strengthen PAT1B or DDX6 dependency for decay. Interestingly, the median GC content of the miRNA targets correlated with the GC content of the miRNA itself (Figure S7G). Thus, despite their small size, the miRNA binding sites tend to have a GC

content similar to that of their full-length host mRNA, which affects their fate in terms of PB localization and mechanism of post-transcriptional control.

GC bias in PBs impacts codon usage and protein yield

The strong bias of GC content in the CDS of PB-enriched and -excluded mRNAs prompted us to investigate its impact on their coding properties. We found that the frequency of amino acids encoded by GC-rich codons (Ala, Gly, Pro) was lower in PB-stored than in PB-excluded mRNAs, while the frequency of those encoded by AU-rich codons (Lys, Asn) was higher (Figure 7A).

In addition to this impact on the amino acid composition of encoded proteins, we observed dramatic differences of codon usage between the two mRNA subsets. For every amino acid encoded by synonymous codons, the relative codon usage in PBs versus out of PBs was systematically biased towards more AU-rich codons in PB mRNAs (log2 of the ratio >0, Figure 7B). Some additional codon bias independent of base composition (NNA/U or NNG/C) were also observed for 4 and 6-fold degenerated codons (Figures S8A,B). For instance, Leu was encoded twice more often by CTT than CTA in PB-enriched mRNAs, whereas the use of both codons was low in PB-excluded mRNAs. Similarly, Gly was encoded more often by GGG than GGC in PB mRNAs, whereas the use of both codons was similar in PB-excluded mRNAs (Figure 7C).

In human, the less frequent synonymous codons of an amino acid generally end with an A or U (with the exception of Thr, Ser, Pro, Ala). Consistent with the GC bias, 23 of the 29 less frequently used synonymous codons (normalized relative usage <1) were overrepresented in PB mRNAs (Figure S8C). We calculated the frequency of the less frequent codons (called rare codons thereafter) for each CDS, and plotted it as a function of the GC content at the third position (GC3) to avoid any confounding effects of the amino acid bias (Figure 7D). The frequency of rare codons correlated strongly and negatively with GC3, with AU-rich CDS having a higher frequency of rare codons than GC-rich CDS (Figure 7D). As expected, PB mRNAs had a higher frequency of rare codons than PB-excluded mRNAs. However, a similar coefficient correlation was measured for each mRNA subsets (0.87 for PB-enriched; 0.84 for PB-excluded mRNAs), meaning that the higher frequency of rare codons in PB mRNAs could be largely explained by their GC bias alone.

We previously reported that protein yield, defined as the ratio between protein and mRNA abundance in HEK293 cells, was 20-times lower for PB-enriched than PB-excluded mRNAs. This was not due to translational repression within the PBs, as the proportion of a given mRNA in PBs hardly exceeded 15%, but rather to some intrinsic mRNA property (Hubstenberger et al., 2017). Searching for such an underlying mRNA feature, we first considered the frequency of rare codons. Its correlation with the protein yield was much lower than with PB localization (r=0.21 versus 0.56). Moreover, CDS tended to be longer in PB-stored than in PB-excluded mRNAs and, interestingly, this was independent of their GC3 content (Figure 7E). This feature correlated much better with the protein yield than with PB localization (r=0.40 versus 0.16). However, combining the frequency of rare codons with the CDS length, that is, considering the absolute number of rare codons per CDS, was a shared parameter of both protein yield (r=0.43, Figure 7F) and PB localization (r=0.31). Thus, strikingly, CDS with more than 100 rare codons were mostly enriched in PBs, while those under 100 were mostly excluded (Figure 7F).

In conclusion, the high GC bias in PB mRNAs affected the amino acid composition of encoded mRNAs, as well as codon usage. Furthermore, we showed that the number of rare codons in PB mRNAs is a likely determinant of their low protein yield. We speculate that such suboptimal translation makes them optimal targets for translation regulation, since any control mechanism has to rely on a limiting step. Conversely, optimally translated transcripts would be better controlled by decay. One prediction of this model (Figure 7G) is that proteins produced in limiting amounts, such as those encoded by

haplo-insufficiency genes, are more likely to be encoded by PB mRNAs. Using haplo-insufficiency prediction scores (Huang et al., 2010; Steinberg et al., 2015), we found that haplo-insufficient mRNAs were indeed enriched in PBs (Figure 7H).

Discussion

An integrated model of post-transcriptional regulation

Our combined analysis of transcriptomes following the silencing of broadly acting storage and decay factors, including DDX6, XRN1 and PAT1B, together with the transcriptome of purified PBs, provided a general landscape of post-transcriptional regulation in human cells, where the GC content of mRNAs plays a central role. As schematized in Figure 7G, GC-rich mRNAs are excluded from PBs and mostly controlled at the level of decay, by a mechanism involving the helicase DDX6 and the 5' exonuclease XRN1. In contrast, AU-rich mRNAs are enriched in PBs and controlled at the level of translation, by a mechanism also involving DDX6, and at the level of decay by the DDX6 partner PAT1B, a decay likely to be from the 3' extremity. Accordingly, NMD and m6A-associated mRNA decay pathways tend to target GC-rich mRNAs, while sequence-specific translation regulators and miRNAs tend to target AU-rich mRNAs. The distinct fate of GC-rich and AU-rich mRNAs correlates with a contrasting protein yield resulting from the combined impact of the GC content on codon usage and on the CDS length. Thus, translation regulation is used to control mRNAs with limiting translational efficiency, which are AU-rich, whereas 5' mRNA decay is used to control mRNAs with optimal translation, which are mostly GC-rich.

It should be stressed that our analysis focused on trends common to most transcripts, which does not preclude that particular mRNAs could be exceptions to this general model, being GC-rich and translationally controlled, or AU-rich and subject to DDX6-dependent decay. Moreover, while the analysis was consistent in proliferating cells of various origins, giving rise to a general model, it is possible that changes in cell physiology at particular developmental stages or during differentiation, rely on a different mechanism.

Distinct decay pathways

Our analysis revealed the existence of separate decay pathways, depending on the GC content of mRNAs. Interestingly, our previous analysis of the read coverage of PB and non-PB RNAs also suggested the existence of distinct decay pathways related to susceptibility of 3' or 5' mRNA ends, respectively (Hubstenberger et al., 2017). Nevertheless, the targeting of mRNAs towards the DDX6/XRN1 versus the PAT1B pathway cannot entirely rely on PB exclusion of GC-rich mRNAs versus PB enrichment of AU-rich mRNAs. Indeed: (i) TOP mRNAs, which were strongly excluded from PBs, were unaffected by either DDX6 or XRN1 silencing; (ii) all PBs disappeared after DDX6 silencing (Minshall et al., 2009), which causes the release of AU-rich mRNAs into the cytosol, yet only GC-rich mRNAs were stabilized. However, only AU-rich mRNAs enriched in PBs were susceptible to PAT1B-dependent decay (not shown).

Focusing on DDX6-mediated decay, the observation that the 5'UTR GC content has little impact compared to the CDS and 3'UTR, indicates that involvement of this helicase is related to the structure of the mRNA over its entire length, rather than simply allowing XRN1 access to the 5' region. It is therefore tempting to propose that the higher dependency of GC-rich mRNAs on DDX6 for decay reflects its activity in unwinding double-stranded regions and facilitating XRN1 progression along the mRNA. UPF1, another RNA helicase involved in mRNA decay, has also been shown to preferentially affect GC-rich mRNAs (Imamachi et al., 2017). Although the bias in the case of UPF1 was restricted to the 3'UTR regions of its targets, it suggests that DDX6 could act in concert with other helicases for

mRNA decay of GC-rich mRNAs. Accordingly, targets of the NMD pathway, which involves UPF1, have GC-rich 3'UTRs (Colombo et al., 2017). For AU-rich mRNAs, either such an active unfolding would be dispensable, or it would rely on other helicases that remain to be identified, with potential candidates being those enriched in purified PBs (Hubstenberger et al., 2017).

Concerning XRN1, however, since PBs are enlarged following its silencing (Cougot, 2004), we cannot exclude that AU-rich mRNAs are also targets of XRN1, but escape its activity due to their protection in PBs. Nevertheless, there are several lines of evidence that 5' decay shows some specificity in vivo. In *Drosophila*, mutations in XRN1 have specific phenotypes, including wound healing, epithelial closure and stem cell renewal in testes, suggesting that it specifically degrades a subset of mRNAs (Pashler et al., 2016). In zebrafish, there is a broad but nevertheless specific contribution of Dcp2 to mRNA decay at the maternal to zygotic transition (Mishima and Tomari, 2017).

Turning to the function of PAT1B in mRNA decay, previous studies reported that tethered PAT1B decreases the abundance of a reporter mRNA as a result of enhanced deadenylation and decapping (Kamenska et al., 2014; Ozgur et al., 2010; Totaro et al., 2011). However, in our PAT1B silencing experiment, we did not find any evidence of PAT1B-dependent 5' to 3' decay, as the read coverage of PAT1B targets in the 5' extremity tended to be high compared to other mRNAs. We therefore favor the possibility of a PAT1B-dependent 3' to 5' decay mechanism. Such a mechanism could involve the CCR4/CNOT deadenylase and the LMS1-7 complexes, as, despite their low abundance or small size, CNOT1 and LSM2/4 had high scores in our recent PAT1B interactome analysis (Vindry et al., 2017). Interestingly, a previous study in yeast showed that two mRNA decay reporters, MFA2 and PGK, were distinctly susceptible to 5' and 3' decay, and that the 3' decay was dependent on Pat1 and Lsm1-7 (Wu et al., 2014). Along the same lines, CLIP experiments in yeast show a preference for Pat1 and Lsm1 binding to the 3' end of mRNAs (Mitchell et al., 2012).

In our previous study, we had also noted that AU-rich motifs (AREs) are overrepresented in the 3'UTR of PAT1B targets (Vindry et al., 2017). While such an enrichment could simply result from the low GC content of PAT1B targets over their entire length, with no particular link between PAT1B and ARE-BPs, many studies have demonstrated a link between AREs and mRNA stability, and its striking importance for processes such as inflammation (Wells et al., 2017). Most ARE-BPs promote mRNA destabilization while some ARE-BPs, such as HuR (Lebedeva et al., 2011; Mukherjee et al., 2011) and AUF1 for a subset of mRNAs (Yoon et al., 2014), can stabilize mRNAs. Altogether, these observations raise the possibility that ARE-BPs behave either as enhancers or inhibitors of PAT1B-dependent decay.

Translation repression and PB accumulation

DDX6 activity in translation repression has been documented in several control pathways in a variety of biological contexts. In *Xenopus,* DDX6 contributes to the repression of maternal mRNAs with a CPE motif prior to oocyte activation, as a component of the well characterized CPEB complex (Minshall et al., 2007). In *Drosophila*, Me31B/DDX6 represses the translation of thousands of mRNAs during the early stages of the maternal to zygotic transition (Wang et al., 2017). It also collaborates with FMRP and AGO proteins for translation repression in fly neurons (Barbee et al., 2006). In mammals, DDX6 is a general co-factor of the miRNA pathway (Chen et al., 2014; Chu and Rana, 2006; Kamenska et al., 2016; Mathys et al., 2014). The intriguing finding of our analysis was that the targets of all tested specific translation repressors (FRX1-2, FMR1, PUM1-2, etc.), including the targets of most miRNAs, were AU-rich. Moreover, unexpectedly, these targets had generally the same mean behavior in all datasets as other mRNAs of similar GC content. It was as if specific regulators all contributed to similar extents to the fate of most of their targets, in terms of both stability and translation repression. A provocative possibility would be that RBPs bind a large number of targets, but only affect a minority

of them, depending on other associated RBPs. Alternatively, the specific activity of these RBPs become significant only when their expression is turned on or off. In any case, the general mRNA decay and translation repression revealed by DDX6, XRN1 and PAT1B silencing seems mainly driven by their GC content rather than by specific regulators.

PBs add another layer to translation regulation, by storing untranslated mRNAs (Hubstenberger et al., 2017). In agreement with the analysis of repressor targets, they contained only AU-rich mRNAs. It was already known that ARE-containing mRNAs bound to ARE-BPs such as TTP and BRF were recruited to PBs (Franks and Lykke-Andersen, 2007) and that miRNA targets accumulate in PBs upon miRNA binding in a reversible manner (Bhattacharyya et al., 2006; Liu et al., 2005). As DDX6 is a key factor of PB assembly in mammalian cells (Ayache et al., 2015; Minshall et al., 2009), it raises the question of whether DDX6 contributes to translation repression by triggering the recruitment of arrested mRNA to PBs, or by mediating translation arrest, which then results in mRNA recruitment to PBs. The same question arises for 4E-T, a translation repressor and DDX6 partner also required for the assembly of mammalian PBs (Ayache et al., 2015; Ferraiuolo et al., 2005; Kamenska et al., 2014). While we have no answer for DDX6, we observed that 4E-T targets were particularly enriched in PBs, though rather insensitive to DDX6 or PAT1B knock-down. First, this suggests that 4E-T function in PBs is partly independent of DDX6, agreeing with the previous observation that some PBs can still form when the DDX6 interaction domain of 4E-T is mutated (Kamenska et al., 2016). Second, it indicates that PB localization and translation repression by DDX6 can be separated, at least to some extent.

In addition to their high AU content, we observed that PB mRNAs were longer than mRNAs excluded from PBs. While a longer 3'UTR would increase the probability of binding a translation repressor, and therefore favor PB recruitment, it should be noted that the CDS were also longer in PB mRNAs. This instead suggests that protein binding over the entire length of mRNA contributes to translation repression and/or PB recruitment. This would also explain why the GC content of the 5'UTR has little impact compared to CDS and 3'UTR, as leader sequences, being considerably shorter, would bind less of such proteins. In this regard, it is interesting to note that we and others have previously proposed from biochemical experiments and electron microscopy imaging that DDX6 and its partner LSM14A coat repressed mRNAs at multiple positions, according to their length (Ernoult-Lange et al., 2012; Götze et al., 2017).

GC content and codon usage

As the redundancy of the genetic code enables amino-acids to be encoded by synonymous codons of different GC content, the strong variation in mRNA GC content observed in and out of PBs has consequences on the amino-acid composition of encoded proteins. It also strongly impacts the identity of the wobble base: in PB mRNAs, the increased frequency of A/U at position 3 of the codon mechanically results in an increased usage of rare codons. As CDS are also longer in PB mRNAs, it further increases the number of rare codons per CDS in these mRNAs. These results provide a molecular mechanism to a previously unexplained feature of PB mRNAs, which is their particularly low protein yield, which we showed was an intrinsic property of these mRNAs and not simply the result of their sequestration in PBs (Hubstenberger et al., 2017). We now report that the absolute number of rare codons per CDS best correlates with low protein yield. Interestingly, the mRNAs of haplo-insufficiency genes, which by definition are expected to have a limited protein yield, are indeed enriched in PBs.

In addition to the GC-dependent codon usage, we also observed some GC-independent codon bias in PB-enriched versus PB-excluded mRNAs. Interestingly, some important post-transcriptional regulation programs involve codon usage. This was shown for proliferation- versus differentiation-specific

transcripts in human cells (Gingold et al., 2014) and for maternal mRNA clearance during the maternal-to-zygotic transition in zebrafish, *Xenopus*, mouse, and *Drosophila* (Bazzini et al., 2016). Codon usage could also enable the regulation of small subsets of mRNAs, depending on cellular requirements. In man, half of GO sets show more variability in codon usage than expected by chance (Rudolph et al., 2016). Based on GO analysis, we previously demonstrated that PB mRNAs tend to encode regulatory proteins, while PB-excluded mRNAs encode basic functions. Furthermore, proteins of the same complexes tended to be encoded by mRNA co-regulated in or out of PBs, in so-called post-transcriptional regulons (Hubstenberger et al., 2017; Standart and Weil, 2018). We speculate that specific codon usage could also underlie these post-transcriptional regulons.

Evolutionary issues

Our results raise intriguing issues in terms of evolution. While PBs are found in all eukaryotes, animal and vegetal, the GC-rich part of the human genome only emerged in amniotes (the ancestor of birds and mammals) (Duret et al., 2002). In more distant organisms, such as yeast, *C. elegans*, *Drosophila* or *Xenopus*, genes have a narrow and most often AU-rich GC content distribution (Figure S9). Thus, despite the conservation of the DDX6, XRN1 and to a lesser extent PAT1B proteins in eukaryotes, distinct mRNA decay pathways depending on their GC content may have evolved more recently. Moreover, the enzymatic properties of DDX6 could have adapted to the higher GC content of human transcripts.

The GC-rich part of the human genome was acquired through GC biased gene conversion (gBGC), a non-selective process linked with meiotic recombination affecting GC content evolution (Duret and Galtier, 2009). We considered the possibility that meiotic recombination occurred more frequently in genomic regions containing genes involved in basic functions, leading to stronger gBGC and, consequently, to higher GC content of PB-excluded mRNA. However, our analysis showed that mRNA base composition and PB enrichment are associated independently of meiotic recombination or the genomic context. We therefore put forward a model where the genome of higher eukaryotes has evolved partly to facilitate the control of regulators at the translation level, by limiting their protein yield. Regardless, the overall outcome of our study is that in human the GC content, a feature written in the genome, shapes mRNA fate and its control in a strikingly coherent system.

Acknowledgments

We thank Michèle Ernoult-Lange for scientific discussion, Marina Pinskaya and Marc Gabriel for scientific discussions and technical assistance. We also thank Virginie Magnone, Kevin Lebrigand (NGS platform, UCA Genomix), Sylvain Baulande, Patricia Legoix-Né and Virginie Raynal (NGS platform, Institut Curie).

Funding

This work was supported by the Association pour la Recherche sur le Cancer and the Agence Nationale pour la Recherche contract ANR-14-CE09-0013-01 to D.W.'s laboratory, the European Research Council (DARK consolidator grant) to A.M.'s laboratory, the LABEX SIGNALIFE [ANR-11-LABX-0028-01], Canceropole PACA to P.B.'s laboratory, BBSRC, Newton Trust and Foundation Wiener – Anspach (C.V.) to N.S.'s laboratory. This work has benefited from the facilities and expertise of the NGS platform of Institut Curie (supported by the Agence Nationale de la Recherche [ANR-10-EQPX-03, ANR10-INBS-09-08] and the Canceropôle Ile-de-France), and of the "UCA GenomiX" Functional Genomics Platform of the University of Nice Sophia Antipolis, a partner of the National Infrastructure France Génomique,

supported by Commissariat aux Grands Investissements (ANR-10-INBS-09-03, ANR-10-INBS-09-02) and Canceropôle PACA.

Methods

Cell culture and transfection

Human embryonic kidney HEK 293 cells and epithelioid carcinoma HeLa cells were maintained in DMEM supplemented with 10% (v/v) fetal calf serum. HCT116 were grown in McCoy's 5A modified medium supplemented with 10% (v/v) fetal bovine serum, 5% (v/v) sodium pyruvate and 5% (v/v) non-essential amino acids.

For DDX6 silencing, 7.10^5 cells were transfected at the time of their plating (reverse transfection) with 50 pmoles DDX6 or control β-globin siRNAs (Minshall et al., 2009) per 3 cm diameter well, using Lipofectamin 2000 (Life Technologies, France), and split in two 24 hours later. Cells were lyzed 48 h after transfection.

For XRN1 silencing with siRNAs, 2.10^5 cells/well were plated in 6-well plates and transfected 24 h later with 50 nM siRNA negative Control or Silencer Pre-designed siRNA XRN1, using Lipofectamine RNAiMAX (Life Technologies). Cells were lyzed 48 h after transfection.

For XRN1 silencing with shRNA, a doxocyclin inducible construct provided by Dharmacon (TRIPZ) with shRNA against Xrn1 or non-silencing shRNA was introduced by lentiviral transduction (MOI 0.5). After 10 days of puromycin selection, cells were tested for expression of the construct. For shRNA induction cells were grown to 30% confluency in 10 cm plates before adding 1 µg/ml doxocycline . After 24h, cells were split in three and doxocycline was maintained untill 48h.

Western blot analysis

Total cell lysates were obtained as described previously (Courel et al., 2006). Proteins were separated on SDS-PAGE on 4–12% polyacrylamide gel (NuPage, Invitrogen) and transferred onto nitrocellulose membrane (PerkinElmer, France). After blocking in 5% (w/v) nonfat dry milk in PBS for 30 min at room temperature, the membrane was incubated for 1 h at 37°C with primary antibodies. After washing in PBS containing 0.05% (v/v) tween-20, blots were incubated for 40 min at room temperature with horseradish peroxidase-conjugated secondary anti-rabbit antibody (1:10000; Jackson Immunoresearch Laboratories). Immunoreactive bands were visualized by chemiluminescence detection of peroxidase activity (SuperSignal West Pico, Pierce) and exposure to CL-XPosure film (Pierce). Protein expression was evaluated by densitometry (NIH ImageJ).

For the shRNA Xrn1 experiment, proteins were isolated using RIPA buffer with Halt™ Protease and Phosphatase Inhibitor Cocktail (Thermo Scientific), precipitated with acetone and separated on Tris-Acetate 3-8% polyacrylamide gel (NuPage, Invitrogen) before transfer to nitrocellulose membrane (GE Healthcare). After blocking with 5% (w/v) nonfat dry milk in TBST for 30 min at room temperature the membrane was incubated for 1h at RT or o/n in 4°C with primary antibodies. After washing in TBST containing 0.05% (v/v) tween-20, blots were subsequently incubated for 1h at room temperature with horseradish peroxidase-conjugated secondary anti-rabbit/mouse antibodies (1:10000; Sigma). Immunoreactive bands were visualized by chemiluminescence detection of peroxidase activity (SuperSignal West Dura, Pierce) and imaged using ImageQuant LAS 4000 (GE Healthcare). Protein expression was evaluated by densitometry (NIH ImageJ).

Primary antibodies were: rabbit polyclonal anti-DDX6 (1:15000; Novus Biological), rabbit polyclonal anti-ribosomal S6 (1:5000; Cell Signaling Technology), anti-XRN1 (1:1000, Bethyl Laboratories), anti-XRN1 (1:5000 Novus Bioscience), anti-PolII (1:100, Santa Cruz).

Polysome profiling experiment

The polysome profiling experiment was described in (Hubstenberger et al., 2017). For q-RT-PCR experiments, 1 µg RNA was reverse transcribed for 1 h at 50°C using the SuperScript III First-Strand Synthesis System for RT-PCR (Life Technologies) with 1 µg random primers. No amplification was detected in negative controls omitting the reverse transcriptase. qPCR amplifications (12 µl) were done in duplicates using 1 µl of cDNA and the GoTaq Probe 2X Master Mix (Promega) on a LightCycler 480 (Roche), programmed as follows: 5 min, 95°C; 40 cycles [10 sec, 95°C; 15 sec, 60°C; 10 sec, 72°C]. A last step (5 min progressive 95°C to 72°C) calculating the qPCR product Tm allowed for reaction specificity check. Primers for ACTB, APP, BACE1, LSM14A, LSM14B, MFN2, PNRC1, TIMP2 and TRIB1 were designed using the Primer 3 software (Untergasser et al., 2012). The results were normalized using HPRT.

Library preparation and RNA-Seq data processing

For polysome profiling after DDX6 silencing in HEK293 cells, rRNA was depleted using the Ribo-Zero kit Human/Mouse/Rat (Epicentre), and libraries were prepared using random priming. Triplicate libraries were generated from three independent experiments and processed as detailed previously (Hubstenberger et al., 2017). For transcriptome after PAT1B silencing, rRNA was depleted using the Ribo-Zero kit and libraries were prepared using random priming. Duplicate libraries were generated from two independent experiments and processed as described previously (Vindry et al., 2017).

For the transcriptome after Xrn1 silencing in HeLa cells, libraries were prepared from 500 ng of total RNAs and oligo(dT) primed using TruSeq Stranded Total RNA kit (Illumina) with two technical replicates for each sample. Libraries were then quantified with KAPA Library Quantification kit (Kapa Biosystems) and pooled. 4 nM of this pool were loaded on a high output flowcell and sequenced on a NextSeq500 platform (Illumina) with 2x75nt paired-end chemistry.

For the shRNA Xrn1 experiment RNA was isolated using Quiazol and miRNase Mini Kit (Quiagen), next subjected to DNase treatment (Quiagen) and quality control with Bioanalyzer. The rRNA was removed using rRNA Removal Mix. Libraries were prepared from 1 ug of RNA following TruSeq Stranded Total RNA kit (Illumina) with two technical replicates for each sample. 100nt paired-end RNA-Seq was performed on HiSeq - Rapid Run (Illumina). The results were aligned using hg19 genome and DESeq2, with standard settings, was used for determining FC and p-values.

For PB enrichment, libraries were prepared without prior elimination of rRNA and using random priming. Triplicate libraries were generated from three independent experiments and processed using the same pipeline as for DDX6 silencing (Hubstenberger et al., 2017).

For the transcriptome after induction of a stably transfected DDX6 shRNA for 48h in K562 cells, the .fastq files from experiments ENCSR119QWQ (DDX6 shRNA) and ENCSR913CAE (control shRNA) were processed according to the same pipeline as DDX6 polysome profiling, except that the control and DDX6 shRNA experiments were not paired to compute the corrected p-values of RNA differential expression in EdgeR_3.6.2.

The ENCODE dataset of mRNA clipped to DDX6 in K562 cells was generated using ENCODE .bam files aligned on the hg19 genome corresponding to (i) the DDX6 eClip experiment ENCSR893EFU, and (ii) the total RNA-seq of K562 experiment ENCSR109IQO. The enrichment in the CLIP dataset compared to the total RNA sample was calculated as in the DDX6 shRNA experiment.

Bioinformatic analysis

For the pie chart in Fig. 1A, mRNAs with a cpm > 1 and a log2 FC > 0.25 or < -0.25 in at least total or polysomal RNA (n=4538) were categorized as changing (i) in total mRNA only, (ii) in polysomal mRNA only, (iii) accordingly in both total and polysomal mRNA or (iv) divergently in total and polysomal mRNA.

Briefly, the read coverage was computed as follows. Raw reads were processed using trimmomatics. Alignment was performed on the longest transcript isoforms of Ensembl annotated genes with bowtie2 aligner. Isoforms shorter than 500 nucleotides were not considered. Only unique mapped reads were qualified for counting. Each transcript was subdivided in 10 bins from transcription start site (TSS) to transcript end site (TES) and the proportion of reads for each bin was computed. For the metagene analysis, the average distribution of read along transcript length was computed so that each gene had the same weight independently of it expression level.

The protein yield was calculated as the ratio between protein abundance in HEK293 cells, taken from (Geiger et al., 2012), and mRNA abundance in HEK293 cells, taken from the control sample of our DDX6 polysome profiling experiment.

For GC profiling of the transcripts in various organisms (Figures 3B and S8E), transcripts were downloaded from ENSEMBL (version 92) with their associated GC content.

Boxplot representations and statistical Mann-Whitney-U tests were performed using the GraphPad Prism software (GraphPad software, Inc.). Other graphical representations and Pearson correlation coefficients were generated using Excel and the Excel Analysis ToolPak (Microsoft) and the R suite (https://www.R-project.org) (R Core Team 2018. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria).

Gene meiotic recombination rates were computed as crossover rates between gene start and gene end using the genetic map from the HapMap project (International HapMap Consortium et al., 2007). Rates were computed as the weighted average of crossover rates of chromosomal regions that overlap the window.

Datasets used in the bioinformatics analysis

The following datasets were downloaded from the supplementary material of the corresponding papers: 1) For mRNA-containing cis-regulatory motifs: (i) in silico identification: AREs (Halees et al., 2007); (ii) experimental determination: CPEs (Piqué et al., 2008), G4-containing (Huppert et al., 2008) and TOP mRNAs (Thoreen et al., 2012). G4-containing genes were restricted to those harboring a G4 in 5'UTR. 2) For RBP targets: (i) PARE in HeLa cells: SMG6 (mRNAs actually cleaved by SMG6) (Schmidt et al., 2015) (ii) CLIP in HEK293 cells: HuR and TTP (Mukherjee et al., 2011, 2014); (iii) CLIP in HeLa cells: YTHDF2 (Wang et al., 2013); (iv) RIP-CHIP in HeLa cells: PUM1 (Galgano et al., 2008); (v) RIP-CHIP in neurons: 4E-T (Yang et al., 2014). In the case of HuR, the transcripts clipped only in 5'UTR/CDS/introns were removed, and the target list was restricted to transcripts clipped more than once. In the case of TTP, the transcripts clipped only in introns were removed.

For other RBP targets (ATXN2, MOV10, IGF2BP1-3, PUM2, FMR1, FXR1-2, AGO1-4, YTHDF2), CLIP data from different laboratories were previously processed through the same pipeline in the CLIPdb 1.0 database using the Piranha method (Yang et al., 2015). We retained those performed in epithelial cells (HeLa, HEK293, HEK293T). Moreover, when replicates were available, we selected the RNA-protein interactions detected in at least 50% of the replicates. Except for FXR1-2 targets, we determined whether protein-RNA interactions occurred in UTR or CDS by intersecting coordinates of the read peaks with the v19 gencode annotation, and we removed the transcripts clipped in their CDS.

For miRNA targets, we extracted the list of all experimentally documented targets from miRTarBase (http://mirtarbase.mbc.nctu.edu.tw/php/index.php) (Hsu et al., 2014), and selected the targets of the 22 miRNAs of interest.

Data availability

RNA-Seq gene expression data and raw fastq files are available on the SRA repository https://www.ebi.ac.uk/ena under accession number: E-MTAB-4091 for the polysome profiling after DDX6 silencing, E-MTAB-5577 for the transcriptome after PAT1B silencing, E-MTAB-5477 for the PB transcriptome.
RNA-Seq gene expression data from XRN1 silencing experiments is available on the GEO repository: GSE115471 and GSE114605 for the HeLa and HCT116 datasets, respectively.
All ENCODE experiments are available at https://www.encodeproject.org.

References

Ayache, J., Bénard, M., Ernoult-Lange, M., Minshall, N., Standart, N., Kress, M., and Weil, D. (2015). P-body assembly requires DDX6 repression complexes rather than decay or Ataxin2/2L complexes. Mol. Biol. Cell *26*, 2579–2595.

Barbee, S.A., Estes, P.S., Cziko, A.-M., Hillebrand, J., Luedeman, R.A., Coller, J.M., Johnson, N., Howlett, I.C., Geng, C., Ueda, R., et al. (2006). Staufen- and FMRP-Containing Neuronal RNPs Are Structurally and Functionally Related to Somatic P Bodies. Neuron *52*, 997–1009.

Bashkirov, V.I., Scherthan, H., Solinger, J.A., Buerstedde, J.-M., and Heyer, W.-D. (1997). A mouse cytoplasmic exoribonuclease (mXRN1p) with preference for G4 tetraplex substrates. J. Cell Biol. *136*, 761–773.

Bazzini, A.A., del Viso, F., Moreno-Mateos, M.A., Johnstone, T.G., Vejnar, C.E., Qin, Y., Yao, J., Khokha, M.K., and Giraldez, A.J. (2016). Codon identity regulates mRNA stability and translation efficiency during the maternal-to-zygotic transition. EMBO J. *35*, 2087–2103.

Bhattacharyya, S.N., Habermacher, R., Martine, U., Closs, E.I., and Filipowicz, W. (2006). Relief of microRNA-Mediated Translational Repression in Human Cells Subjected to Stress. Cell *125*, 1111–1124.

Bish, R., Cuevas-Polo, N., Cheng, Z., Hambardzumyan, D., Munschauer, M., Landthaler, M., and Vogel, C. (2015). Comprehensive Protein Interactome Analysis of a Key RNA Helicase: Detection of Novel Stress Granule Proteins. Biomolecules *5*, 1441–1466.

Bonnerot, C., Boeck, R., and Lapeyre, B. (2000). The two proteins Pat1p (Mrt1p) and Spb8p interact in vivo, are required for mRNA decay, and are functionally linked to Pab1p. Mol. Cell. Biol. *20*, 5939–5946.

Bouveret, E., Rigaut, G., Shevchenko, A., Wilm, M., and Séraphin, B. (2000). A Sm-like protein complex that participates in mRNA degradation. EMBO J. *19*, 1661–1671.

Braun, J.E., Tritschler, F., Haas, G., Igreja, C., Truffault, V., Weichenrieder, O., and Izaurralde, E. (2010). The C-terminal α–α superhelix of Pat is required for mRNA decapping in metazoa. EMBO J. *29*, 2368–2380.

Chen, Y., Boland, A., Kuzuoğlu-Öztürk, D., Bawankar, P., Loh, B., Chang, C.-T., Weichenrieder, O., and Izaurralde, E. (2014). A DDX6-CNOT1 Complex and W-Binding Pockets in CNOT9 Reveal Direct Links between miRNA Target Recognition and Silencing. Mol. Cell *54*, 737–750.

Chu, C., and Rana, T.M. (2006). Translation repression in human cells by microRNA-induced gene silencing requires RCK/p54. PLoS Biol. *4*, e210.

Coller, J., and Parker, R. (2005). General Translational Repression by Activators of mRNA Decapping. Cell *122*, 875–886.

Colombo, M., Karousis, E.D., Bourquin, J., Bruggmann, R., and Mühlemann, O. (2017). Transcriptome-wide identification of NMD-targeted human mRNAs reveals extensive redundancy between SMG6- and SMG7-mediated degradation pathways. RNA N. Y. N *23*, 189–201.

Cougot, N. (2004). Cytoplasmic foci are sites of mRNA decay in human cells. J. Cell Biol. *165*, 31–40.

Courel, M., Rodemer, C., Nguyen, S.T., Pance, A., Jackson, A.P., O'connor, D.T., and Taupenot, L. (2006). Secretory granule biogenesis in sympathoadrenal cells: identification of a granulogenic determinant in the secretory prohormone chromogranin A. J. Biol. Chem. *281*, 38038–38051.

Duret, L., and Galtier, N. (2009). Biased gene conversion and the evolution of mammalian genomic landscapes. Annu. Rev. Genomics Hum. Genet. *10*, 285–311.

Duret, L., Semon, M., Piganeau, G., Mouchiroud, D., and Galtier, N. (2002). Vanishing GC-rich isochores in mammalian genomes. Genetics *162*, 1837–1847.

ENCODE Project Consortium (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74.

Ernoult-Lange, M., Wilczynska, A., Harper, M., Aigueperse, C., Dautry, F., Kress, M., and Weil, D. (2009). Nucleocytoplasmic traffic of CPEB1 and accumulation in Crm1 nucleolar bodies. Mol. Biol. Cell *20*, 176–187.

Ernoult-Lange, M., Baconnais, S., Harper, M., Minshall, N., Souquere, S., Boudier, T., Bénard, M., Andrey, P., Pierron, G., Kress, M., et al. (2012). Multiple binding of repressed mRNAs by the P-body protein Rck/p54. RNA N. Y. N *18*, 1702–1715.

Ferraiuolo, M.A., Basak, S., Dostie, J., Murray, E.L., Schoenberg, D.R., and Sonenberg, N. (2005). A role for the eIF4E-binding protein 4E-T in P-body formation and mRNA decay. J. Cell Biol. *170*, 913–924.

Franks, T.M., and Lykke-Andersen, J. (2007). TTP and BRF proteins nucleate processing body formation to silence mRNAs with AU-rich elements. Genes Dev. *21*, 719–735.

Galgano, A., Forrer, M., Jaskiewicz, L., Kanitz, A., Zavolan, M., and Gerber, A.P. (2008). Comparative Analysis of mRNA Targets for Human PUF-Family Proteins Suggests Extensive Interaction with the miRNA Regulatory System. PLoS ONE *3*, e3164.

Geiger, T., Wehner, A., Schaab, C., Cox, J., and Mann, M. (2012). Comparative Proteomic Analysis of Eleven Common Cell Lines Reveals Ubiquitous but Varying Expression of Most Proteins. Mol. Cell. Proteomics *11*, M111.014050-M111.014050.

Gingold, H., Tehler, D., Christoffersen, N.R., Nielsen, M.M., Asmar, F., Kooistra, S.M., Christophersen, N.S., Christensen, L.L., Borre, M., Sørensen, K.D., et al. (2014). A Dual Program for Translation Regulation in Cellular Proliferation and Differentiation. Cell *158*, 1281–1292.

Götze, M., Dufourt, J., Ihling, C., Rammelt, C., Pierson, S., Sambrani, N., Temme, C., Sinz, A., Simonelig, M., and Wahle, E. (2017). Translational repression of the Drosophila nanos mRNA involves the RNA helicase Belle and RNA coating by Me31B and Trailer hitch. RNA N. Y. N *23*, 1552–1568.

Hafner, M., Landthaler, M., Burger, L., Khorshid, M., Hausser, J., Berninger, P., Rothballer, A., Ascano, M., Jungkamp, A.-C., Munschauer, M., et al. (2010). Transcriptome-wide Identification of RNA-Binding Protein and MicroRNA Target Sites by PAR-CLIP. Cell *141*, 129–141.

Halees, A.S., El-Badrawi, R., and Khabar, K.S.A. (2007). ARED Organism: expansion of ARED reveals AU-rich element cluster variations between human and mouse. Nucleic Acids Res. *36*, D137–D140.

Hentze, M.W., Castello, A., Schwarzl, T., and Preiss, T. (2018). A brave new world of RNA-binding proteins. Nat. Rev. Mol. Cell Biol. *19*, 327–341.

Horvathova, I., Voigt, F., Kotrys, A.V., Zhan, Y., Artus-Revel, C.G., Eglinger, J., Stadler, M.B., Giorgetti, L., and Chao, J.A. (2017). The Dynamics of mRNA Turnover Revealed by Single-Molecule Imaging in Single Cells. Mol. Cell *68*, 615-625.e9.

Hsu, S.-D., Tseng, Y.-T., Shrestha, S., Lin, Y.-L., Khaleel, A., Chou, C.-H., Chu, C.-F., Huang, H.-Y., Lin, C.-M., Ho, S.-Y., et al. (2014). miRTarBase update 2014: an information resource for experimentally validated miRNA-target interactions. Nucleic Acids Res. *42*, D78-85.

Huang, N., Lee, I., Marcotte, E.M., and Hurles, M.E. (2010). Characterising and predicting haploinsufficiency in the human genome. PLoS Genet. *6*, e1001154.

Hubstenberger, A., Courel, M., Bénard, M., Souquere, S., Ernoult-Lange, M., Chouaib, R., Yi, Z., Morlot, J.-B., Munier, A., Fradet, M., et al. (2017). P-Body Purification Reveals the Condensation of Repressed mRNA Regulons. Mol. Cell *68*, 144-157.e5.

Huppert, J.L., Bugaut, A., Kumari, S., and Balasubramanian, S. (2008). G-quadruplexes: the beginning and end of UTRs. Nucleic Acids Res. *36*, 6260–6268.

Imamachi, N., Salam, K.A., Suzuki, Y., and Akimitsu, N. (2017). A GC-rich sequence feature in the 3' UTR directs UPF1-dependent mRNA decay in mammalian cells. Genome Res. *27*, 407–418.

International HapMap Consortium, Frazer, K.A., Ballinger, D.G., Cox, D.R., Hinds, D.A., Stuve, L.L., Gibbs, R.A., Belmont, J.W., Boudreau, A., Hardenbol, P., et al. (2007). A second generation human haplotype map of over 3.1 million SNPs. Nature *449*, 851–861.

Kamenska, A., Lu, W.-T., Kubacka, D., Broomhead, H., Minshall, N., Bushell, M., and Standart, N. (2014). Human 4E-T represses translation of bound mRNAs and enhances microRNA-mediated silencing. Nucleic Acids Res. *42*, 3298–3313.

Kamenska, A., Simpson, C., Vindry, C., Broomhead, H., Bénard, M., Ernoult-Lange, M., Lee, B.P., Harries, L.W., Weil, D., and Standart, N. (2016). The DDX6–4E-T interaction mediates translational repression and P-body assembly. Nucleic Acids Res. *44*, 6318–6334.

Khong, A., Matheny, T., Jain, S., Mitchell, S.F., Wheeler, J.R., and Parker, R. (2017). The Stress Granule Transcriptome Reveals Principles of mRNA Accumulation in Stress Granules. Mol. Cell *68*, 808-820.e5.

Łabno, A., Tomecki, R., and Dziembowski, A. (2016). Cytoplasmic RNA decay pathways - Enzymes and mechanisms. Biochim. Biophys. Acta *1863*, 3125–3147.

Lebedeva, S., Jens, M., Theil, K., Schwanhäusser, B., Selbach, M., Landthaler, M., and Rajewsky, N. (2011). Transcriptome-wide Analysis of Regulatory Interactions of the RNA-Binding Protein HuR. Mol. Cell *43*, 340–352.

Liu, J., Valencia-Sanchez, M.A., Hannon, G.J., and Parker, R. (2005). MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. Nat. Cell Biol. *7*, 719–723.

Marnef, A., and Standart, N. (2010). Pat1 proteins: a life in translation, translation repression and mRNA decay. Biochem Soc Trans *38*, 1602–1607.

Mathys, H., Basquin, J., Ozgur, S., Czarnocki-Cieciura, M., Bonneau, F., Aartse, A., Dziembowski, A., Nowotny, M., Conti, E., and Filipowicz, W. (2014). Structural and Biochemical Insights to the Role of the CCR4-NOT Complex and DDX6 ATPase in MicroRNA Repression. Mol. Cell *54*, 751–765.

Minshall, N., Reiter, M.H., Weil, D., and Standart, N. (2007). CPEB interacts with an ovary-specific eIF4E and 4E-T in early Xenopus oocytes. J Biol Chem *282*, 37389–37401.

Minshall, N., Kress, M., Weil, D., and Standart, N. (2009). Role of p54 RNA helicase activity and its C-terminal domain in translational repression, P-body localization and assembly. Mol Biol Cell *20*, 2464–2472.

Mishima, Y., and Tomari, Y. (2017). Pervasive yet nonuniform contributions of Dcp2 and Cnot7 to maternal mRNA clearance in zebrafish. Genes Cells Devoted Mol. Cell. Mech. *22*, 670–678.

Mitchell, S.F., Jain, S., She, M., and Parker, R. (2012). Global analysis of yeast mRNPs. Nat. Struct. Mol. Biol. *20*, 127–133.

Mukherjee, N., Corcoran, D.L., Nusbaum, J.D., Reid, D.W., Georgiev, S., Hafner, M., Ascano, M., Tuschl, T., Ohler, U., and Keene, J.D. (2011). Integrative Regulatory Mapping Indicates that the RNA-Binding Protein HuR Couples Pre-mRNA Processing and mRNA Stability. Mol. Cell *43*, 327–339.

Mukherjee, N., Jacobs, N.C., Hafner, M., Kennington, E.A., Nusbaum, J.D., Tuschl, T., Blackshear, P.J., and Ohler, U. (2014). Global target mRNA specification and regulation by the RNA-binding protein ZFP36. Genome Biol *15*, R12.

Nissan, T., Rajyaguru, P., She, M., Song, H., and Parker, R. (2010). Decapping activators in Saccharomyces cerevisiae act by multiple mechanisms. Mol. Cell *39*, 773–783.

Nonhoff, U., Ralser, M., Welzel, F., Piccini, I., Balzereit, D., Yaspo, M.-L., Lehrach, H., and Krobitsch, S. (2007). Ataxin-2 interacts with the DEAD/H-box RNA helicase DDX6 and interferes with P-bodies and stress granules. Mol. Biol. Cell *18*, 1385–1396.

Ozgur, S., Chekulaeva, M., and Stoecklin, G. (2010). Human Pat1b Connects Deadenylation with mRNA Decapping and Controls the Assembly of Processing Bodies. Mol. Cell. Biol. *30*, 4308–4323.

Ozgur, S., Basquin, J., Kamenska, A., Filipowicz, W., Standart, N., and Conti, E. (2015). Structure of a Human 4E-T/DDX6/CNOT1 Complex Reveals the Different Interplay of DDX6-Binding Proteins with the CCR4-NOT Complex. Cell Rep. *13*, 703–711.

Pashler, A.L., Towler, B.P., Jones, C.I., and Newbury, S.F. (2016). The roles of the exoribonucleases DIS3L2 and XRN1 in human disease. Biochem. Soc. Trans. *44*, 1377–1384.

Piqué, M., López, J.M., Foissac, S., Guigó, R., and Méndez, R. (2008). A Combinatorial Code for CPE-Mediated Translational Control. Cell *132*, 434–448.

Radhakrishnan, A., Chen, Y.-H., Martin, S., Alhusaini, N., Green, R., and Coller, J. (2016). The DEAD-Box Protein Dhh1p Couples mRNA Decay and Translation by Monitoring Codon Optimality. Cell *167*, 122-132.e9.

Rudolph, K.L.M., Schmitt, B.M., Villar, D., White, R.J., Marioni, J.C., Kutter, C., and Odom, D.T. (2016). Codon-Driven Translational Efficiency Is Stable across Diverse Mammalian Cell States. PLOS Genet. *12*, e1006024.

Schmidt, S.A., Foley, P.L., Jeong, D.-H., Rymarquis, L.A., Doyle, F., Tenenbaum, S.A., Belasco, J.G., and Green, P.J. (2015). Identification of SMG6 cleavage sites and a preferred RNA cleavage motif by global analysis of endogenous NMD targets in human cells. Nucleic Acids Res. *43*, 309–323.

Standart, N., and Minshall, N. (2008). Translational control in early development: CPEB, P-bodies and germinal granules. Biochem. Soc. Trans. *36*, 671.

Standart, N., and Weil, D. (2018). P-Bodies: Cytosolic Droplets for Coordinated mRNA Storage. Trends Genet. TIG.

Steinberg, J., Honti, F., Meader, S., and Webber, C. (2015). Haploinsufficiency predictions without study bias. Nucleic Acids Res. *43*, e101.

Thoreen, C.C., Chantranupong, L., Keys, H.R., Wang, T., Gray, N.S., and Sabatini, D.M. (2012). A unifying model for mTORC1-mediated regulation of mRNA translation. Nature *485*, 109–113.

Totaro, A., Renzi, F., La Fata, G., Mattioli, C., Raabe, M., Urlaub, H., and Achsel, T. (2011). The human Pat1b protein: a novel mRNA deadenylation factor identified by a new immunoprecipitation technique. Nucleic Acids Res. *39*, 635–647.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., and Rozen, S.G. (2012). Primer3--new capabilities and interfaces. Nucleic Acids Res. *40*, e115.

Vindry, C., Marnef, A., Broomhead, H., Twyffels, L., Ozgur, S., Stoecklin, G., Llorian, M., Smith, C.W., Mata, J., Weil, D., et al. (2017). Dual RNA Processing Roles of Pat1b via Cytoplasmic Lsm1-7 and Nuclear Lsm2-8 Complexes. Cell Rep. *20*, 1187–1200.

Wang, M., Ly, M., Lugowski, A., Laver, J.D., Lipshitz, H.D., Smibert, C.A., and Rissland, O.S. (2017). ME31B globally represses maternal mRNAs by two distinct mechanisms during the Drosophila maternal-to-zygotic transition. ELife *6*.

Wang, X., Lu, Z., Gomez, A., Hon, G.C., Yue, Y., Han, D., Fu, Y., Parisien, M., Dai, Q., Jia, G., et al. (2013). N6-methyladenosine-dependent regulation of messenger RNA stability. Nature *505*, 117–120.

Wells, M.L., Perera, L., and Blackshear, P.J. (2017). An Ancient Family of RNA-Binding Proteins: Still Important! Trends Biochem. Sci. *42*, 285–296.

Wu, D., Muhlrad, D., Bowler, M.W., Jiang, S., Liu, Z., Parker, R., and Song, H. (2014). Lsm2 and Lsm3 bridge the interaction of the Lsm1-7 complex with Pat1 for decapping activation. Cell Res. *24*, 233–246.

Yang, G., Smibert, C.A., Kaplan, D.R., and Miller, F.D. (2014). An eIF4E1/4E-T Complex Determines the Genesis of Neurons from Precursors by Translationally Repressing a Proneurogenic Transcription Program. Neuron *84*, 723–739.

Yang, Y.-C.T., Di, C., Hu, B., Zhou, M., Liu, Y., Song, N., Li, Y., Umetsu, J., and Lu, Z.J. (2015). CLIPdb: a CLIP-seq database for protein-RNA interactions. BMC Genomics *16*, 51.

Yoon, J.-H., De, S., Srikantan, S., Abdelmohsen, K., Grammatikakis, I., Kim, J., Kim, K.M., Noh, J.H., White, E.J.F., Martindale, J.L., et al. (2014). PAR-CLIP analysis uncovers AUF1 impact on target RNA fate and genome integrity. Nat. Commun. *5*, 5248.

Figure legends

Figure 1: DDX6 silencing reveals opposite effects on mRNA decay and translation depending on their GC content.

(A) Polysome profiling after DDX6 silencing in HEK293 cells. The pie chart represents the fold-changes in total and polysomal mRNA, using a threshold of +/- 0.25 (see methods). Both (in stripped gray) means concordant positive or negative fold-changes in total and polysomal mRNA. 3.9% of the transcripts had inverse fold-changes the two datasets (in black), probably reflecting non-significant variations.

(B) Dot plot representation of the changes in translation rate (Polysome/Total FC) as a function of changes in mRNA stability (Total FC).

(C) mRNA stabilization after DDX6 silencing in HEK293 and K562 cells applies to GC-rich mRNAs. Transcripts were subdivided into six classes depending on the GC content of their gene (from <40 to >60%). The boxplots represent the distribution of their respective fold-changes in total mRNA (in light green). The distribution of the changes of the complete dataset (all mRNAs, in grey) was added for comparison. The boxes represent the 25-75 percentiles and the whiskers the 10-90 percentiles. r is the Pearson correlation coefficient.

(D) mRNA translation derepression after DDX6 silencing in HEK293 cells applies to AU-rich mRNAs. The fold-changes in translation rate (in orange) were analyzed as in (C).

(E) mRNA stabilization after DDX6 silencing in HEK293 cells mostly depends on the GC content of their CDS and 3'UTR. The analysis was repeated as in (C) using the GC content of the 5'UTR, CDS or 3'UTR, as indicated. For 5'UTRs, the >60% class was subdivided into three classes to take into account their higher GC content compared to CDSs and 3'UTRs.

(F) GC-rich mRNAs are particularly enriched in the DDX6 CLIP experiment (in dark green).

See also Figures S1 and S2.

Figure 2: XRN1 and PAT1B involvement in mRNA decay depend on their GC content.

(A) mRNA stabilization after XRN1 silencing in HeLa and HCT116 cells applies to GC-rich mRNAs (in brown). The analysis was performed as in Figure 1C.

(B) mRNA stabilization after PAT1B silencing in HEK293 cells applies to AU-rich mRNAs (in peach). The analysis was performed as in Figure 1C.

(C) Read coverage of PAT1B targets (FC>O.6, n=616) and non-targets (FC<-0.6, n=493), as defined in the siPAT1B dataset. Their average read coverage was analyzed in control cells (left panel) and after PAT1B silencing (right panel), and normalized as described in the Methods.

(D) Read coverage of XRN1 targets (FC>1, n=333) and non-targets (FC<-1, n=139) was analyzed as in (C).

See also Figure S3.

Figure 3: PB mRNAs are AU-rich and longer than average.

(A) AU-rich mRNAs are particularly enriched in PBs (in dark red). The analysis was performed as in Figure 1C.

(B) PBs only contain transcripts from AU-rich genes. The human transcriptome was subdivided in 77 classes depending on their GC content (0.5% GC increments). The graph represents the number of PB-enriched (PB-in, in red) and PB-excluded (PB-out, in blue) transcripts in each class. The distribution of all genes is shown for comparison (in grey). The median GC value is indicated below for each group.

(C) PB mRNAs are longer than PB-excluded ones. The distribution of length of PB-enriched (PB-in) and PB-excluded (PB-out) mRNAs was analyzed as in Figure 1C.

(D) AU-rich mRNAs are better XRN1 targets when excluded from PBs than when enriched in PBs. The whole transcriptome was subdivided in bins of 500 mRNAs after ranking by GC content. For each bin, the median GC content of mRNAs was represented as a function of their median fold-change after XRN1 silencing (in grey). The same analysis was performed for PB-excluded mRNAs (PB-out), except that mRNAs under 50% GC were binned by 200 to account for their low number. The dashed line indicates the median GC content of all mRNAs.
(E) AU-rich mRNAs are not better DDX6 decay targets when they are excluded from PBs. The dataset after DDX6 silencing in HEK293 cells was analysed as in (D).
See also Figure S4.

Figure 4: Targets of various RNA decay factors are GC-rich.
(A) SMG6 targets, YTHDF2 targets and G4-containing mRNAs are particularly stabilized after DDX6 silencing in HEK293 and K562 cells. While the presented analysis uses the YTHDF2 targets reported in (Yang et al., 2015), the same pattern was observed using the targets reported in (Wang et al., 2013). Two-tail Mann-Whitney test was performed with respect to all mRNAs: **, $p<0.0001$; ns, non significant.
(B) The same mRNA subsets as in (A) are not enriched in PBs. SMG6 targets and TOP mRNAs are particularly excluded from PBs.
(C) SMG6 targets, YTHDF2 targets and G4-containing mRNAs are stabilized after DDX6 silencing mostly like other PB-excluded mRNAs. The same analysis as in (A) was conducted separately on PB-excluded (PB-out) and PB-enriched (PB-in) mRNAs.
(D) SMG6 targets, YTHDF2 targets and G4-containing mRNAs are GC-rich and stabilized after DDX6 depletion like mRNAs of the same GC content. The graph represents the median GC content of the various mRNA subsets as a function of their median fold-change after DDX6 silencing (green dots). The values observed for the whole transcriptome, subdivided in bins of 500 mRNAs as in Figure 3F (in grey), is presented for comparison. The dashed line indicates the median GC content of all mRNAs.
(E) The same mRNA subsets as in (D) are excluded from PBs. The exclusion is particularly dramatic in the case of TOP mRNAs.
See also Figure S5.

Figure 5: Targets of various translation regulators are AU-rich.
(A-C) Targets of indicated translation regulators are stabilized following PAT1B silencing (A), enriched in PBs (B) and their translation increase following DDX6 silencing (C).
(D) Summary of the features shared by the targets of the translation regulators studied in (A-C).
(E-H) Targets of translation regulators are AU-rich and behave similarly to mRNAs of the same GC content in terms of stability following PAT1B silencing (E), PB enrichment (F), translation derepression (G) and stabilization (H) following DDX6 silencing. The graphs are built as in Figure 4D.
See also Figure S6.

Figure 6: Targets of the miRNA pathway are AU-rich.
(A-C) The translation of AGO1-4 targets increases following DDX6 silencing (A), they are enriched in PBs (B) and stabilized following PAT1B silencing (C). The targets of the 22 miRNAs the most expressed in HEK293 cells (top22 miR) were also analyzed.
(D-G) AGO targets are AU-rich and behave similarly to mRNAs of the same GC content in terms of PB enrichment (D), stability following PAT1B silencing (E), translation derepression (F) and stabilization (G) following DDX6 silencing. The graphs are built as in Figure 4D.
(H-K) miRNA targets tend to be AU-rich and behave similarly to mRNAs of the same GC content in terms of PB enrichment (H), stability following PAT1B silencing (I), and translation (J) and stability (K)

following DDX6 silencing. The targets of the 22 miRNAs most expressed in HEK1293 cells were analyzed separately, as in (D-G).
See also Figure S7.

Figure 7: Codon usage is strongly biased in PBs.
(A) PB mRNAs encode proteins with a different amino-acid composition from PB-excluded mRNAs. The graph represents the frequency of each amino-acid in the proteins encoded by mRNAs enriched or excluded from PBs, using the indicated PB enrichment thresholds.
(B) Codon usage bias in and out of PBs follows their GC content. The relative codon usage for each amino-acid was calculated in mRNAs enriched (PB-in) and excluded (PB-out) mRNAs, using a PB enrichment threshold of +/- 1 (in log2). The graph represents the log2 of their ratio (PB-in/PB-out) ranked by increasing values for each amino acid. The GC content of each codon is grey-coded below, using the scale indicated in the right panel.
(C) The usage of some codons is biased in mRNAs enriched (PB-in) or excluded from (PB-out) P-bodies. Two examples are shown encoding Leucine (L) and Glycine (G).
(D) The frequency of rare codons is strongly correlated with the GC content of the CDS, independently of their PB localization. The frequency of rare codons was calculated for mRNAs excluded (PB-out) and enriched (PB-in) in PBs using a PB enrichment threshold of +/- 1 (in log2). It was expressed as a function of the CDS GC content at position 3 (GC3). Note that the slope of the tendency curves are similar for PB-enriched and -excluded transcripts.
(E) PB mRNAs have longer CDS than PB-excluded mRNAs. The analysis was performed as in Figure 3C.
(F) The number of rare codons per CDS is a good determinant of both protein yield and PB localization. The protein yield was expressed as a function of the number of rare codons for PB-enriched (PB-in) and PB-excluded (PB-out) mRNAs.
(G) Schematic representation recapitulating the features of the mRNAs regulated at the level of decay and translation depending on their GC content.
(H) The transcripts of haplo-insufficiency genes are enriched in PBs. The haplo-insufficiency score is the probability that a gene is haplo-insufficient, as taken from the Huang et al. study (Huang et al., 2010). The analysis was performed for PB-enriched (PB-in) and PB-excluded (PB-out) mRNAs. The results were similar using Steinberg et al. scores (Steinberg et al., 2015).
See also Figures S8 and S9.

Figure 1

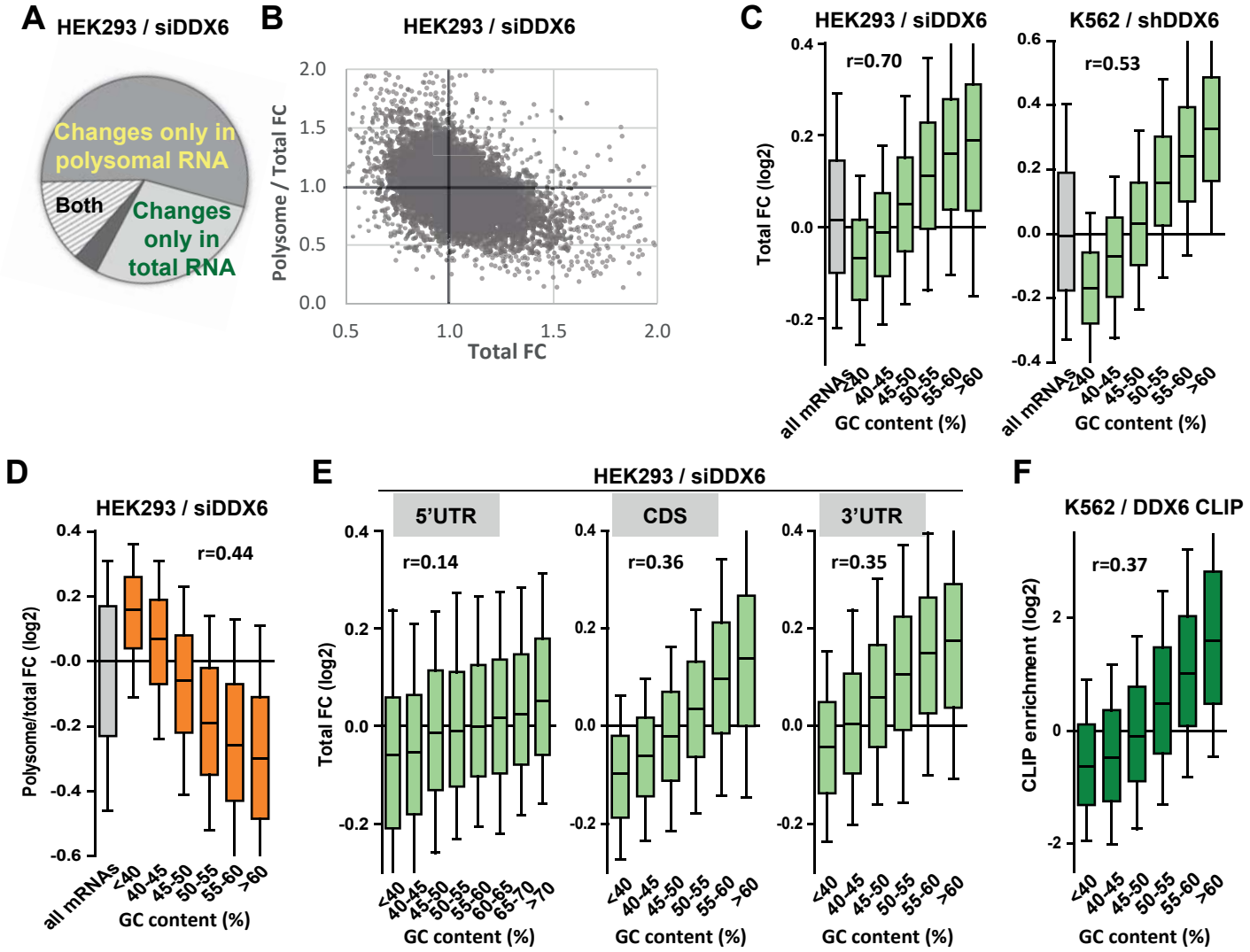

A
HEK293 / siDDX6
Changes only in polysomal RNA
Both
Changes only in total RNA
B
HEK293 / siDDX6
Polysome / Total FC
Total FC
C
HEK293 / siDDX6
r=0.70
Total FC (log2)
K562 / shDDX6
r=0.53
all mRNAs
<40
40-45
45-50
50-55
55-60
>60
GC content (%)
D
HEK293 / siDDX6
r=0.44
Polysome/total FC (log2)
E
HEK293 / siDDX6
5'UTR
r=0.14
CDS
r=0.36
3'UTR
r=0.35
Total FC (log2)
60-65
65-70
>70
F
K562 / DDX6 CLIP
r=0.37
CLIP enrichment (log2)

Figure 2

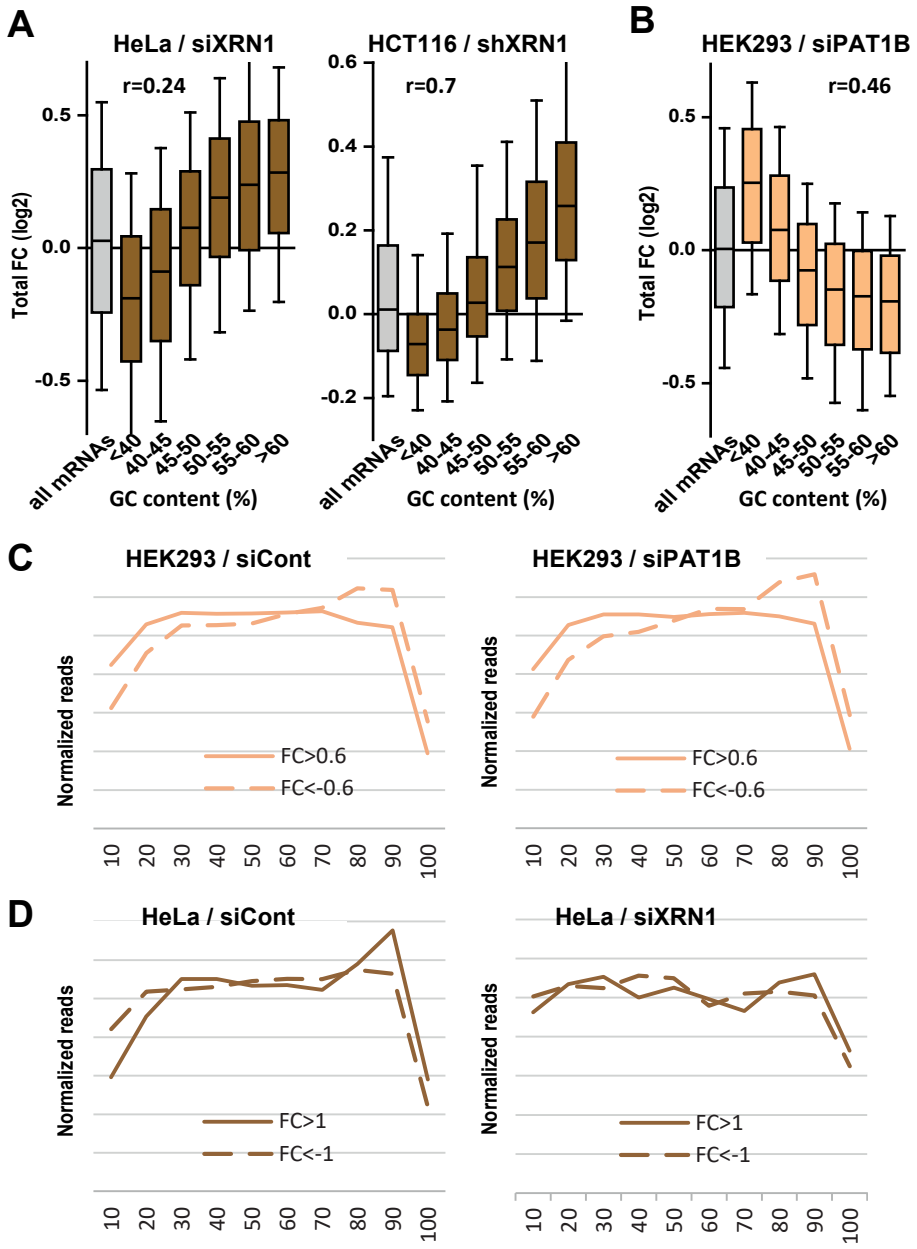

A
HeLa / siXRN1
r=0.24
HCT116 / shXRN1
r=0.7
B
HEK293 / siPAT1B
r=0.46
Total FC (log2)
0.5
0.0
-0.5
0.6
0.4
0.2
-0.2
all mRNAs
<40
40-45
45-50
50-55
55-60
>60
GC content (%)
C
HEK293 / siCont
HEK293 / siPAT1B
Normalized reads
FC>0.6
FC<-0.6
10
20
30
40
50
60
70
80
90
100
D
HeLa / siCont
HeLa / siXRN1
FC>1
FC<-1

Figure 3

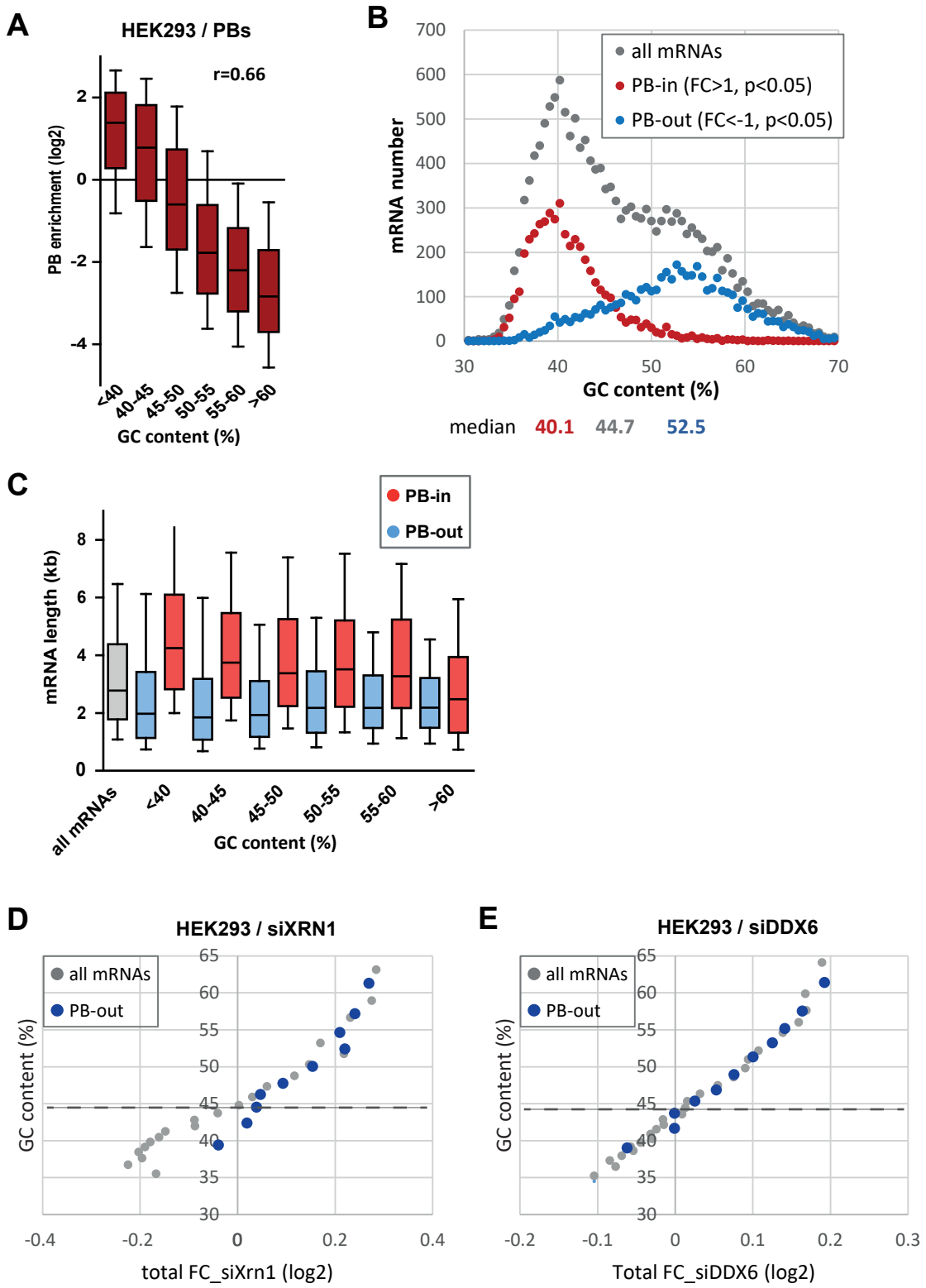

A
HEK293 / PBs
r=0.66
PB enrichment (log2)
2
0
-2
-4
<40
40-45
45-50
50-55
55-60
>60
GC content (%)
B
700
600
500
400
300
200
100
0
mRNA number
all mRNAs
PB-in (FC>1, p<0.05)
PB-out (FC<-1, p<0.05)
30
40
50
60
70
GC content (%)
median 40.1 44.7 52.5
C
PB-in
PB-out
mRNA length (kb)
8
6
4
2
0
all mRNAs
<40
40-45
45-50
50-55
55-60
>60
GC content (%)
D
HEK293 / siXRN1
all mRNAs
PB-out
GC content (%)
65
60
55
50
45
40
35
30
-0.4
-0.2
0
0.2
0.4
total FC_siXrn1 (log2)
E
HEK293 / siDDX6
all mRNAs
PB-out
GC content (%)
65
60
55
50
45
40
35
30
-0.2
-0.1
0
0.1
0.2
0.3
Total FC_siDDX6 (log2)

Figure 4

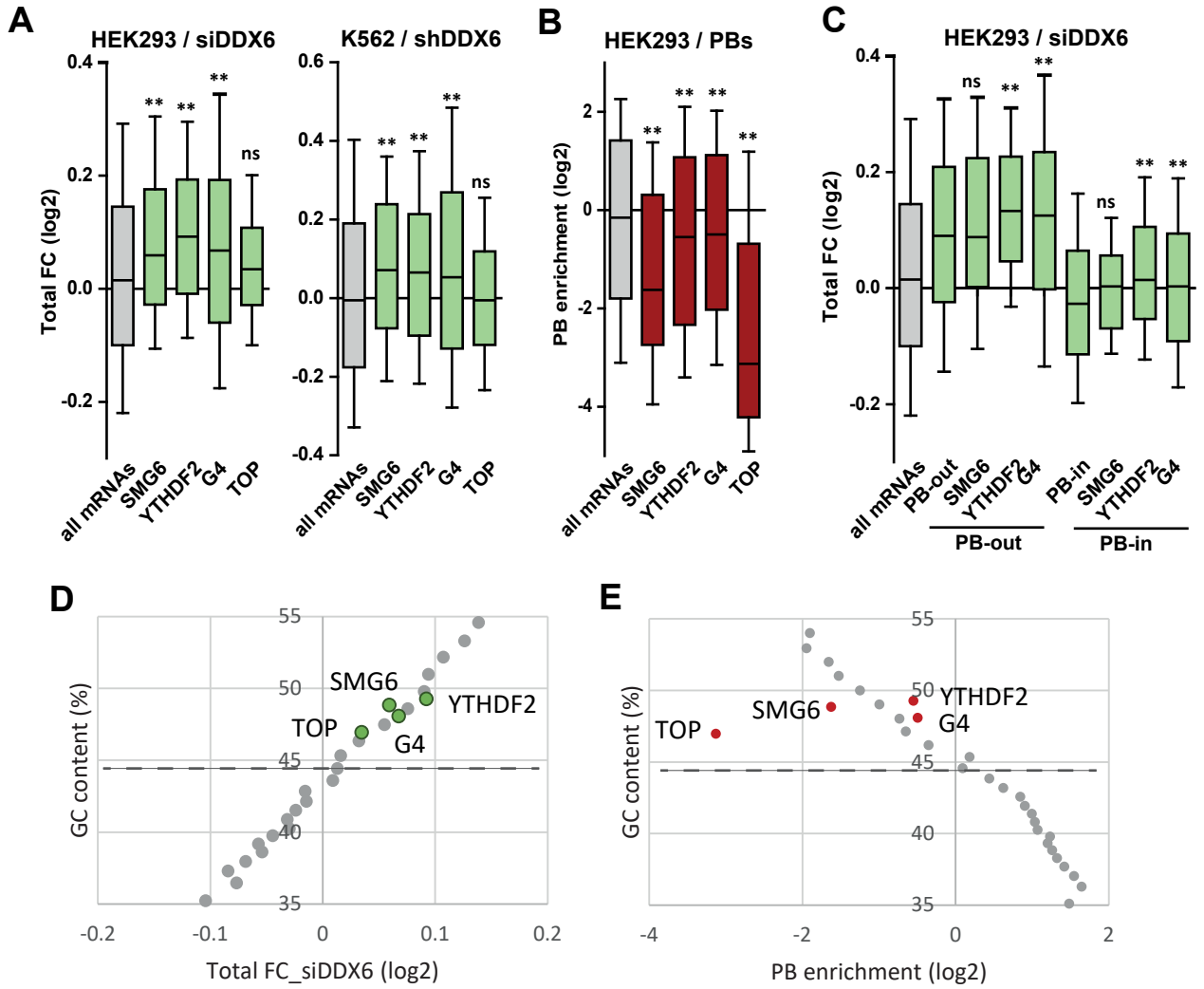
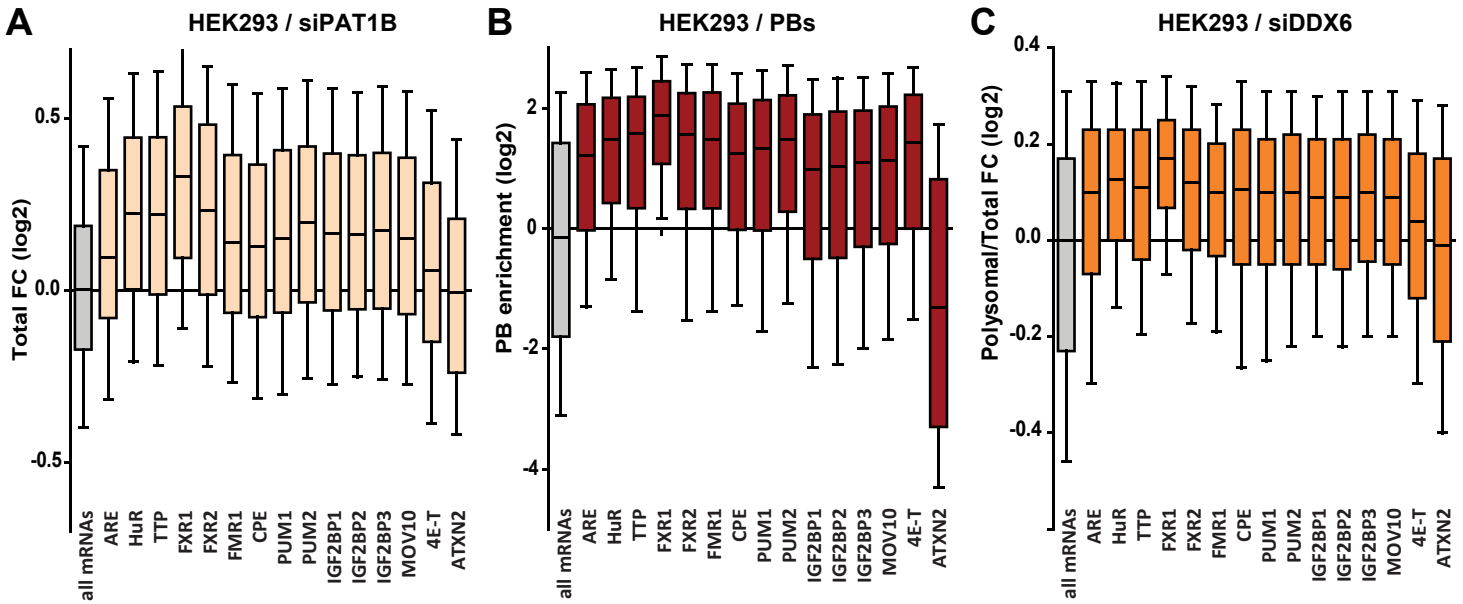

A
HEK293 / siDDX6
K562 / shDDX6
Total FC (log2)
all mRNAs
SMG6
YTHDF2
G4
TOP
B
HEK293 / PBs
PB enrichment (log2)
C
HEK293 / siDDX6
PB-out
PB-in
D
GC content (%)
Total FC_siDDX6 (log2)
E
PB enrichment (log2)

Figure 5



D

		HuR	TTP	FXR1-2	FMR1	CPEB	PUM1-2	IGF2BP1-3	MOV10	4E-T	ATXN2
siPat1b	Target Stability	up	up	up	up	up	up	up	up	+/-	-
siDDX6	Target Translation	up	up	up	up	up	up	up	up	+/-	-
siDDX6	Target Stability	-	-	-	-	-	-	-	-	-	-
DDX6 CLIP	Target binding	-	-	-	-	-	-	-	-	-	-
DDX6 interactome	Protein binding	+		+	+		+	+	+	+	+
PB localization	Protein	+			+	+	+	+	+	+	-
PB localization	Targets	+	+	+	+	+	+	+	+	+	-
Sequence	Targets	AU	AU	AU	AU	AU	AU	AU	AU	AU	AU GC

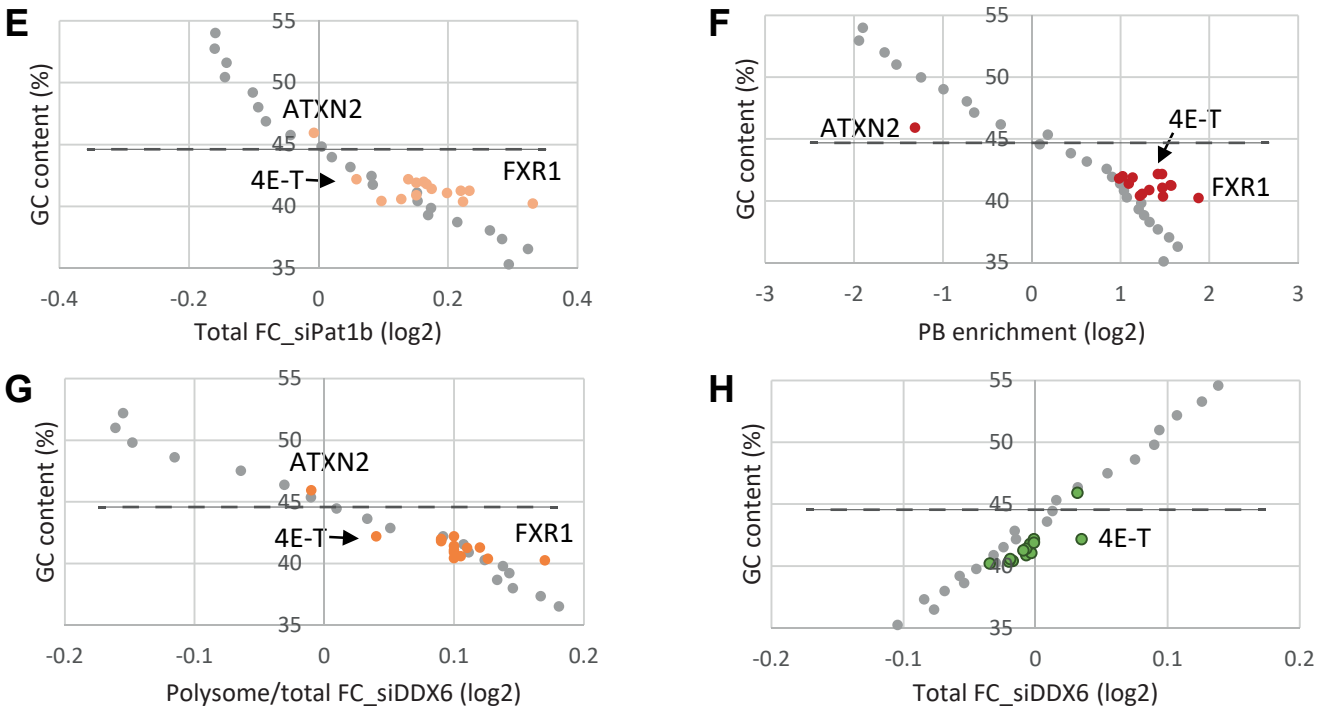


Figure 6

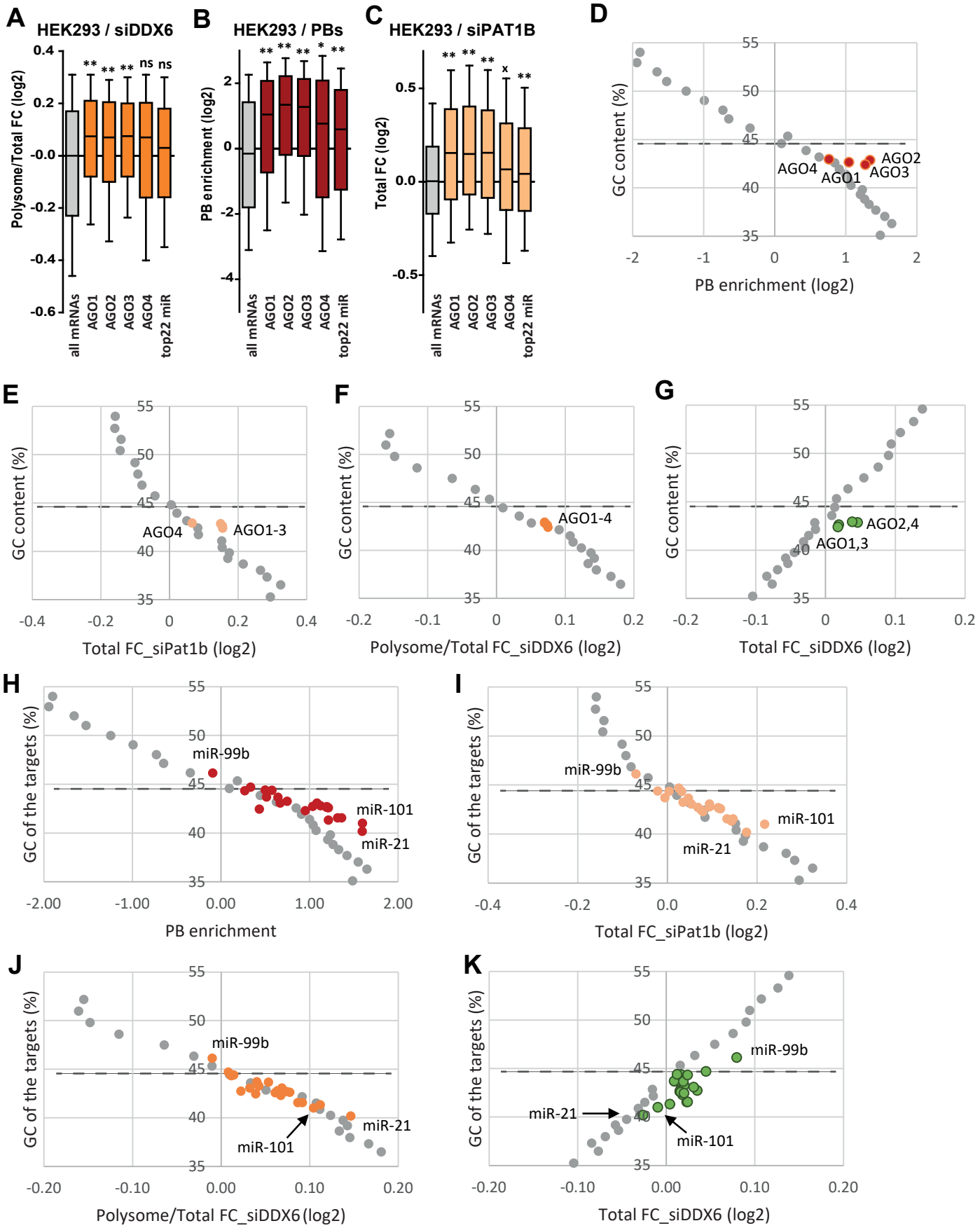
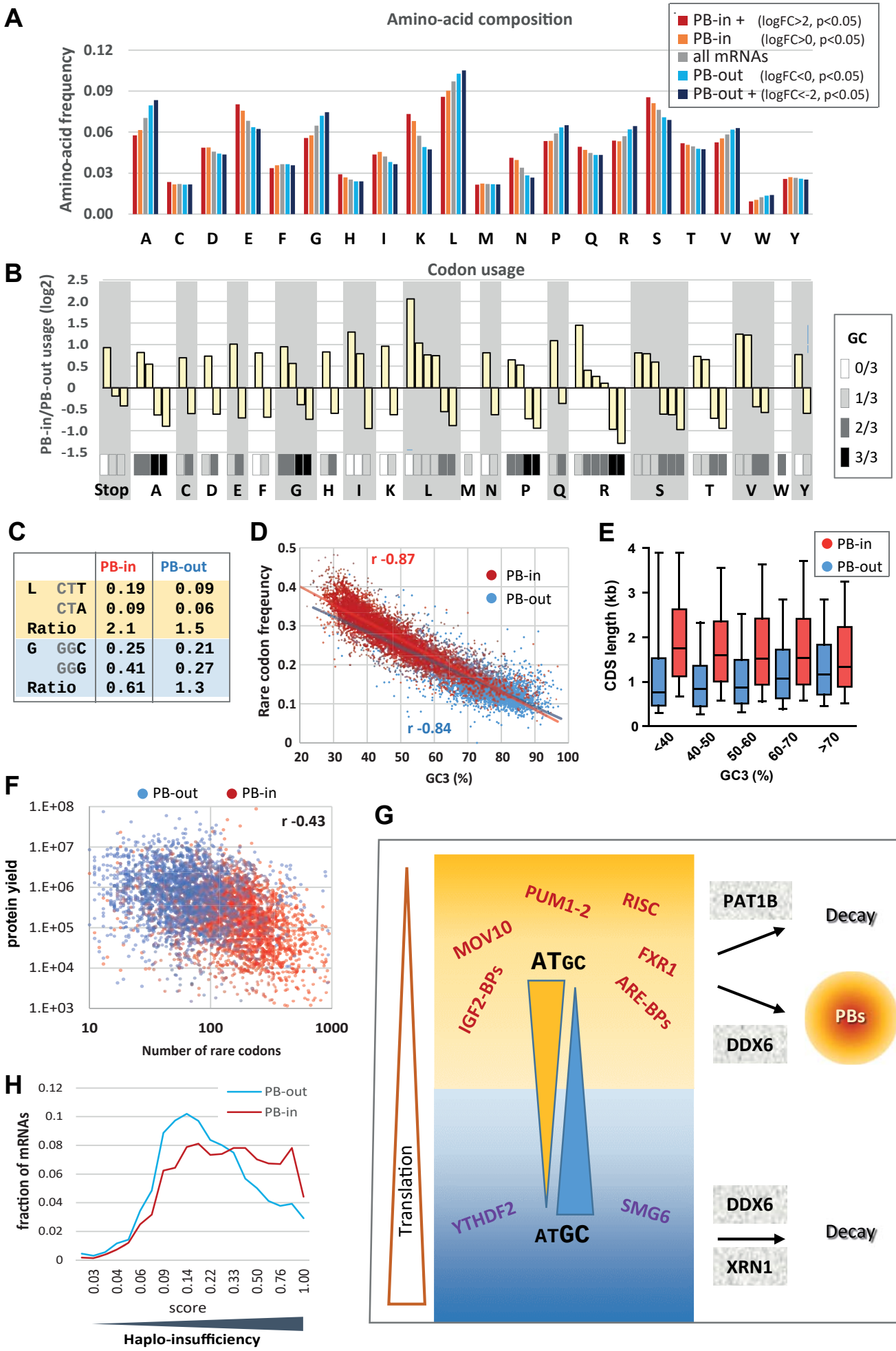

A
HEK293 / siDDX6
Polysome/Total FC (log2)
all mRNAs
AGO1
AGO2
AGO3
AGO4
top22 miR
B
HEK293 / PBs
PB enrichment (log2)
C
HEK293 / siPAT1B
Total FC (log2)
D
GC content (%)
PB enrichment (log2)
AGO4
AGO1
AGO2
AGO3
E
Total FC_siPat1b (log2)
AGO1-3
F
Polysome/Total FC_siDDX6 (log2)
AGO1-4
G
Total FC_siDDX6 (log2)
AGO2,4
AGO1,3
H
GC of the targets (%)
PB enrichment
miR-99b
miR-101
miR-21
I
Total FC_siPat1b (log2)
J
Polysome/Total FC_siDDX6 (log2)
K
Total FC_siDDX6 (log2)

Figure 7


A
Amino-acid composition
PB-in + (logFC>2, p<0.05)
PB-in (logFC>0, p<0.05)
all mRNAs
PB-out (logFC<0, p<0.05)
PB-out + (logFC<-2, p<0.05)
Amino-acid frequency
0.12
0.09
0.06
0.03
0.00
A C D E F G H I K L M N P Q R S T V W Y
B
Codon usage
PB-in/PB-out usage (log2)
2.5
2.0
1.5
1.0
0.5
0
-0.5
-1.0
-1.5
Stop A C D E F G H I K L M N P Q R S T V W Y
GC
0/3
1/3
2/3
3/3
C
PB-in
PB-out
L CTT 0.19 0.09
CTA 0.09 0.06
Ratio 2.1 1.5
G GGC 0.25 0.21
GGG 0.41 0.27
Ratio 0.61 1.3
D
r -0.87
PB-in
PB-out
r -0.84
Rare codon freqeuncy
0.5
0.4
0.3
0.2
0.1
0
20 30 40 50 60 70 80 90 100
GC3 (%)
E
PB-in
PB-out
CDS length (kb)
4
3
2
1
0
<40 40-50 50-60 60-70 >70
GC3 (%)
F
PB-out
PB-in
r -0.43
protein yield
1.E+08
1.E+07
1.E+06
1.E+05
1.E+04
1.E+03
10 100 1000
Number of rare codons
G
PUM1-2
RISC
MOV10
ATGC
FXR1
IGF2-BPs
ARE-BPs
PAT1B
Decay
PBs
DDX6
Translation
YTHDF2
SMG6
ATGC
DDX6
Decay
XRN1
H
PB-out
PB-in
fraction of mRNAs
0.12
0.1
0.08
0.06
0.04
0.02
0
0.03 0.04 0.06 0.09 0.14 0.22 0.33 0.50 0.76 1.00
score
Haplo-insufficiency