The DYRK1A protein kinase regulates ribosome mass and translation rates through transcriptional control of ribosomal protein genes expression

Chiara Di Vona[1,2]*, Laura Barba[1,2]*, Roberto Ferrari[3] and Susana de la Luna[1,2,4,5]

[1]Centre for Genomic Regulation (CRG), The Barcelona Institute of Science and Technology (BIST), Dr Aiguader 88, 08003 Barcelona, Spain.

[2]Centro de Investigación Biomédica en Red en Enfermedades Raras (CIBERER), Spain.

[3]Department of Chemistry, Life Sciences and Environmental Sustainability, University of Parma, Viale delle Scienze 23/A, 43124 Parma, Italy.

[4]Universitat Pompeu Fabra (UPF), Dr Aiguader 88, 08003 Barcelona, Spain.

[5]Institució Catalana de Recerca i Estudis Avançats (ICREA), Passeig Lluís Companys 23, 08010 Barcelona, Spain.

*, these authors contributed equally to this work

Corresponding authors: chiara.divona@crg.eu; susana.luna@crg.eu

Running title: DYRK1A regulates ribosomal protein expression

Key words: DYRK1A, ribosomal protein genes, transcription, ribosome, translation, cell growth.

ABSTRACT

Ribosomal proteins (RPs) are evolutionary conserved proteins that are essential for protein translation. RP expression must be tightly regulated, both to ensure the appropriate assembly of ribosomes and to respond to the growth demands of cells. The elements regulating the transcription of RP genes (RPGs) have been characterized in yeast and *Drosophila*, yet how cells regulate the production of RPs in mammals is less well understood. The dual-specificity tyrosine-regulated kinase 1A (DYRK1A) is known to participate in cell proliferation and differentiation in mammals, and its dysregulation is associated with disease in humans. Here, we show that DYRK1A marks a specific subset of proximal RPG promoters, which are characterized by the presence of the palindromic TCTCGCGAGA motif. The presence of DYRK1A at these promoters is associated with enhanced binding of the TATA-binding protein, TBP, and it is negatively correlated with the binding of the GABP transcription factor, establishing at least two clusters of RPGs that could be coordinately regulated. Indeed, DYRK1A depletion leads to a reduction of both RP mRNA and protein. Significantly, cells in which DYRK1A is depleted have fewer ribosomes, reduced global protein synthesis and they are smaller. Based on these results, we propose a novel role for DYRK1A in coordinating the expression of genes encoding RPs, thereby controlling cell growth in mammals.

INTRODUCTION

Ribosomes are cellular machines that translate mRNA into protein, and in mammals they are formed by the large 60S subunit and the small 40S subunit. The 60S subunit is comprised of the 5S, 5.8S and 28S rRNAs associated with 52 ribosomal proteins (RPs), and the smaller 40S subunit is made up of the 18S rRNA plus 35 RPs. Ribosome biogenesis is a complex process that involves more than 200 different factors: rRNAs, small nucleolar RNAs, and canonical and auxiliary RPs (Pena et al. 2017). The three RNA polymerases (Pol) participate in the transcription of the ribosomal components, with Pol I responsible for transcribing the 28S, 18S and 5.8S rRNAs, Pol III transcribing the 5S rRNAs and Pol II responsible for the transcription of all the protein coding genes involved in ribosome biogenesis, including the RP genes (RPGs). Therefore, the coordinated expression of these components is required to ensure the correct assembly and proper functioning of ribosomes. Indeed, dysregulation of ribosome biogenesis is associated with a group of human diseases that are collectively known as ribosomopathies (Aspesi and Ellis 2019) and moreover, alterations to RP expression contribute to cancer cell growth (Bustelo and Dosil 2018).

The coding sequences of RPGs have been highly conserved over evolution, unlike the features of their promoters and the machinery involved in their transcriptional regulation. As such, RPGs are organized into operons in prokaryotes (McGary and Nudler 2013), whereas the situation is much more complex in the case of eukaryotes, with multiple genes widely scattered across the genome (Gupta and Warner 2014). The main elements involved in the transcriptional regulation of RPGs have been characterized thoroughly in *Saccharomyces cerevisiae* (Bosio et al. 2017), in which the Repressor Activator Protein 1 (Rap1p) and the Fhl1p forkhead transcription factor (TF) are constitutively bound to the RPG promoters, coordinating RPG expression (Knight et al. 2014). In higher eukaryotes, most studies have focused on the differential enrichment of TF binding motifs within RPG promoters (Roepcke et al. 2006; Ma et al. 2009; Perina et al. 2011; Sleumer et al. 2012). In particular, several DNA sequences are found over-represented in human RPG promoters. The polypyrimidine TCT motif is found close to the transcription start site (TSS) of RPGs, and

it is thought to play a dual role in the initiation of both transcription and translation (Parry et al. 2010). This motif is recognized by the TATA-box Binding Protein (TBP)-Related Factor 2 (TRF2) in *Drosophila* (Wang et al. 2014; Baumann and Gilmour 2017), yet it remains unclear whether there is functional conservation with its human TBP like 1 (TBPL1) homolog. Around 35% of the RPG promoters contain a TATA-box in the -25 region and an additional 25% contain A/T-rich sequences in this region (Perry 2005). Other motifs frequently found are those for SP1, the GA-Binding Protein (GABP) and the Yin Yang 1 (YY1) TFs, the latter predominantly found downstream of the TSS in around 50% of all human RPGs (Perry 2005). In addition, the E-box TF MYC is a key regulator of ribosomal biogenesis, enhancing the expression of RPGs (van Riggelen et al. 2010). Finally, a *de novo* motif (M4 motif) was found enriched in human and mouse RPG promoters (Roepcke et al. 2006). RPG mRNA expression displays tissue- and development-specific patterns, both in human and mouse (Kondrashov et al. 2011; Gupta and Warner 2014; Panda et al. 2020). Hence, RPG expression could be regulated by specific combinations of TFs in different organisms and/or physiological conditions.

The M4 motif matches the palindromic sequence that is bound by the dual-specificity tyrosine-regulated kinase 1A (DYRK1A) protein kinase (Di Vona et al. 2015). DYRK1A fulfills many diverse functions by phosphorylating a wide range of substrates (Arbones et al. 2019; Boni et al. 2020) and it is a kinase with exquisite gene dosage dependency. On the one hand, DYRK1A overexpression in individuals with trisomy 21 has been associated to several of the pathological symptoms associated with Down syndrome (DS) (Stringer et al. 2017). On the other hand, *de novo* mutations in one *DYRK1A* allele cause a rare clinical syndrome known as DYRK1A haploinsufficiency syndrome (OMIM#614104) (van Bon et al. 2016; Arranz et al. 2019; and references therein). DYRK1A has also been proposed as a pharmacological target for neurodegenerative disorders, diabetes and cancer (Branca et al. 2017; Luna et al. 2019; Boni et al. 2020; Scavuzzo and Borowiak 2020). We have shown that DYRK1A is a transcriptional activator when recruited to proximal promoter regions of a subset of genes that are enriched for the palindromic motif TCTCGCGAGA (Di Vona et al.

2015). DYRK1A phosphorylates serine residues 2, 5 and 7 within the C-terminal domain (CTD) of the catalytic subunit of Pol II (Di Vona et al. 2015). The interaction of this kinase with the CTD depends on a run of histidine residues in its non-catalytic C-terminus, which also promotes the nucleation of a phase-separated compartment that is functionally associated with transcriptional elongation (Lu et al. 2018). Here, we have analyzed the occupancy of RPG promoters by DYRK1A in more depth, performing a comprehensive analysis of the promoter occupancy by other factors whose binding motifs are differentially enriched in human RPG promoter regions. Our results indicate that most of these factors are found at almost all RPG promoters, irrespective of the presence of their cognate binding sites. By contrast, DYRK1A associates with a subset of human and mouse RPG promoters that contain the TCTCGCGAGA motif. Physiological levels of DYRK1A are required to maintain RPG transcript levels, and this effect could at least in part contribute to the global reduction in ribosome mass and protein synthesis when DYRK1A is silenced. Therefore, our results expand the functional spectrum of the DYRK1A kinase, indicating that it contributes to the regulation of cell growth in mammalian cells.

RESULTS

DYRK1A is recruited to the proximal promoter regions of canonical RPGs

Our prior analysis of DYRK1A recruitment to chromatin showed an enrichment in gene ontology terms related to ribosome biogenesis and translational regulation (Di Vona et al. 2015). In addition, a *de novo* motif analysis found a sequence similar to the DYRK1A-associated motif to be over-represented in the promoter regions of human RPGs (Roepcke et al. 2006). Accordingly, we examined whether DYRK1A was recruited to human RPGs, considering the genes encoding the 80 canonical RPs plus 10 paralogues (Gupta and Warner 2014). In experiments on different human cell lines, including glioblastoma T98G, bone osteosarcoma U2OS and cervical adenocarcinoma HeLa cells (Fig. 1A and Supplemental Table S1), genome-wide chromatin immunoprecipitation followed by deep sequencing (ChIP-Seq) demonstrated DYRK1A bound upstream of the TSS and in general, within 500 bp of the TSS of these RPGs (Fig. 1A and Supplemental Fig. S1A). Moreover, no ChIP signal was detected at either the gene bodies or at the transcriptional termination sites (TTS: Fig. 1A, Supplemental Fig. S1A and S1B). To validate the recruitment of DYRK1A to RPG promoters, genome-wide DYRK1A occupancy was assessed in both control cells and in cells where the levels of DYRK1A were depleted by lentiviral delivery of a short hairpin RNA (shRNA) targeting DYRK1A. A clear reduction in chromatin-associated DYRK1A was observed at target RPG promoters, reflecting the specific depletion of this kinase (Fig. 1B and Supplemental Fig. S1A and S1B). Finally, the presence of DYRK1A at specific promoter regions of RPGs was further confirmed in independent ChIP-qPCR experiments (Fig. 1C).

Around 25% of the promoters of all RPGs were occupied by DYRK1A in the different human cell lines analyzed, with considerable overlap among them (Fig. 1D and 1E). No association with any particular ribosomal subunit was observed, since DYRK1A occupancy was detected similarly at the promoters of RPGs from both the large and small ribosome subunits (Fig. 1E). Furthermore, DYRK1A ChIP-Seq data from mouse embryonic stem cells (mESCs) also showed the presence of this kinase at the promoters of a subset of mouse

RPGs, which coincided well with the DYRK1A-positive subset in human cells (Fig. 1F and 1G, Supplemental Table S2). Together, these results indicate that DYRK1A is recruited to proximal promoter regions of a subset of RPGs in different human and mouse cell lines, suggesting that the association of DYRK1A with these RPGs at the chromatin level represents a general and conserved function for the kinase.

DYRK1A-bound RPG promoters are positive for the TCTCGCGAGA motif

Around 25% of the human RPG promoters contain a TCTCGCGAGA motif (Fig. 2A and 2B), and DYRK1A bound to these in one, two or in all the human cell lines analyzed, except for *RPL10* and *RPS5* and their paralogs *RPL10L*, *RPL22L1* and *RPS4Y2* that are not expressed in any of these cell lines (Supplemental Table S1). Likewise, DYRK1A was almost exclusively detected at RPG promoters containing the TCTCGCGAGA motif in mESCs (Supplemental Table S2), further evidence of its functional conservation. Therefore, the TCTCGCGAGA motif appears to be a distinctive feature of DYRK1A-binding RPG promoters in different cell lines.

A central model motif analysis (CMEA; Bailey and Machanick 2012) showed the TCTCGCGAGA motif to be positioned precisely around the center of the DYRK1A-associated peaks within the RPG promoters (Fig. 2C), suggesting it serves as a platform for DYRK1A recruitment. The kinase was also found to associate with a small number of RPG promoters without this palindromic sequence (Supplemental Table S1), although its detection was poorer in these cases than when the RPG promoter contained the TCTCGCGAGA motif (Fig. 2D). Hence, this palindromic sequence appears to be a major determinant for DYRK1A efficient recruitment to RPGs. Given that the variability in DYRK1A occupancy in the different cell lines was mostly restricted to the RPG promoters without the TCTCGCGAGA motif, it is possible that DYRK1A associates with promoters containing this motif independently of the cell-type, whereas occupancy of other RPG promoters might be context-specific.

The TCTCGCGAGA motif has been shown to work as a promoter element that drives bidirectional transcription (Mahpour et al. 2018). In around 30% of RPGs, the TSS lies within

a 500 bp window from the TSS of genes transcribed in the opposite direction and in these cases, the RPG is always transcribed more strongly (Supplemental Table S3). However, no specific bias for the presence of the TCTCGCGAGA motif could be observed, since some RPGs with bidirectional transcription contained this motif (*RPL12*, *RPL15*, *RPL23A*) while others did not (*RPL34*, *RPL9*, *RPS18*: Supplemental Table S3). The transcriptional repressor Zinc finger and BTB domain containing 33 (ZBTB33, also known as KAISO) also binds directly to the TCTCGCGAGA motif *in vitro* when methylated (Raghav et al. 2012). Indeed, this palindromic motif is included in the Jaspar database of curated TF binding profiles as a ZBTB33-motif (http://jaspar.genereg.net/matrix/MA0527.1/). ChIP-Seq experiments in T98G cells did detect ZBTB33 at the majority of RPG promoters containing this motif (Supplemental Fig. S2A and Supplemental Table S1), with signals overlapping those of DYRK1A (Supplemental Fig. S2B). Moreover, the TCTCGCGAGA motif is positioned centrally within the ZBTB33 bound regions of these RPGs (Supplemental Fig. S2C). Hence, this promoter motif might not only favor DYRK1A recruitment but also, its interaction with other proteins.

DYRK1A and GABP are differentially distributed at RPG promoters

Given the presence of DYRK1A at RPG promoters, we wondered whether the promoters that bind DYRK1A were also characterized by any other feature. Most studies on the regulation of RPG expression in higher eukaryotes have used *Drosophila* as a model system, identifying several sequence elements and TFs that regulate RPG transcription (Parry et al. 2010; Wang et al. 2014; Baumann and Gilmour 2017). One of them is the TCT motif that is a specific promoter element for the expression of RPGs (Parry et al. 2010). Using high confidence human TSS data, we scanned human RPG promoters and we found the TCT consensus sequence (YC+1TYTYY: Fig. 3A) in 77 of the 86 promoters analyzed (Supplemental Table S4). However, we did not find any association between the TCT motif and the presence of DYRK1A at these promoters (Fig. 3B). In *Drosophila*, the TCT element drives the recruitment of Pol II and consequently, the transcription of RPGs through the

coordinated action of TRF2 but not of TBP, and the TFs Motif 1 Binding Protein (M1BP) and DNA Replication-related Element (DRE) Factor (DREF) (Wang et al. 2014; Baumann and Gilmour 2017). Thus, we asked whether the transcriptional regulators of RPGs were functionally conserved between *Drosophila* and humans, and if so, how they might be related to the promoter association of DYRK1A. We analyzed the presence of the homologs of the fly TFs at human RPG promoters: TBPL1, the homolog of TRF2; Zinc Finger Protein 281 (ZNF281), the homolog of M1BP; and Zinc finger BED-type containing 1 (ZBED1), the homolog of DREF. This analysis showed that TBPL1 binds to the promoter of almost all RPGs (Supplemental Fig. S3A and Supplemental Table S4), similar to its behavior in *Drosophila* (Baumann and Gilmour 2017). ZNF281 was only detected at 7 RPG promoters (Supplemental Fig. S3A and S3B, Supplemental Table S4), although the sequence motif enriched within the ZNF281 ChIP-Seq dataset differed from the M1BP binding motif (Motif 1, Supplemental Fig. S3C; Ma et al. 2009) and thus, we cannot assume that M1BP and ZNF281 are fully functional homologs. ZBED1 binds to several human RPG promoters (Supplemental Fig. S3A and S3B, Supplemental Table S4), and the motif enriched in the ChIP-Seq dataset partially overlaps with the *Drosophila* DREF motif (http://jaspar.genereg.net/matrix/MA1456.1/, Supplemental Fig. S3D). However, no enrichment for this motif was found within the human RPG promoters. ZBED1 may recognize the TCTCGCGAGA motif (Yamashita et al. 2007), yet the CMEA analysis did not find a unimodal and centered distribution of the TCTCGCGAGA motif within the ZBED1 ChIP-peaks (Supplemental Fig. S3E). Hence, ZBED1 did not appear to bind directly to the TCTCGCGAGA motif, consistent with data suggesting that the TCTCGCGAGA motif and the human DRE are distinct regulatory elements (Mahpour et al. 2018). Finally, ZBED1 binding to human RPGs was not affected by the presence of DYRK1A (Supplemental Fig. S3F). Together, these data suggest there is little functional conservation of the core promoter elements of RPGs between *Drosophila* and humans, either in *cis* or *trans*. Furthermore, the association of the TFs homologs to those identified in *Drosophila* does not appear to depend

on the presence of the TCTCGCGAGA motif or the binding of DYRK1A to human RPG promoters.

We then asked whether DYRK1A recruitment was associated with the presence of TFs whose binding sites are known to be overrepresented at human RPG promoters, such as TBP, MYC, SP1, GABP and YY1. However, instead of using motif occurrence as in previous studies that analyzed RPG promoter architecture (Perry 2005; Perina et al. 2011), we took advantage of genome-wide mapping data for each of the TFs. Thus, though TBP was predicted to be differentially enriched at human RPG promoters based on the presence of the TATA-box (Perry 2005; Supplemental Table S4), the analysis of TBP occupancy found TBP bound to nearly all human RPG promoters (93%) in the different cell lines (Fig. 3C and 3D, Supplemental Table S4). Hence, TBP would appear to be a general component of the human RPG transcriptional machinery, irrespective of the presence of a TATA consensus, a TATA-like sequence or the complete absence of such motif (Fig. 3E). Consistent with a role for TBP in the assembly of the pre-initiation complex (PIC) at RPG promoters, the presence of the TFIID subunit TBP Associated Factor 1 (TAF1) was strongly correlated with that of TBP (Fig. 3C and Supplemental Table S4). Notably, we observed stronger TBP binding at DYRK1A-enriched RPG promoters than at RPG promoters devoid of DYRK1A (Fig. 3F and 3G), suggesting that these two factors might cooperate. The presence of YY1, SP1 and MYC was also detected in almost all RPG promoters (Supplemental Fig. S4A and Supplemental Table S4), and an unbiased clustering analysis showed no differential distribution of any of these factors based on the presence of DYRK1A (Supplemental Fig. S4B). By contrast, DYRK1A and GABP were distributed into one cluster that included those RPG promoters with high DYRK1A occupancy and low GABP presence (Fig. 3H, cluster 1), and another in which promoters were depleted of DYRK1A but there was a strong GABP occupation (Fig. 3H, cluster 2). Indeed, while the DYRK1A-associated TCTCGCGAGA motif was mostly overrepresented in cluster 1, the GABP-motif is a hallmark of cluster 2, and both motifs are distributed equally within the third cluster (Fig. 3I). These results suggest that the presence of

DYRK1A labels a specific subset of RPGs that might respond distinctly to those labeled by GABP.

DYRK1A regulates the expression of RPGs

Based on the ability of DYRK1A to regulate transcription when recruited to chromatin (Di Vona et al. 2015), we wondered whether the transcription of RPGs might be modulated by DYRK1A. Indeed, the TCTCGCGAGA motif regulates the expression of several RPGs, such as *RPL7A* (Colombo and Fried 1992), *RPS6*, *RPL10A*, *RPL12* (Yamashita et al. 2007) and *RPS11* (Di Vona et al. 2015). Furthermore, we demonstrated that the *RPS11* promoter responds to DYRK1A in a kinase- and motif-dependent manner (Di Vona et al. 2015). To evaluate the DYRK1A-dependent response of RPGs, we assessed whether the presence of DYRK1A has functional consequences on its target genes by profiling gene expression in different cell lines. Globally, the expression of genes that have promoters occupied by DYRK1A is stronger than that of the genes in which DYRK1A is absent (Fig. 4A), supporting a role for DYRK1A as a positive regulator of transcription (Di Vona et al. 2015). The same tendency towards stronger expression of DYRK1A bound targets was observed when assessing RPGs, although the differences were not statistically significant in any of the cell lines analyzed (Fig. 4B and Supplemental Fig. S5A and S5B). As described in *Drosophila* (Baumann and Gilmour 2017), the transcripts of most human RPGs were in the group of the top 5% most strongly transcribed genes in the cell lines analyzed (Supplemental Fig. S5C-S5E), which could dilute the significant differences in expression between DYRK1A positive and negative RPGs. ChIP-Seq data of Pol II for RPGs revealed a profile corresponding to actively transcribed genes, with Pol II occupancy detected all along gene bodies as an indication of transcriptionally engaged Pol II (Fig. 4C). Moreover, and consistent with the RNA data, there were no differences in the distribution of transcribing Pol II between the DYRK1A positive and negative RPGs (Fig. 4B and 4C).

We then assessed whether the absence of DYRK1A affected the expression of its target genes by comparing the RNA expression of cells infected with a control lentivirus or a

lentivirus expressing a shRNA to DYRK1A, and we used *Drosophila* RNA for spike-in normalization. Interestingly, the majority of differentially expressed genes were downregulated in response to DYRK1A depletion (Supplemental Fig. S5F). Indeed, these downregulated genes were enriched in the subset of genes with DYRK1A at their promoters (Supplemental Fig. S5G), supporting the role of DYRK1A in transcriptional activation. A general reduction in RPG transcripts was observed in DYRK1A-silenced cells (Fig. 4D), which affected RPGs whose promoters are occupied by DYRK1A and those without DYRK1A to similar extent (Fig. 4D and Supplemental Fig. S5H). The reduction in transcripts of selected target RPGs was validated by RT-qPCR using two different shRNAs directed against DYRK1A (Fig. 4E and Supplemental Fig. S5I). All these results indicated that DYRK1A modulates steady state levels of RPG mRNA transcripts through different mechanisms: on the one hand, at the transcriptional level through its recruitment to the proximal promoter regions of a subset of RPGs; and on the other hand, by an as yet unknown mechanism that may act at transcriptional and/or post-transcriptional levels.

DYRK1A depletion causes a reduction in ribosome content

We wondered whether the downregulation of RPGs transcripts levels induced by DYRK1A silencing would be reflected at the protein level. As such, we used mass spectrometry (MS) to quantify the ribosome composition in control cells and cells depleted in DYRK1A. A cytosolic fraction enriched in ribosomes was prepared from T98G cells by ultracentrifugation through a sucrose cushion and we analyzed two independent experiments with three biological replicates each, and considered for quantification only those proteins identified in at least three samples in any of the conditions (Supplemental Table S5). Our dataset had a strong overlap with those published in studies defining the human riboproteome (Reschke et al. 2013; Imami et al. 2018) (Supplemental Fig. S6A). Indeed, not only were the biological functions related to protein synthesis (RPs and other well-known translation-associated proteins) among those enriched but also, those related to oxidative phosphorylation, RNA transport/processing or endocytosis (Supplemental Fig. S6B and Supplemental Table S6),

probably reflecting co-sedimenting complexes. An enrichment in nuclear proteins related with splicing has also been described (Aviner et al. 2017). Nevertheless, core RPs represent more than 75% of the protein mass in the fraction analyzed (Fig. 5A).

The MS data allowed for the detection and quantification of all RPs except for the RPL10L and RPL39L paralogs, which showed low mRNA levels or not expressed at all in T98G cells, and for RPL41 that is usually not detected by MS approaches (Reschke et al. 2013). In addition, this quantification also indicated a lower relative abundance of the RPS27L, RPL3L, RPL7L1, RPL22L1, RPL26L1 and RPL36AL paralogs than their corresponding pair (Supplemental Fig. S6C and S6D and Supplemental Table S5), suggesting that they are under-represented in ribosomes from T98G cells. However, we cannot rule out that the variation in stoichiometry could be due to some RPs being loosely bound to the ribosome and thus, their presence may be affected by the method for cell extract preparation or pools of RPs performing extra-ribosomal functions might not be present in the cell fraction analyzed. We observed variability in the protein/mRNA ratio in T98G cells, with extreme cases for most of the weakly expressed paralogs like RPL3L, RPL7L1 or RPL22L1 (Supplemental Fig. S6F). Although we cannot rule out that some RPs are underrepresented in the fraction analyzed, the results are consistent with published data showing that the amounts of RPs correlate poorly with their corresponding mRNA levels in other cellular contexts (Franks et al. 2017). Our analysis did not reveal significant differences in the relative abundance of RPs encoded by DYRK1A-positive or DYRK1A-negative genes (Fig. 5B, Supplemental Fig. S6C and S6D).

We next focused on the alterations induced by silencing DYRK1A. The datasets revealed significant differences in the amount of protein associated with specific functional categories in the ribosomal-enriched fractions from control and DYRK1A depleted cells. Proteins in the oxidative phosphorylation category were enriched in the shDYRK1A cell fraction (Fig. 5C), including components of the mitochondrial electron transport chain (COX6B1, COX7A2L, SDHB) or mitochondrial ATP synthases (Supplemental Table S6). Conversely, the ribosome category was enriched in the protein dataset reduced in the

fraction from shDYRK1A cells (Fig. 5C and Supplemental Table S6). Indeed, DYRK1A depletion significantly reduced the total ribosome mass when calculated relative to the total amount of protein in the fraction analyzed by MS (Fig. 5A). In accordance with the effect of DYRK1A on RPG transcript levels (Fig. 4D and 4E and Supplemental Fig. S5I), less RP amounts were for RPGs containing or lacking DYRK1A at promoters (Fig. 5B and Supplemental Fig. S6C and S6D). Notably, rRNA content tended to fall (Supplemental Fig. S6G), supporting the decrease in ribosome amounts. The protein/mRNA ratios correlated strongly between control and DYRK1A silenced cells (Fig. 5D), suggesting that changes in protein abundance were largely due to the changes in RNA abundance upon DYRK1A depletion. Finally, differences in RP stoichiometry were observed between shControl and shDYRK1A cells (Fig. 5E), which might suggest differences in ribosome composition. All these results indicate that cells respond to DYRK1A depletion by reducing the steady state levels of RPs and that this effect occurs, at least in part, at the transcriptional level.

DYRK1A plays a role in cell size control by regulating protein synthesis

Altered ribosome biogenesis can lead to major defects in translation and thus, we assessed the impact of DYRK1A on the translational status of cells. The functional status of ribosomes was first analyzed by polysome profiling on sucrose gradients upon DYRK1A silencing. Downregulation of DYRK1A diminished the polysome fractions, which corresponded to those ribosomes engaged in active translation, with a concomitant increase in the monosome peak (Fig. 6A and Supplemental Fig. S7A). These results indicate that DYRK1A depletion leads to polysome disorganization. Indeed, there was a significant reduction of translational rates in DYRK1A-depleted cells when measured through radiolabeled ^{35}S-methionine incorporation (Fig. 6B). This effect was specific, as it was clearly observed with two distinct shRNAs targeted to DYRK1A (Fig. 6C). In eukaryotes, cell growth is coupled to cell cycle progression and therefore, global translation rates change during the cell cycle (Lloyd 2013). Defects in the cell cycle have been associated with alterations in DYRK1A levels in different cellular environments (Chen et al. 2013). Indeed, we found that DYRK1A silencing alters the cell

cycle balance in T98G cells, augmenting the population of cells in the G1 phase (Supplemental Fig. S7B). Hence, we next checked whether the shift in cell cycle phases was associated with the reduced translation rates. As such, T98G cells were arrested in G1 by serum deprivation (Supplemental Fig. S7C), which led to a strong reduction in the rate of translation (Fig. 6C). Reversing serum depletion for 30 min induced protein synthesis (Fig. 6C), although the cell cycle profiles were not changed (Supplemental Fig. S7C). In these conditions, lower rates of translation persisted in cells with silenced DYRK1A relative to control cells (Fig. 6C), suggesting that the reduction in protein synthesis upon DYRK1A silencing is independent of the alterations to the cell cycle.

As DYRK1A is a highly pleiotropic kinase, we wondered whether the reduction in translation could be due to an effect of DYRK1A on other signaling pathways that regulate protein synthesis. Therefore, we analyzed the effect of DYRK1A on one of the major signaling pathways that regulates protein translation, the mechanistic Target of Rapamycin (mTOR) pathway (Laplante and Sabatini 2012), assessing the Thr389 phosphorylation of the RPS6 kinase beta-1 (RPS6KB1 or p70S6K) that represents a late event in mTOR pathway activation. DYRK1A depletion did not alter Thr389-p70S6K phosphorylation (Fig. 6D and 6E) and likewise, the MS data showed no differences in the amount of RPS6 peptides phosphorylated at Ser235, Ser236 and Ser240 (Supplemental Fig. S7E), targets of p70S6K downstream of mTOR (Laplante and Sabatini 2012). Accordingly, DYRK1A does not appear to affect the mTOR pathway under regular growth conditions. Other signaling pathways like cellular stress and unfolded protein response can inhibit protein synthesis through the Ser51 phosphorylation of the translation initiation factor eIF2α (Koromilas 2015). However, the levels of Ser51-eIF2α phosphorylation remained unchanged in the absence of the DYRK1A (Supplemental Fig. S7F and S7G), indicating that DYRK1A-dependent inhibition of translation is not mediated by eIF2α phosphorylation. Moreover, DYRK1A was not detected in the ribosome enriched fraction by MS or in Western blots of polysome-associated fractions

(Supplemental Fig. S7H), suggesting that it is not tightly bound to actively translating ribosomes and that it probably does not act directly on polysomes.

Finally, as reduced protein synthesis might affect cell mass, we checked whether cell size was affected by the loss of DYRK1A activity. Indeed, DYRK1A-silenced HeLa and T98G cells were both significantly smaller than their controls in terms of cell volume (Fig. 6F and Supplemental Fig. S7I). Hence, the fine-tuning of cellular DYRK1A levels is important to assure the proper size of human cells is maintained.

DISCUSSION

In this study, we show that the protein kinase DYRK1A occupies the proximal promoter regions of a common subset of RPGs in different human and mouse cell lines. We also find that the RPGs whose promoters are occupied by DYRK1A are more strongly expressed than the subset devoid of DYRK1A, suggesting that DYRK1A could be required for maximal RPG expression in optimal growth conditions. In addition, DYRK1A could be part of signaling modules that respond by augmenting the expression of RPGs, such as Wnt signaling (Madan et al. 2018), to which DYRK1A activity has also been associated (Deshmukh et al. 2019; Granno et al. 2019). Interestingly, and despite the lack of conservation of the transcriptional regulatory networks between yeast and higher eukaryotes, a link between transcriptional regulation of RPGs and DYRK kinases has already been established, with the *S. cerevisiae* DYRK family member Yak1p acting downstream of TOR to repress RPG transcription (Martin et al. 2004).

The palindromic TCTCGCGAGA motif appears to be a distinctive feature of RPG promoters that interact with DYRK1A in different cell lines. A motif enrichment analysis did not find the TCTCGCGAGA motif within RPG promoters in yeast, basal metazoa or plants (Perina et al. 2011), although a similar motif (CGCGGCGAGACC) was found within the proximal promoter regions of 28 RPGs in *Caenorhabditis elegans* (Sleumer et al. 2012). In *Drosophila*, no similar motifs were enriched in RPG promoters (Ma et al. 2009) but nevertheless, the DRE sequence has been proposed to mirror such a motif (Yamashita et al. 2007). By contrast, the TCTCGCGAGA motif is conserved in vertebrates and it is generally found in DNAse I accessible regions, leading to the proposal that it serves as a core promoter element in TATA-less promoters associated with CpG islands (Mahpour et al. 2018). These findings would suggest that the TCTCGCGAGA motif is a *cis*-regulatory element that arose later in evolution and that it might be linked to co-evolution with regulatory factors acting in *trans*. Besides DYRK1A, the transcriptional repressor ZBTB33 also uses the TCTCGCGAGA motif as a putative chromatin recruitment platform (Raghav et al. 2012), and we confirm here its presence at RPG promoters that contain this motif. Whether DYRK1A

and ZBTB33, or other proteins recruited to chromatin via the TCTCGCGAGA motif, compete or collaborate in the regulation of common target genes, including RPGs, is an issue that merits further exploration.

In *Drosophila*, RPG expression is regulated by a combination of two TFs, M1BP and DREF, which are associated with distinct subsets of RPGs through binding to their corresponding DNA sequence motifs. These two proteins are responsible for recruiting TRF2, which substitutes for TBP in the assembly of the PIC (Parry et al. 2010; Wang et al. 2014; Baumann and Gilmour 2017). Our analysis shows no evidence of such a regulatory network in humans. On the one hand, we observe TBP binding to almost all RPG promoters, regardless of whether or not they contain a TATA sequence, which indicates that TBP would be responsible for PIC assembly in human RPGs, as supported by the presence of the largest subunit of TFIID, TAF1. It is worth noting that the presence of DYRK1A at RPG promoters is correlated with more TBP binding. It is therefore possible that DYRK1A assists in the recruitment of TBP to promote efficient TBP-driven transcription or to facilitate its positioning on the TSS. TBPL1, a TRF2 homolog, also binds to almost all RPGs, consistent with the finding that TBPL1 is recruited to the PIC, not as a substitute for TBP but rather along with it in mouse testis (Martianov et al. 2016). By contrast, we do not detect significant enrichment of the human M1BP and DREF homologs at human RPG promoters, ZNF281 and ZBED1, respectively. Therefore, a completely different set of regulators must exist in mammals to interpret, and to respond dynamically to growth and stress cues.

With the exception of the TCT motif, most of the sequences implicated in regulating the expression of RPGs in humans are only present in a subset of RPGs (Parry et al. 2010), including the binding motifs for the SP1, GABP, MYC and YY1 TFs. We have extended these results by analyzing the presence of these TFs at RPG promoter regions, revealing a more general distribution than that inferred through the presence of their binding sites. Therefore, no specific bias was found for the recruitment of TFs like SP1, YY1 or MYC to DYRK1A positive RPGs. By contrast, the RPGs that associate with DYRK1A are characterized by reduced GABP binding. The distinct distribution of DYRK1A and GABP could allow for

differential regulation of subsets of RPGs in response to a variety of stimuli. Notably, both *DYRK1A* and *GABP* are located on human chromosome 21 and when in trisomy, their overexpression might contribute to the general increase in RPG mRNA transcripts in the brain of individuals with DS (Supplemental Table S7).

Depletion of DYRK1A downregulates RPG transcripts and in terms of RPGs that interact with DYRK1A, this reduction could be a direct effect of the loss of DYRK1A at their proximal promoters and the subsequent reduction of Pol II CTD phosphorylation, as shown for *RPS11* (Di Vona et al. 2015). However, since the reduction in RPG expression occurs in a general manner, an additional and yet unknown DYRK1A-dependent activity may be at play. In any case, our results demonstrate that the production of RPs and ultimately, the number of ribosomes in proliferating cells, depends on the physiological levels of DYRK1A. The shortage of ribosomes upon a loss of DYRK1A provokes translational dysfunction. The reduction in ribosomes could either globally affect translation or it may result in transcript-specific translational control. Indeed, a general reduction in functional ribosomes in different physiological situations reduces the translation of some transcripts more than others (Mills and Green 2017). In this context, the pool of mRNAs associated with polysomes that respond differentially to DYRK1A downregulation still needs to be characterized. Together with the identification of *cis*-regulatory elements in these transcripts, this information will surely help to discriminate between the two possibilities and establish a mechanistic framework.

Although our findings point to a DYRK1A-dependent transcriptional effect in the regulation of the total ribosome content, we cannot rule out the existence of other DYRK1A-dependent effectors that contribute to altered ribosome biogenesis and/or translation. Our results rule out the impact of DYRK1A depletion on two major regulatory pathways, the stress pathway and the mTOR pathway, although other as yet unknown DYRK1A targets might also contribute to the phenotype observed. Further studies will help decipher how the effects of DYRK1A in the nucleus and cytoplasm converge to control cell growth.

Cell growth and proliferation, although separate processes, are intimately linked (Lloyd 2013; Roux and Topisirovic 2018). Since DYRK1A has been linked with the regulation of cell proliferation, this kinase could couple the cell cycle with protein synthesis by maintaining the amounts of RPs. In this regard, DYRK1A plays essential roles in central nervous system development, not only influencing cell numbers but also their differentiation (Arbones et al. 2019). Hence, the effect of DYRK1A on translation might contribute to its role as a regulator of neurite and axonal extension, events requiring post-mitotic cell growth. It is also possible that the effects of DYRK1A on cell mass may affect other tissues, particularly since heterozygous mouse models exhibit a global reduction in body size (Fotaki et al. 2002). In this regard, DYRK1A has been shown to phosphorylate Pol II at gene loci involved in myogenic differentiation (Yu et al. 2019). In a different context, defects in protein production are closely related to cancer since enhanced translation is required to boost cell proliferation (Bustelo and Dosil 2018; Aspesi and Ellis 2019). DYRK1A has both positive and negative effects on cell proliferation depending on the tumor context (Boni et al. 2020), although the impact of DYRK1A on translation rates might influence the tumor phenotype. Finally, mutations in specific RPGs produce very unique phenotypes (Aspesi and Ellis 2019), including craniofacial anomalies and urogenital malformations (Ross and Zarbalis 2014). In addition, *RPL10* mutations have been linked to neurodevelopmental conditions including autism spectrum disorders and microcephaly (Brooks et al. 2014), and indeed, translation is a process targeted in autism-associated disorders (Borrie et al. 2017). All these features are hallmarks of *DYRK1A* haploinsufficiency syndrome in humans or of animal models with *Dyrk1a* dysregulation (Arranz et al. 2019).

In summary, our findings indicate that DYRK1A influences cell growth by the transcriptional regulation of RPs expression.

METHODS

RP nomenclature

A new naming system for RPs has been proposed (Ban et al. 2014) and the equivalences to this nomenclature are listed in Supplemental Table S1.

Cell culture and lentivirus-mediated transduction

The HEK-293T, HeLa, U2OS and T98G cell lines were obtained from the American Type Culture Collection (www.atcc.org). E14 (129/Ola) mESCs were kindly provided by Maria Pia Cosma's laboratory (Centre for Genomic Regulation, Barcelona). Growth details in normal and serum starved conditions are included in the Supplemental Methods. Lentiviral transduction of shRNAs was used to downregulate DYRK1A expression, and the generation of the lentiviral stocks and the infection conditions are detailed in the Supplemental Methods. DYRK1A silencing was confirmed in Western blots (see Supplemental Methods for information and Supplemental Table S8 for antibodies). The protocols to determine the cell cycle profile and cell volume are also included in the Supplemental Methods.

To analyze global protein synthesis, T98G cells were incubated for 90 min in Met-free DMEM (GIBCO) with 10% dialyzed fetal bovine serum (FBS; GIBCO), they were metabolically labelled for 20 min with ^{35}S-Met (50 μCi ^{35}S-Met, 1175 Ci/mmol, Perkin Elmer) and then lysed in SDS-lysis buffer. The protein extracts were resolved by SDS-PAGE and the incorporation of ^{35}S-Met was detected by autoradiography of the dried gel using film or a Phosphoimager (Typhoon Trio, GE Healthcare).

Preparation of polysome and ribosome enriched fractions

Polysome profiles were obtained from approximately 1x10^7 T98G cells. Protein synthesis was arrested by incubation with cycloheximide (CHX 100 μg/ml: Sigma). The cells were washed in PBS containing CHX (100 μg/ml), collected in 1 ml of polysome lysis buffer (10

mM Tris-HCl [pH 7.4], 100 mM NaCl, 10 mM $MgCl_2$, 1% Triton X-100, 20 mM dithiothreitol [DTT], 100 μg/ml CHX, 0.25% sodium deoxycholate) and frozen rapidly in liquid nitrogen. Cell debris and nuclei were eliminated by centrifugation (12,000 *xg*, 5 min, 4 ºC) and the nucleic acid concentration in the supernatants was assessed by measuring the A_{260} in a NanoDrop™ (Thermo Fisher). Samples with A_{260}≈10 were loaded onto a 10-50% linear sucrose gradient prepared in polysome gradient buffer (20 mM Tris-HCl [pH 7.4], 100 mM NH_4Cl, 10 mM $MgCl_2$, 0.5 mM DTT, 100 μg/ml CHX), and centrifuged in a Beckman SW41Ti rotor (35,000 rpm, 3 h, 4 ºC). Profiles were obtained by continuous monitoring of the A_{254} (Econo-UV Monitor and Econo-Recorder model 1327; Bio-Rad). To calculate the polysome:monosome ratio, the polysome and monosome area under the curve was measured with ImageJ (1.50i) (Schneider et al. 2012).

To isolate the total ribosome fraction, cells were collected in sucrose buffer (250 mM sucrose, 250 mM KCl, 5 mM $MgCl_2$, 50 mM Tris-HCl [pH 7.4], 0.7% Nonidet P-40), the cytosol was isolated by centrifugation (750 x*g*, 10 min, 4 ºC), and then centrifuged again to obtain a post-mitochondrial supernatant (12,500 x*g*, 10 min, 4 ºC). The supernatant was adjusted to 0.5 M KCl, and the volume equivalent to OD_{260}=5 was loaded onto a sucrose cushion (1 M sucrose, 0.5 M KCl, 5 mM $MgCl_2$, 50 mM Tris-HCl [pH 7.4]) and centrifuged in a TLA 100.3 rotor (250,000 x*g*, 2 h, 4 ºC).

Mass spectrometry (MS) analysis

Proteins in the ribosome-enriched pellets were identified and quantified by free-label MS analysis using a LTQ-Orbitrap Fusion Lumos mass spectrometer. The sample preparation, chromatography and MS analysis are detailed in the Supplemental Methods. For peptide identification, a precursor ion mass tolerance of 7 ppm was used for MS1, with trypsin as the chosen enzyme and up to three miscleavages allowed. The fragment ion mass tolerance was set to 0.5 Da for MS2. Oxidation of methionine and N-terminal protein acetylation were used as variable modifications, whereas carbamidomethylation on cysteine was set as a

fixed modification. In the analysis of phosphorylated peptides, phosphorylation of serine, threonine and tyrosine were also set as variable modifications. False discovery rate (FDR) was set to a maximum of 5% in peptide identification. Protein quantification was retrieved from the protein TOP3 Area node from Proteome Discoverer (v2.3). For normalization, a correction factor was applied: sum TOP3 replicate “n”/average sum TOP3 all replicates. Based on the estimated normalized abundance, values were $\log_2$ transformed, and the fold change (FC) and *p*-values were calculated. Two independent experiments, each with three biological replicates, were performed on T98G cells transduced with shControl or shDYRK1A lentiviruses, and only those proteins detected in at least three replicates of any condition were quantified. To study RP stoichiometry, the intensity of each RP was defined relative to the intensities for all RPs. The RP protein/mRNA ratios were obtained using the $\log_2$ transformed normalized protein abundance and the $\log_2$ transformed normalized RNA counts from the RNA-seq experiments.

Chromatin immunoprecipitation (ChIP)

Detailed information on sample preparation is provided in the Supplemental Methods. DNA libraries were generated with the Ovation® Ultralow Library System V2 (NuGEN). Libraries were sequenced with 50 bp single end reads on an Illumina Hiseq-2500 sequencer at the CRG Genomics Unit. ChIP-Seq analysis was performed as described (Ferrari et al. 2012), with few modifications (see Supplemental Methods).To analyze RPG promoter occupancy, datasets from The Encyclopedia of DNA Elements (ENCODE) Consortium were used and are listed in Supplemental Table S9 (Consortium 2012; Gertz et al. 2013). Read numbers were normalized to a 1x reads per million (RPM) sequencing depth in both this work and the ENCODE datasets.

RNA-Seq

RNA was isolated with the RNeasy extraction kit and the samples were treated with DNase I (see Supplemental Methods for full details). For T98G cell spike-in normalization, equal

numbers of T98G cells for each condition were mixed with a fixed number of *D. melanogaster* Kc167 cells (1:4 ratio). Libraries were prepared with the TruSeq Stranded mRNA Sample Prep Kit v2 and sequenced with Illumina Hiseq-2500 to obtain 125 bp pair-ended reads. Differential gene expression was assessed with the DESeq2 package in R, filtering genes that had >10 average normalized counts per million (Love et al. 2014). For the spike-in libraries, the size factor of each replicate was calculated according to exogenous *Drosophila* spike-in reads. Expression was considered to be altered when p-value≤0.05, and the FC was above 0.7 and below -0.7 for up- and down-regulated genes, respectively.

Quantitative PCR (qPCR)

PCR reactions were performed in triplicate in 384 well plates with SYBR Green (Roche) and specific primers (Supplemental Tables S10 and S11), using a Roche LC-480 machine. The crossing point was calculated for each sample with the Lightcycler 480 1.2 software (Roche). No PCR products were observed in the absence of template and all primer sets gave narrow single melting point curves. For ChIP-qPCR, a 1/10 dilution of ChIP DNA was used as the template for the PCR reaction, and a 1/1000 dilution of input DNA was used as standard for normalization. For RT-qPCR, 1/10 dilution of the cDNAs was used and expression of the *D. melanogaster* gene *Act42A* was used for normalization.

Computational tools and statistical analysis

Full details for computational tools used to analyze the ChIP-Seq, RNA-Seq and proteomics data are included in the Supplemental Methods. To calculate statistical significance, the normality of the samples was evaluated with the Shapiro-Wilk normality test (Prism 5, v5.0d), and parametric or non-parametric tests were used accordingly. Statistical significance was calculated with two-tailed Mann-Whitney or Student's tests for unpaired samples or with a Wilcoxon matched pairs signed ranks test (Prism 5, v5.0d), and p-value≤0.05 considered as significant. All experiments were performed independently at least three times.

Data access

All the raw and processed sequencing data generated in this study have been submitted to the NCBI GEO repository (https://www.ncbi.nlm.nih.gov/geo/) under accession number GSE155809. The raw proteomics data have been submitted to the Proteomics Identifications Database (PRIDE; https://www.ebi.ac.uk/pride/archive/) under accession number PXD022966.

Competing interest statement

The authors have no competing interests to declare.

Acknowledgments

We thank all the members of Susana de la Luna's laboratory for their helpful discussions, Alicia Raya for technical assistance, the members of Juana Díez's laboratory (Molecular Virology lab, UPF) for assistance with the polysome profiling, Enrique Blanco (Epigenetics Events in Cancer Lab, CRG) for helpful discussions regarding the RNA-Seq analysis, Sarah Bonnin (CRG Bioinformatics Unit) for advice on the bioinformatics analysis, Aránzazu Rosado (Genome architecture Lab, CRG) for kindly providing the Kc167 Drosophila cells, Giorgio Dieci (Dept. of Chemistry, Life Sciences and Environmental Sustainability, University of Parma) for critical reading of the manuscript, and Mark Sefton for English language editing. The proteomics experiments were performed at the CRG/UPF Proteomics Unit, which is part of the Spanish Infrastructure for Omics Technologies (ICTS OmicsTech) and member of the ProteoRed PRB3 Consortium (supported by grant PT17/0019 from the Instituto de Salud Carlos III and ERDF), and we acknowledge the assistance of Eduard Sabidó and Eva Borras. Support from the CRG Genomics Facility is also acknowledged. L.B. was a FPU predoctoral fellow (FPU13/02400). This work has been supported by grants from the Spanish Ministry of Science and Innovation (BFU2016-76141-P, AEI/FEDER) and Secretaria d'Universitats i Recerca del Departament d'Empresa i Coneixement de la Generalitat de Catalunya (2017SGR1163). We also acknowledge the support of the Spanish

Ministry of Science and Innovation to the EMBL partnership, the Centro de Excelencia Severo Ochoa and the support of the CERCA Programme/Generalitat de Catalunya.

Author contributions:

S.d.l.L. supervised the project and acquired funding. C.d.V., L.B. and S.d.l.L. designed the experiments. C.d.V. and L.B. acquired the data. C.d.V., L.B., R.F. and S.d.l.L. analyzed the data. C.d.V. performed high throughput sequencing data analysis. S.d.l.L. wrote the manuscript together with C.d.V. All authors contributed critically and approved the final manuscript.

REFERENCES

- Arbones ML, Thomazeau A, Nakano-Kobayashi A, Hagiwara M, Delabar JM. 2019. DYRK1A and cognition: A lifelong relationship. *Pharmacol Ther* **194**: 199-221.

- Arranz J, Balducci E, Arato K, Sanchez-Elexpuru G, Najas S, Parras A, Rebollo E, Pijuan I, Erb I, Verde G et al. 2019. Impaired development of neocortical circuits contributes to the neurological alterations in DYRK1A haploinsufficiency syndrome. *Neurobiol Dis* **127**: 210-222.

- Aspesi A, Ellis SR. 2019. Rare ribosomopathies: insights into mechanisms of cancer. *Nat Rev Cancer* **19**: 228-238.

- Aviner R, Hofmann S, Elman T, Shenoy A, Geiger T, Elkon R, Ehrlich M, Elroy-Stein O. 2017. Proteomic analysis of polyribosomes identifies splicing factors as potential regulators of translation during mitosis. *Nucleic Acids Res* **45**: 5945-5957.

- Bailey TL, Machanick P. 2012. Inferring direct DNA binding from ChIP-seq. *Nucleic Acids Res* **40**: e128.

- Ban N, Beckmann R, Cate JH, Dinman JD, Dragon F, Ellis SR, Lafontaine DL, Lindahl L, Liljas A, Lipton JM et al. 2014. A new system for naming ribosomal proteins. *Curr Opin Struct Biol* **24**: 165-169.

- Baumann DG, Gilmour DS. 2017. A sequence-specific core promoter-binding transcription factor recruits TRF2 to coordinately transcribe ribosomal protein genes. *Nucleic Acids Res* **45**: 10481-10491.

- Boni J, Rubio-Perez C, Lopez-Bigas N, Fillat C, de la Luna S. 2020. The DYRK family of kinases in cancer: molecular functions and therapeutic opportunities. *Cancers (Basel)* **12**: 2106.

- Borrie SC, Brems H, Legius E, Bagni C. 2017. Cognitive dysfunctions in intellectual disabilities: the contributions of the Ras-MAPK and PI3K-AKT-mTOR pathways. *Annu Rev Genomics Hum Genet* **18**: 115-142.

- Bosio MC, Fermi B, Dieci G. 2017. Transcriptional control of yeast ribosome biogenesis: A multifaceted role for general regulatory factors. *Transcription* **8**: 254-260.

- Branca C, Shaw DM, Belfiore R, Gokhale V, Shaw AY, Foley C, Smith B, Hulme C, Dunckley T, Meechoovet B et al. 2017. Dyrk1 inhibition improves Alzheimer's disease-like pathology. *Aging Cell* **16**: 1146-1154.

- Brooks SS, Wall AL, Golzio C, Reid DW, Kondyles A, Willer JR, Botti C, Nicchitta CV, Katsanis N, Davis EE. 2014. A novel ribosomopathy caused by dysfunction of RPL10 disrupts neurodevelopment and causes X-linked microcephaly in humans. *Genetics* **198**: 723-733.

- Bustelo XR, Dosil M. 2018. Ribosome biogenesis and cancer: basic and translational challenges. *Curr Opin Genet Dev* **48**: 22-29.

- Chen JY, Lin JR, Tsai FC, Meyer T. 2013. Dosage of Dyrk1a shifts cells within a p21-cyclin D1 signaling map to control the decision to enter the cell cycle. *Mol Cell* **52**: 87-100.

- Colombo P, Fried M. 1992. Functional elements of the ribosomal protein L7a (rpL7a) gene promoter region and their conservation between mammals and birds. *Nucleic Acids Res* **20**: 3367-3373.

- Consortium EP. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**: 57-74.

- Deshmukh V, O'Green AL, Bossard C, Seo T, Lamangan L, Ibanez M, Ghias A, Lai C, Do L, Cho S et al. 2019. Modulation of the Wnt pathway through inhibition of CLK2 and DYRK1A by lorecivivint as a novel, potentially disease-modifying approach for knee osteoarthritis treatment. *Osteoarthritis Cartilage* **27**: 1347-1360.

- Di Vona C, Bezdan D, Islam AB, Salichs E, Lopez-Bigas N, Ossowski S, de la Luna S. 2015. Chromatin-wide profiling of DYRK1A reveals a role as a gene-specific RNA polymerase II CTD kinase. *Mol Cell* **57**: 506-520.

- Ferrari R, Su T, Li B, Bonora G, Oberai A, Chan Y, Sasidharan R, Berk AJ, Pellegrini M, Kurdistani SK. 2012. Reorganization of the host epigenome by a viral oncogene. *Genome Res* **22**: 1212-1221.

- Fotaki V, Dierssen M, Alcantara S, Martinez S, Marti E, Casas C, Visa J, Soriano E, Estivill X, Arbones ML. 2002. Dyrk1A haploinsufficiency affects viability and causes developmental delay and abnormal brain morphology in mice. *Mol Cell Biol* **22**: 6636-6647.

- Franks A, Airoldi E, Slavov N. 2017. Post-transcriptional regulation across human tissues. *PLoS Comput Biol* **13**: e1005535.

- Gertz J, Savic D, Varley KE, Partridge EC, Safi A, Jain P, Cooper GM, Reddy TE, Crawford GE, Myers RM. 2013. Distinct properties of cell-type-specific and shared transcription factor binding sites. *Mol Cell* **52**: 25-36.

- Granno S, Nixon-Abell J, Berwick DC, Tosh J, Heaton G, Almudimeegh S, Nagda Z, Rain JC, Zanda M, Plagnol V et al. 2019. Downregulated Wnt/beta-catenin signalling in the Down syndrome hippocampus. *Sci Rep* **9**: 7322.

- Gupta V, Warner JR. 2014. Ribosome-omics of the human ribosome. *RNA* **20**: 1004-1013.

Imami K, Milek M, Bogdanow B, Yasuda T, Kastelic N, Zauber H, Ishihama Y, Landthaler M, Selbach M. 2018. Phosphorylation of the ribosomal protein RPL12/uL11 affects translation during mitosis. *Mol Cell* **72**: 84-98 e89.

- Knight B, Kubik S, Ghosh B, Bruzzone MJ, Geertz M, Martin V, Denervaud N, Jacquet P, Ozkan B, Rougemont J et al. 2014. Two distinct promoter architectures centered on dynamic nucleosomes control ribosomal protein gene transcription. *Genes Dev* **28**: 1695-1709.

- Kondrashov N, Pusic A, Stumpf CR, Shimizu K, Hsieh AC, Xue S, Ishijima J, Shiroishi T, Barna M. 2011. Ribosome-mediated specificity in Hox mRNA translation and vertebrate tissue patterning. *Cell* **145**: 383-397.

- Koromilas AE. 2015. Roles of the translation initiation factor eIF2alpha serine 51 phosphorylation in cancer formation and treatment. *Biochim Biophys Acta* **1849**: 871-880.

- Laplante M, Sabatini DM. 2012. mTOR signaling in growth control and disease. *Cell* **149**: 274-293.

- Lloyd AC. 2013. The regulation of cell size. *Cell* **154**: 1194-1205.

- Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550.

- Lu H, Yu D, Hansen AS, Ganguly S, Liu R, Heckert A, Darzacq X, Zhou Q. 2018. Phase-separation mechanism for C-terminal hyperphosphorylation of RNA polymerase II. *Nature* **558**: 318-323.

- Luna J, Boni J, Cuatrecasas M, Bofill-De Ros X, Nunez-Manchon E, Gironella M, Vaquero EC, Arbones ML, de la Luna S, Fillat C. 2019. DYRK1A modulates c-MET in pancreatic ductal adenocarcinoma to drive tumour growth. *Gut* **68**: 1465-1476.

- Ma X, Zhang K, Li X. 2009. Evolution of *Drosophila* ribosomal protein gene core promoters. *Gene* **432**: 54-59.

- Madan B, Harmston N, Nallan G, Montoya A, Faull P, Petretto E, Virshup DM. 2018. Temporal dynamics of Wnt-dependent transcriptome reveal an oncogenic Wnt/MYC/ribosome axis. *J Clin Invest* **128**: 5620-5633.

- Mahpour A, Scruggs BS, Smiraglia D, Ouchi T, Gelman IH. 2018. A methyl-sensitive element induces bidirectional transcription in TATA-less CpG island-associated promoters. *PLoS One* **13**: e0205608.

- Martianov I, Velt A, Davidson G, Choukrallah MA, Davidson I. 2016. TRF2 is recruited to the pre-initiation complex as a testis-specific subunit of TFIIA/ALF to promote haploid cell gene expression. *Sci Rep* **6**: 32069.

- Martin DE, Soulard A, Hall MN. 2004. TOR regulates ribosomal protein gene expression via PKA and the Forkhead transcription factor FHL1. *Cell* **119**: 969-979.

- McGary K, Nudler E. 2013. RNA polymerase and the ribosome: the close relationship. *Curr Opin Microbiol* **16**: 112-117.

- Mills EW, Green R. 2017. Ribosomopathies: There's strength in numbers. *Science* **358**: eaan2755.

- Panda A, Yadav A, Yeerna H, Singh A, Biehl M, Lux M, Schulz A, Klecha T, Doniach S, Khiabanian H et al. 2020. Tissue- and development-stage-specific mRNA and heterogeneous CNV signatures of human ribosomal proteins in normal and cancer samples. *Nucleic Acids Res* **48**: 7079-7098.

- Parry TJ, Theisen JW, Hsu JY, Wang YL, Corcoran DL, Eustice M, Ohler U, Kadonaga JT. 2010. The TCT motif, a key component of an RNA polymerase II transcription system for the translational machinery. *Genes Dev* **24**: 2013-2018.

- Pena C, Hurt E, Panse VG. 2017. Eukaryotic ribosome assembly, transport and quality control. *Nat Struct Mol Biol* **24**: 689-699.

- Perina D, Korolija M, Roller M, Harcet M, Jelicic B, Mikoc A, Cetkovic H. 2011. Over-represented localized sequence motifs in ribosomal protein gene promoters of basal metazoans. *Genomics* **98**: 56-63.

- Perry RP. 2005. The architecture of mammalian ribosomal protein promoters. *BMC Evol Biol* **5**: 15.

- Raghav SK, Waszak SM, Krier I, Gubelmann C, Isakova A, Mikkelsen TS, Deplancke B. 2012. Integrative genomics identifies the corepressor SMRT as a gatekeeper of adipogenesis through the transcription factors C/EBPbeta and KAISO. *Mol Cell* **46**: 335-350.

- Reschke M, Clohessy JG, Seitzer N, Goldstein DP, Breitkopf SB, Schmolze DB, Ala U, Asara JM, Beck AH, Pandolfi PP. 2013. Characterization and analysis of the composition and dynamics of the mammalian riboproteome. *Cell Rep* **4**: 1276-1287.

- Roepcke S, Zhi D, Vingron M, Arndt PF. 2006. Identification of highly specific localized sequence motifs in human ribosomal protein gene promoters. *Gene* **365**: 48-56.

- Ross AP, Zarbalis KS. 2014. The emerging roles of ribosome biogenesis in craniofacial development. *Front Physiol* **5**: 26.

- Roux PP, Topisirovic I. 2018. Signaling pathways involved in the regulation of mRNA translation. *Mol Cell Biol* **38**: e00070-00018.

- Scavuzzo MA, Borowiak M. 2020. Two drugs converged in a pancreatic beta cell. *Sci Transl Med* **12**: eaba7359.

- Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* **9**: 671-675.

- Sleumer MC, Wei G, Wang Y, Chang H, Xu T, Chen R, Zhang MQ. 2012. Regulatory elements of *Caenorhabditis elegans* ribosomal protein genes. *BMC Genomics* **13**: 433.

- Stringer M, Goodlett CR, Roper RJ. 2017. Targeting trisomic treatments: optimizing Dyrk1a inhibition to improve Down syndrome deficits. *Mol Genet Genomic Med* **5**: 451-465.

- van Bon BW, Coe BP, Bernier R, Green C, Gerdts J, Witherspoon K, Kleefstra T, Willemsen MH, Kumar R, Bosco P et al. 2016. Disruptive de novo mutations of DYRK1A lead to a syndromic form of autism and ID. *Mol Psychiatry* **21**: 126-132.

- van Riggelen J, Yetil A, Felsher DW. 2010. MYC as a regulator of ribosome biogenesis and protein synthesis. *Nat Rev Cancer* **10**: 301-309.

- Wang YL, Duttke SH, Chen K, Johnston J, Kassavetis GA, Zeitlinger J, Kadonaga JT. 2014. TRF2, but not TBP, mediates the transcription of ribosomal protein genes. *Genes Dev* **28**: 1550-1555.

- Yamashita D, Sano Y, Adachi Y, Okamoto Y, Osada H, Takahashi T, Yamaguchi T, Osumi T, Hirose F. 2007. hDREF regulates cell proliferation and expression of ribosomal protein genes. *Mol Cell Biol* **27**: 2003-2013.

- Yu D, Cattoglio C, Xue Y, Zhou Q. 2019. A complex between DYRK1A and DCAF7 phosphorylates the C-terminal domain of RNA polymerase II to promote myogenesis. *Nucleic Acids Res* **47**: 4462-4475.

FIGURE LEGENDS

Figure 1. DYRK1A occupies a subset of RPG promoters. *(A)* Density plot showing the distribution of DYRK1A in T98G, U2OS and HeLa cells relative to the TSS of all human RPGs. The *y*-axis represents the relative protein recruitment (-$\log_{10}$ Poisson *p*-value) and the offset was set to ± 3 kb from the TSS. *(B)* Density plot showing DYRK1A occupancy relative to the TSS of all human RPGs, comparing shControl and shDYRK1A T98G cells (blue and orange lines, respectively). The *y*-axis represents the relative protein recruitment quantified as significant (sig) ChIP-Seq tags. The offset was set to ± 3 kb from the TSS (see Supplemental Fig. S1B for representative examples). *(C)* Validation of selected targets by ChIP-qPCR, representing the data as a percentage of the input recovery (mean ± SD of three technical replicates). *(D)* Venn diagram showing the overlap of DYRK1A-associated RPG promoters in T98G, U2OS and HeLa cells. *(E)* List of DYRK1A-positive RPGs common to the three human cell lines. *(F)* Venn diagram indicating the overlap between common DYRK1A-bound RPGs in the human cell lines and mESCs. *(G)* List of RPGs with DYRK1A detected at their promoters in mESCs. The asterisk indicates the coincident occupancy in the three human cell lines analyzed.

Figure 2. The TCTCGCGAGA motif correlates with significant DYRK1A binding at human RPG promoters. *(A)* Distribution of the TCTCGCGAGA motif (DYRK1A motif) in human RPGs based on the *q*-value (see Methods). The DYRK1A motif associated logo is included. *(B)* List of human RPGs with the *bona-fide* palindromic motif detected in their promoters. The RPG promoters containing more than one DYRK1A motif are highlighted in bold. *(C)* CentriMo plot showing the distribution of the DYRK1A motif for the RPG promoters that bind DYRK1A. The solid curve shows the positional distribution (averaged over bins of 10 bp width) of the best site of the DYRK1A motif at each position in the RPG-ChIP-Seq peaks (500 bp). The *p*-value for the central enrichment of the motif is included. *(D)* Box plot and scatter plots (dots represent individual RPGs) showing the correlation between DYRK1A

ChIP binding ($\log_2$ *p*-value, *x*-axis) and the conservation of the DYRK1A motif (*q*-value, *y*-axis) in T98G, U2OS and HeLa cells. The colors represent the motif classification indicated.

Figure 3. Differential recruitment of DYRK1A and GABP to RPG promoters. *(A)* Sequence logo of the human TCT motif found in human RPGs (see Methods). *(B)* Percentage of TCT positive or negative human RPGs promoters distributed according to the presence of DYRK1A (see also Supplemental Table S4). *(C)* The heat maps show TBP and TAF1 (K562 cells, ENCODE data) occupancy in all human RPGs, showing the genomic region considered on the *x*-axis. *(D)* Venn diagram indicating the overlap of TBP occupancy at RPGs in different human cell lines (ENCODE data). *(E)* Relationship of the presence of a TATA-box or TA-like sequences in human RPG promoters according to (Perry 2005) and TBP binding in HeLa cells (ND, not determined). *(F)* Unbiased *k*-mean clustering of the average binding of DYRK1A (HeLa, this work) and TBP (HeLa, ENCODE data) on all known human RPGs. The color scale bar indicates the binding score and the genomic region considered is shown on the *x*-axis. *(G)* Metagene plot showing TBP occupancy relative to all human RPGs in HeLa cells according to the clusters shown in Fig. 3F. The *y*-axis represents relative protein recruitment quantified as ChIP-Seq reads. *(H)* Unbiased *k*-mean clustering of the average binding of DYRK1A (T98G, this work), GABP and YY1 (SK-N-SH, ENCODE data) to all RPGs. The color scale bar indicates the binding score and the genomic region considered is shown on the *x*-axis. *(I)* The histogram represents the percentage of RPG promoters containing a DYRK1A or a GABP motif in each of the clusters shown Fig. 3H.

Figure 4. Depletion of DYRK1A causes a general reduction in RPG expression. *(A)* Box plot indicating the RNA levels of all human genes expressed in T98G cells (normalized counts) separated into two clusters depending on the presence of DYRK1A at their promoters (unpaired, two-tailed Student's test). *(B)* Box plot indicating the RNA levels of all the human RPGs (normalized counts) separated into two clusters, as in Fig. 4A (unpaired, two-tailed, Mann-Whitney test, ns=not significant). *(C) Bottom panel*, heat map showing

DYRK1A and Pol II chromatin occupancy relative to human RPGs depicted as metagenes, and separated into two clusters according to the presence/absence of DYRK1A at the promoters. The binding score (-$\log_{10}$ Poisson *p*-value) is indicated by the color scale bar. *Top panel*, density plot corresponding to the mean value of the heat map (blue and light blue line correspond to DYRK1A positive or DYRK1A negative RPGs, respectively). The binding score (-$\log_{10}$ Poisson *p*-value) is indicated on the *y*-axis. *(D)* Box plot indicating the expression of RPGs (normalized counts) in T98G cells classified according to the presence (positive) or absence (negative) of DYRK1A at their promoters, and comparing cells infected with a lentivirus expressing a shRNA Control (blue and light blue, respectively) or shDYRK1A (orange and light orange, respectively: Wilcoxon matched-pairs signed ranks test, *** $p<0.00001$). The reduction in DYRK1A levels is shown in Supplemental Fig. S5J. *(E)* RT-qPCR validation of differentially expressed RPGs in T98G cells transduced with shControl or shDYRK1A lentiviruses. DYRK1A and one translation-related non-RPG factor (*EIF4E2*) RNA levels are also shown (see also Supplemental Fig. S5I). The data was corrected by *Drosophila* spike in values and represented as the RNA levels relative to the shControl arbitrarily set as 1 (mean ± SD of independent experiments: Mann-Whitney test ***$p\leq0.001$, **$p\leq0.01$, *$p\leq0.05$, ns=not significant).

Figure 5. DYRK1A depletion causes a reduction in ribosome content. *(A)* The relative RP values (sum of all RP TOP3 values/sum all TOP3 values) for each sample of shControl (blue) and shDYRK1A (brown) T98G cells represented, also showing the median and IQRs (n=6, two-tailed Mann-Whitney test). *(B)* Box plot indicating the levels of the RPs ($\log_2$ of normalized peptide intensities) separated into two clusters according to the presence of DYRK1A at the promoters of their corresponding genes, both in shDYRK1A and shControl T98G cells (Wilcoxon matched-pairs signed rank test: ***$p<0.0001$). The reduction in DYRK1A is shown in Supplemental Fig. S6E. *(C)* Functional categorization of the proteins found more abundant in the shDYRK1A cells (UP) or in the shControl cells (DOWN) (proteins

with *p*-value<0.05 in the comparisons or present in only one of the conditions were used). The number of proteins identified in each category is shown. See also Supplemental Table S6. *(D, E)* Correlation analysis of the ratio of RP protein and mRNA abundances (*D*) and RP stoichiometry (*E*) of shControl and shDYRK1A T98G cells. A color code was used to indicate the presence (+) / absence (-) of DYRK1A at the RPG promoter regions. The Spearman's correlation coefficient is shown for each subset. In *E*, the value for the adjusted R^2 for each set is also included.

Figure 6. DYRK1A controls cell size by regulating protein synthesis. (*A*) Polysome profile obtained by sucrose gradient sedimentation of T98G cells in which DYRK1A is down-regulated via lentiviral delivery of shRNAs (shDYRK1A, orange line) and relative to the control cells (shControl, blue line). The position of the 40S, 60S, 80S and polysome peaks is indicated. The *y*-axis shows absorbance at 254 nm in arbitrary units and the *x*-axis corresponds to the different fractions, with the top of the gradient to the left. See Supplemental S7A for the quantification of the polysome:monosome ratios in independent experiments. (*B*) Global protein synthesis assays were performed by metabolic ^{35}S-methionine labeling in shDYRK1A.1 or shControl T98G cells, visualizing the incorporation of the labeled amino acid by SDS-PAGE of the total cell extracts and autoradiography. The presence of DYRK1A was analyzed in Western blots using tubulin as a loading control. Quantification of the average radioactive intensity of independent experiments is shown at the bottom of the images, with the values of the shControl condition set as 1 (mean ± SEM, n=3, Wilcoxon matched-pairs signed ranks test, *p*<0.001). (*C*) Autoradiography of protein extracts prepared from proliferating, serum starved cells for 48 h (SS), or serum starved cells re-incubated with FBS for 30 min (SS+FBS) and pulse-labeled with ^{35}S-Met for 20 min. Equal numbers of T98G cells infected with the indicated shRNA-lentivectors were used. The reduction in DYRK1A was assessed in Western blots using vinculin as a control and a Coomassie stained gel as a loading control is shown in Supplemental Fig. S7D. Cell cycle profiles are shown in Supplemental Fig. S7C. (*D, E*) Analysis of the Thr389 phosphorylation

of p70S6K (p-p70S6K) assessed in Western blots of shControl or shDYRK1A T98G cells. A representative experiment (*D*) and the quantification of the p-p70S6K signal relative to the total p70S6K (*E*) are shown (mean ± SD, n = 4, Wilcoxon matched-pairs signed ranks test, ns=not significant). The dots and lines show independent experiments. (*F*) Cell volume represented in arbitrary units with the control cells set as 1 (mean ± SEM; Mann-Whitney test, **p*<0.05, ***p*<0.01). See Supplemental Fig. S7G for the results with two different shRNAs against DYRK1A in T98G cells.

Figure 1
Di Vona, Barba *et al*.

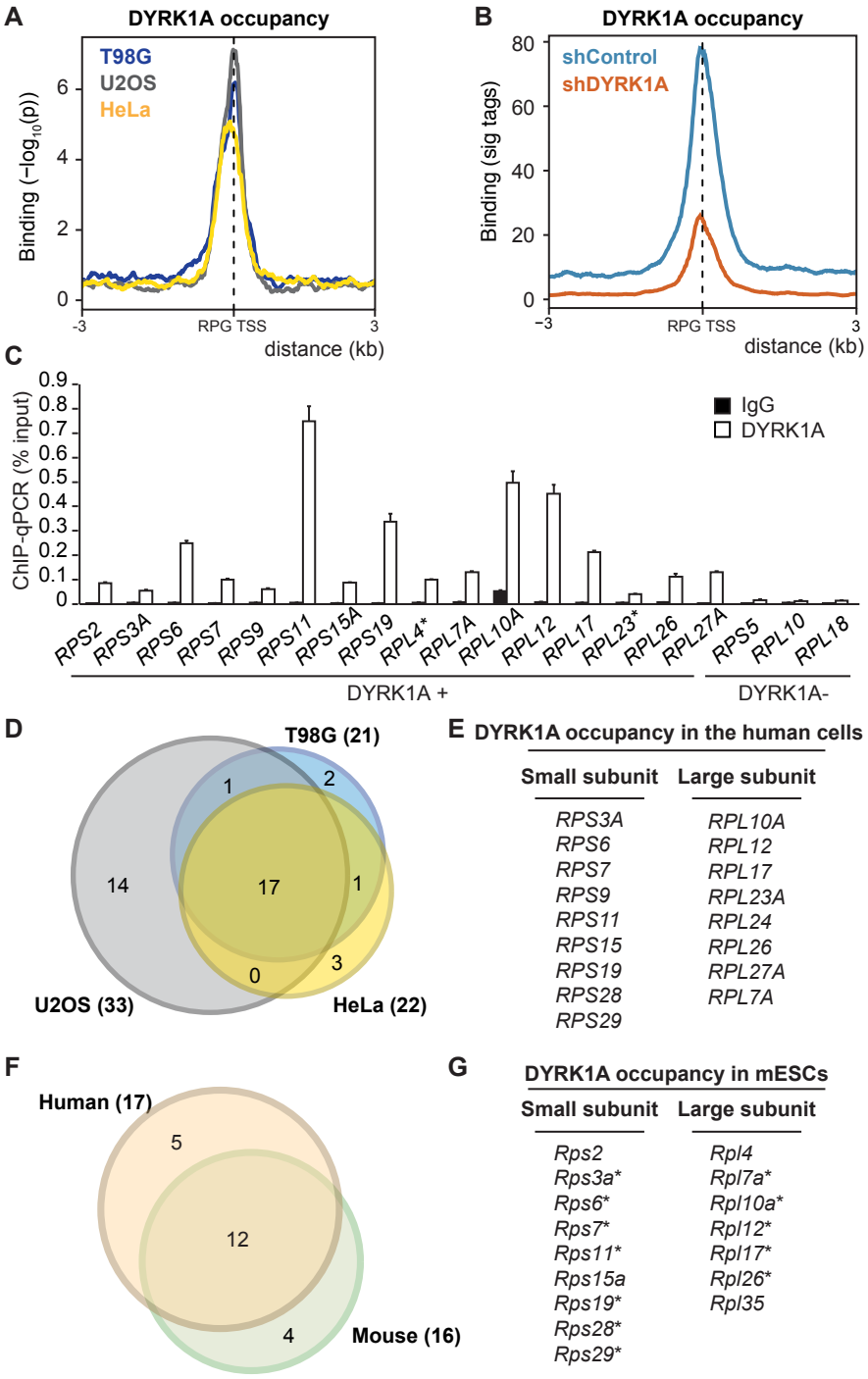

Figure 2
Di Vona, Barba *et al*.

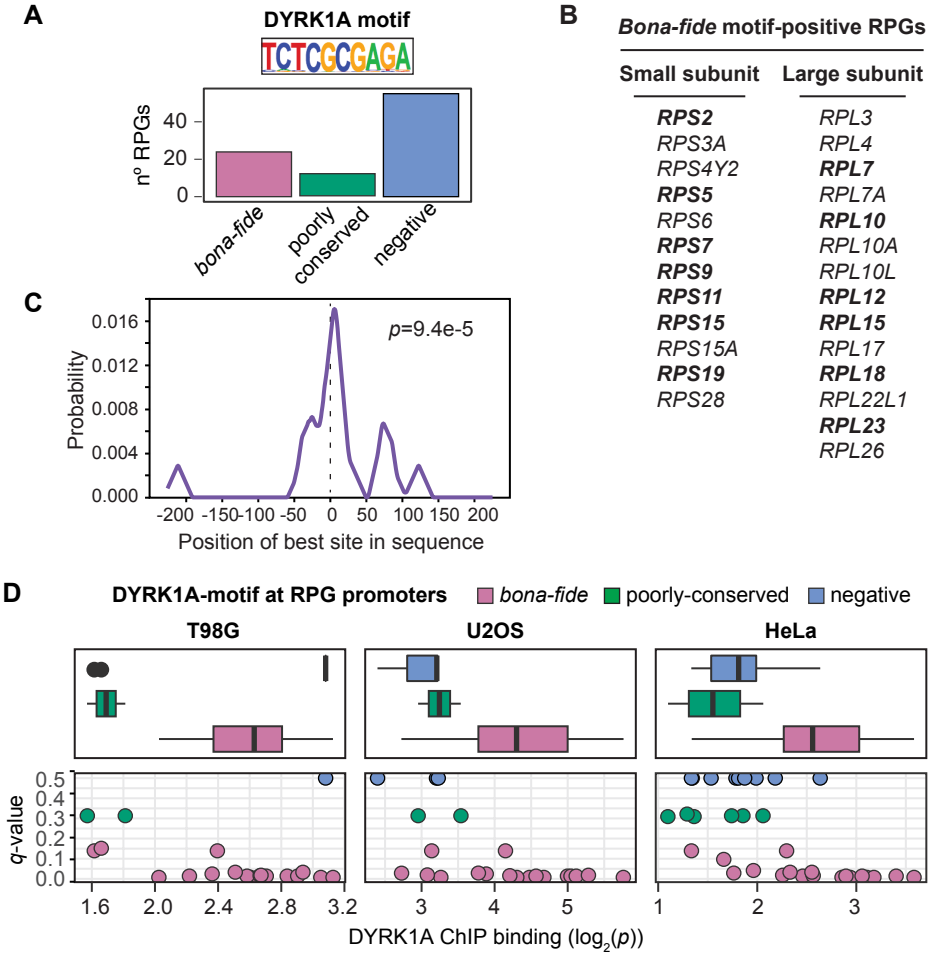

Figure 3
Di Vona, Barba *et al*.

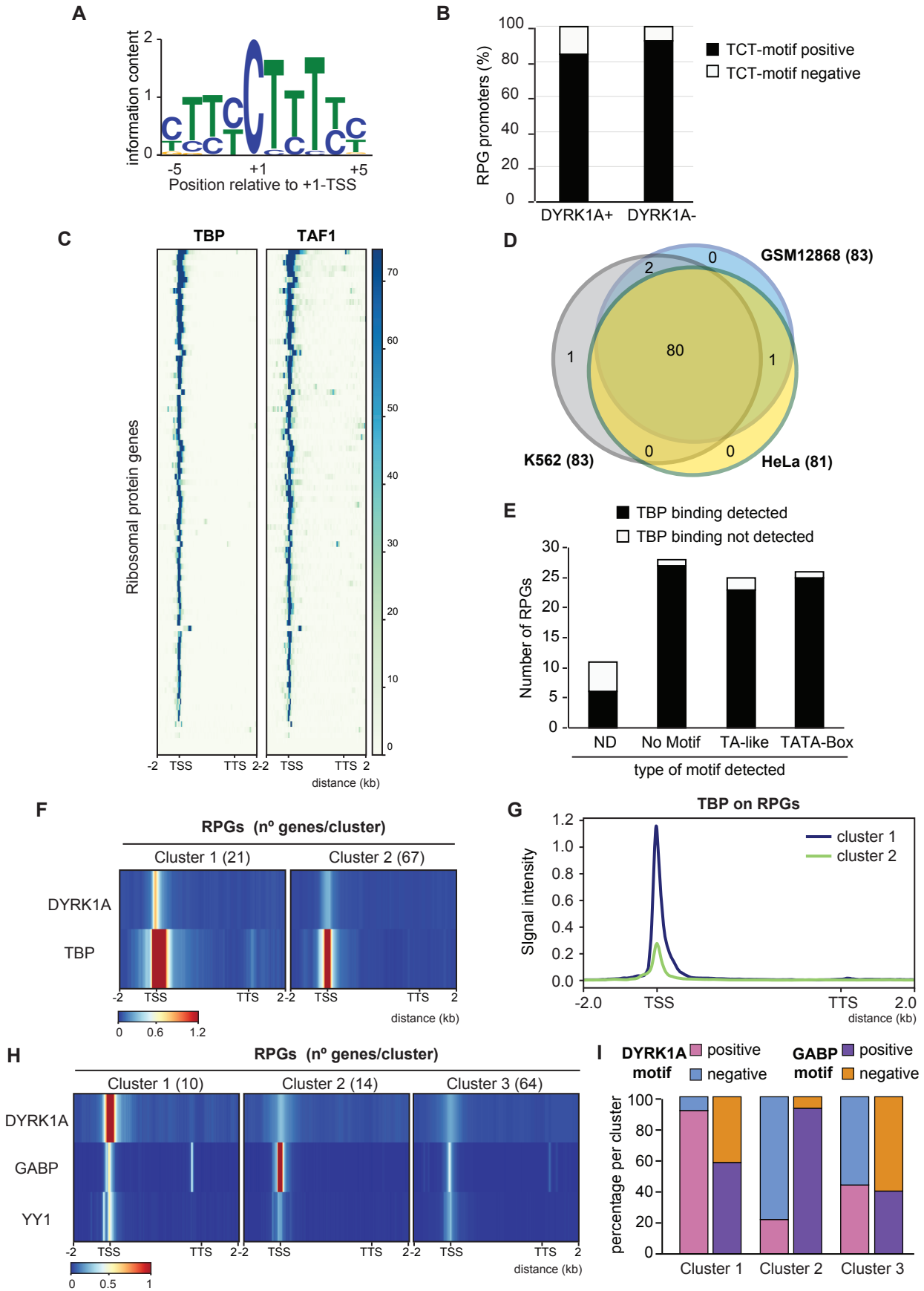

Figure 4
Di Vona, Barba *et al.*

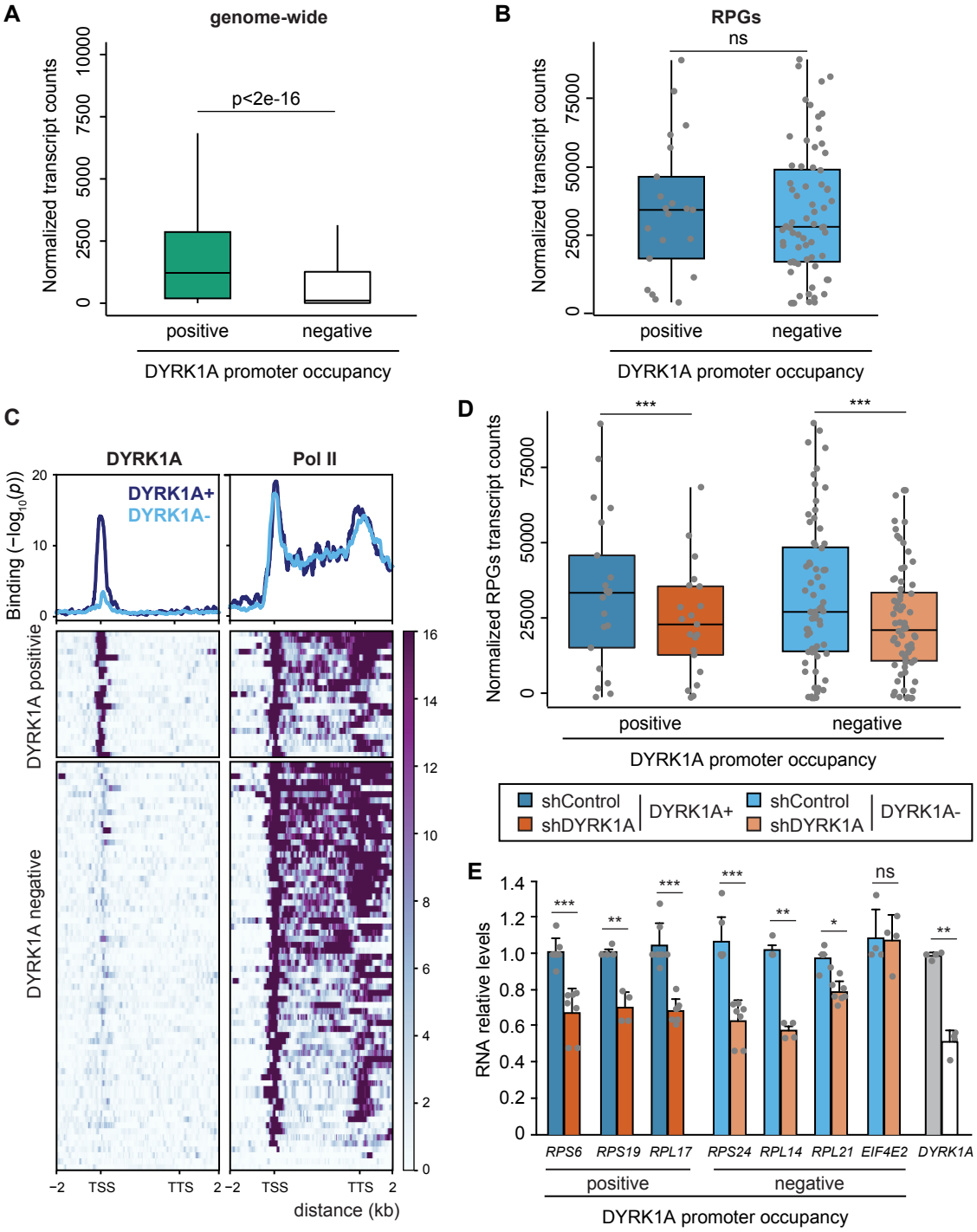

Figure 5
Di Vona, Barba *et al.*

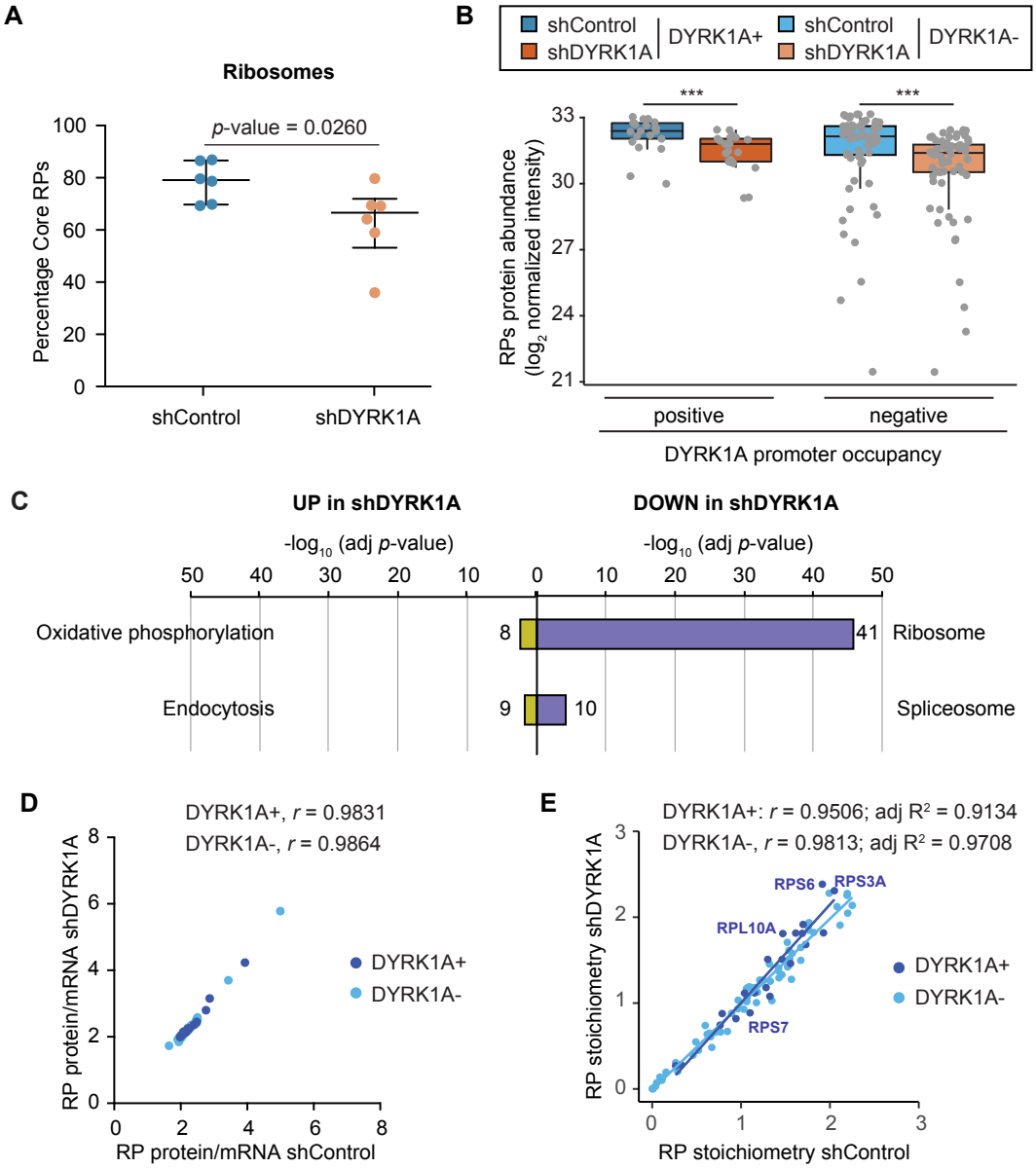

Figure 6
Di Vona, Barba *et al.*