Check for updates

Regulation of translation by site-specific ribosomal RNA methylation

Martin D. Jansson [●1✉], Sophia J. Häfner [●1], Kübra Altinel[1,5], Disa Tehler[1,5], Nicolai Krogh[2], Emil Jakobsen [●3], Jens V. Andersen [●3], Kasper L. Andersen[1], Erwin M. Schoof [●4], Patrice Ménard[1], Henrik Nielsen [●2] and Anders H. Lund [●1✉]

Ribosomes are complex ribozymes that interpret genetic information by translating messenger RNA (mRNA) into proteins. Natural variation in ribosome composition has been documented in several organisms and can arise from several different sources. A key question is whether specific control over ribosome heterogeneity represents a mechanism by which translation can be regulated. We used RiboMeth-seq to demonstrate that differential 2′-O-methylation of ribosomal RNA (rRNA) represents a considerable source of ribosome heterogeneity in human cells, and that modification levels at distinct sites can change dynamically in response to upstream signaling pathways, such as _MYC_ oncogene expression. Ablation of one prominent methylation resulted in altered translation of select mRNAs and corresponding changes in cellular phenotypes. Thus, differential rRNA 2′-O-methylation can give rise to ribosomes with specialized function. This suggests a broader mechanism where the specific regulation of rRNA modification patterns fine tunes translation.

Given the complexity of the ribosome, heterogeneity in its composition can emanate from many potential sources. These include variations in the nucleotide sequences of rRNAs expressed from distinct loci, post-transcriptional rRNA modifications, as well as the complement of associated ribosomal proteins and their post-translational modifications[1–5]. Whether this heterogeneity gives rise to ribosomes with specialized function is currently a matter of debate[6–8].

During ribosome biogenesis, ribose methylation at the 2′-O position occurs only at certain defined rRNA nucleotide positions, which generally cluster in highly conserved and functionally important parts of the ribosome[9]. In Eukarya, almost all 2′-O-methylations (2′-O-me) are added by the methyltransferase fibrillarin, which is guided to specific sites by different C/D box small nucleolar RNAs (snoRNAs) (SNORDs) via base-pairing interactions with their target rRNA[10]. Studies in bacteria and yeast indicate 2′-O-me can influence ribosome biogenesis, stabilization of rRNA structures and translational fidelity[11–13]. The 2′-O-me is the most abundant post-transcriptional modification in human rRNA; however, little is known about the impact of differential methylation on ribosome function and translation[3].

Results

Variation and dynamics in rRNA 2′-O-methylation. We previously used RiboMeth-seq (RMS)[14] to map and quantify rRNA 2′-O-me in two human cancer cell lines, HeLaS3 and HCT116 (ref. [3]). To obtain a broader perspective, we profiled additional cell lines (Extended Data Fig. 1). The result was that a new 2′-O-me site at 28S:U2031 was detected in nontumor-derived fibroblasts, probably guided by SNORD123, and a previous candidate at 28S:G3606 was also verified (Extended Data Figs. 1a and 2a–c). In human rRNA, 113 2′-O-me sites (42 in 18S, 69 in 28S and 2 in 5.8S) have now been identified in analyses of various cell lines by RMS or mass

spectrometry (MS)[4,15]. Here we focus on the 39 18S and 68 28S sites, for which methylation was detected in this study, and for which both RMS and MS evidence exists.

Comparing the rRNA 2′-O-me signatures across multiple cell lines, it became apparent that substantial variation exists between levels at many sites (Extended Data Fig. 1). Interestingly, heterogeneity in rRNA 2′-O-me has also recently been observed in samples from patients with diffuse large B cell lymphoma and breast cancer[16,17]. This raised the question of whether these differences are hard-wired and cell type specific, or whether they could arise due to the action of distinct regulatory pathways within the cell. As some of the largest methylation changes at particular sites were apparent between tumor- and nontumor-derived cell lines, such as HeLaS3 adenocarcinoma and BJ[hTERT] fibroblasts, for example (Extended Data Fig. 2d–g), we aimed to explore the potential impact of a key cancer-relevant pathway on rRNA 2′-O-me levels.

The _MYC_ proto-oncogene is a master transcription factor, promoting cell proliferation and also translation[18]. As a model for oncogene activation, we engineered BJ[hTERT] fibroblasts to express _MYC_ in an inducible manner (BJ[MYC]). Following treatment with doxycycline for 72 or 120 hours, increased MYC mRNA and protein levels are observed (Extended Data Fig. 2h,i). As expected, oncogene expression in these nontumor cells resulted in slightly elevated p53 protein levels, as well as mildly increased expression of cell cycle inhibitors p21 and p16 (Extended Data Fig. 2i,j). However, no cell cycle arrest was detected under these conditions (Extended Data Fig. 2k). MYC expression resulted in upregulation of nucleophosmin and fibrillarin, both known target genes with established roles in ribosome biogenesis[19] (Extended Data Fig. 2l). The 2′-O-me profiles of 18S and 28S rRNA in uninduced BJ[MYC] cells were compared with those following 72 and 120 hours of MYC induction (Fig. 1a,b and Extended Data Fig. 1b–d). Particularly robust changes were detected at three sites, 2′-O-me increased at 18S:C174, while levels decreased at the

[1]Biotech Research and Innovation Centre, University of Copenhagen, Copenhagen, Denmark. [2]Department of Cellular and Molecular Medicine, University of Copenhagen, Copenhagen, Denmark. [3]Department of Drug Design and Pharmacology, University of Copenhagen, Copenhagen, Denmark. [4]Department of Biotechnology and Biomedicine, Technical University of Denmark, Lyngby, Denmark. [5]These authors contributed equally: Kübra Altinel, Disa Tehler. ✉e-mail: martin.jansson@bric.ku.dk; anders.lund@bric.ku.dk
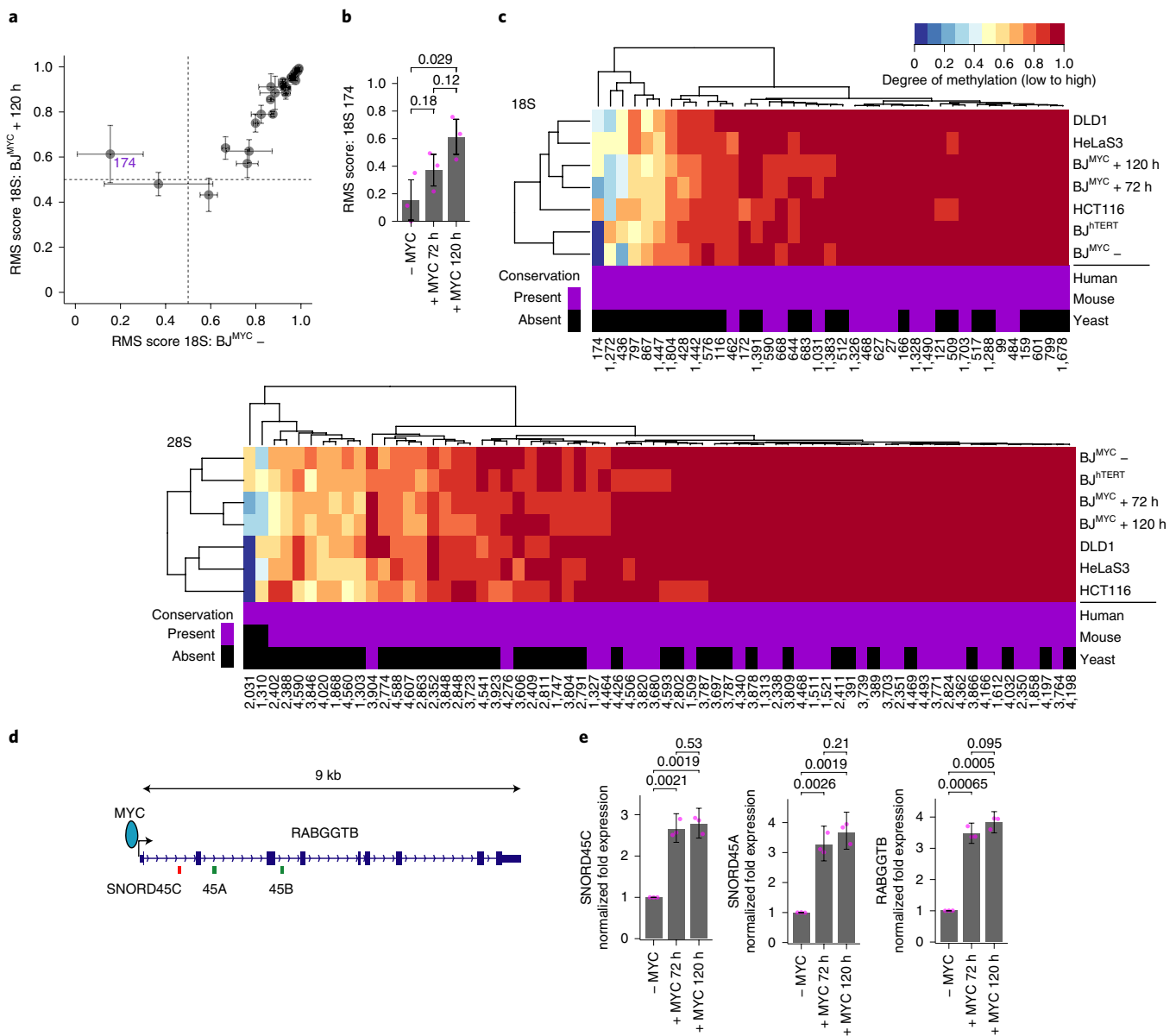
Fig. 1 | MYC expression induces dynamic changes in 2′-O-methylation of human ribosomal rRNA. a, Comparison of RMS scores representing the fraction of 18S rRNA molecules 2′-O-methylated at each site in BJ^{MYC} cells, where MYC expression is induced for 120 h or not induced (−). Points denote mean RMS score (*n* = 3 libraries from individual cultures) per site. Horizontal and vertical error bars indicate ±s.d. for each condition. Sites exhibiting significant change are labeled with nucleotide position (magenta). **b**, RMS scores for 2′-O-me at 18S:C174 in BJ^{MYC} cells, without MYC induction or with MYC induction for 72 or 120 h. Columns indicate mean RMS score for each condition of *n* = 3 libraries from individual cultures, points (magenta) denote each value separately. Error bars represent ±s.d. *P* values (Welch's *t*-test, two-tailed) are shown for each pairwise comparison, indicated by brackets. **c**, Heat map showing unsupervised clustering of RMS scores at 2′-O-me sites (columns) in 18S (top panel) and 28S (bottom panel) rRNAs from the multiple cell lines indicated (rows). Color scale represents degree of methylation from 0 (low, blue) to 1 (high, red). Dendrograms define identified clusters. Directly below each RMS heat map, conservation of human 2′-O-me sites is indicated for equivalent nucleotide positions known to be methylated in mouse and yeast (purple, present; black, absent). **d**, Schematic showing location of SNORD45 family members in the snoRNA host gene *RABGGTB*. MYC transcription factor is depicted by the cyan oval. **e**, Expression of RABGGTB and SNORD45 family members (C and A) in BJ^{MYC} cells where MYC expression is not induced or induced for 72 or 120 h. Columns indicate mean (*n* = 3 independent experiments) fold change relative to the control condition, normalized to snRNA U6 expression, as determined by RT–qPCR analysis. Error bars indicate 95% confidence intervals. *P* values (Welch's *t*-test, two-tailed) are shown for pairwise comparisons, indicated by brackets.

newly identified site 28S:U2031 and at 28S:U2402 (Fig. 1a,b and Extended Data Fig. 2m,n).

Using unsupervised hierarchal clustering, we examined the RMS scores of 18S and 28S rRNA across the datasets (Fig. 1c). Around two-thirds of all sites are close to being fully methylated and are generally invariant, while the remainder are fractionally methylated and exhibit a higher degree of variation between cell types and conditions.

The most variable methylated sites tend to be poorly conserved with those occurring at equivalent positions in the yeast ribosome (Fig. 1c and Supplementary Table 1). Furthermore, the 2′-O-me profiles of the fibroblasts after MYC induction cluster alongside the cancer cell lines for both the 18S and 28S rRNAs (Fig. 1c). Methylation levels at 18S:C174 and 28S:U2031 are among the principal determinants of variation in the 18S and 28S rRNA clusters, respectively.

From in silico predictions, SNORD45C was proposed to guide methylation at 18S:C174 (Cm174) and also at 18S:A159, using two different antisense elements[20] (Extended Data Fig. 2o). SNORD45A, however, may also guide methylation at 18S:A159, as well as at 18S:U172. The SNORD45 family members are processed from introns of their host gene *RABGGTB*, which interestingly has been implicated as a MYC transcriptional target[21] (Fig. 1d). Notably, expression of both SNORD45C and SNORD45A is elevated along with the host gene, following MYC induction in fibroblasts (Fig. 1e). So, in this case there appears to be a direct link between induction of SNORD45C by MYC and increased Cm174 levels. We therefore selected Cm174 as the primary candidate for a proof-of-principle study. These results support the idea that discrete cellular pathways can influence rRNA methylation signatures.

Differential 2′-O-methylation of rRNA influences translation.
We selected the cervical carcinoma-derived cell line HeLaS3 to perform characterization of a role for Cm174 in ribosome function, since we required a system where the manifold pleiotropic effects of MYC expression on ribosome biogenesis, translation or transcription[18] would not confound analysis of any changes occurring following alteration of Cm174-modified ribosome levels in the cell.

Typical of many cancer cells, MYC expression is elevated in HeLaS3. The proportion of ribosomes modified at 18S:C174 is similar to that of BJ^MYC at 120 hours post MYC induction (Extended Data Fig. 3a–d). Akin to the situation in fibroblasts, MYC expression also influences SNORD45C, SNORD45A and *RABGGTB* in HeLaS3 cells, since short interfering RNA (siRNA)-mediated knockdown of *MYC* reduced their expression (Extended Data Fig. 3e).

To generate a system in which we could investigate the function of the MYC-responsive rRNA 2′-O-me at 18S:C174, we abolished it by using CRISPR–Cas9 gene editing to delete SNORD45C in HeLaS3 (HeLaS3^45CKO). Expression of SNORD45C was lost in the knockout (KO) cells, with modest reductions in *RABGGTB* and SNORD45A (Extended Data Fig. 3e). RMS analysis showed that Cm174 was completely absent in HeLaS3^45CKO cells, while methylation at 18S:A159, 18S:U172 and other sites remained largely unaffected (Fig. 2a,b and Extended Data Fig. 3f–h). This confirms SNORD45C as the guide for Cm174 and implies that Am159 is predominantly guided by SNORD45A.

Loss of Cm174 did not affect global translation, as determined by both [35S]methionine incorporation and by examining ribosome distribution in polysome gradients (Fig. 2c–e). Furthermore, the levels of rRNA processing intermediates were also unaltered, indicating no defects in rRNA biogenesis (Extended Data Fig. 3i).

To ascertain whether Cm174 status might influence the translation of specific mRNAs, we performed ribosome profiling of HeLaS3

and HeLaS3^45CKO cells (Extended Data Fig. 3j–q). Transcripts were assigned to one of five categories depending on their differential regulation in HeLaS3^45CKO cells: those that displayed significant change in mRNA expression (transcription); ribosome occupancy (translation); both transcription and translation together; opposite change in transcription and translation; or were unchanged. While the majority of transcripts were unaffected (8,388), abrogation of Cm174 resulted in a large number of transcripts changing at the level of translation alone (2,071), while few changes in transcription (106) were detected (Fig. 2f and Supplementary Table 2). A total of 1,168 transcripts displayed decreased translation (TL-DN) and 903 showed increased translation (TL-UP) in HeLaS3^45CKO (Fig. 2f). This predominantly translational regulation implicates modulation of ribosome function by Cm174 as a probable mechanism.

mRNAs in the TL-DN and TL-UP sets are associated with different gene ontology (GO) categories for biological processes (Fig. 2g,h). For example, cell cycle- and mitosis-related terms are significantly enriched in the TL-DN set. On the other hand, metabolism-related, oxidation-reduction and intracellular transport terms are over-represented in the TL-UP set. In contrast to the translationally regulated sets, no enriched GO categories were found associated with transcripts displaying increased mRNA expression (TC-UP), and only four significant categories were found for those with decreased mRNA expression (TC-DN) (Extended Data Fig. 3r).

Metagene read coverage analysis indicates that ribosome occupancy is globally reduced in TL-DN and increased in the TL-UP sets, respectively, in accordance with altered ribosome density in the translationally regulated transcripts (Fig. 2i). No major changes in the distribution of ribosomes over the coding sequence (CDS) regions are apparent, although we noticed a slight reduction in ribosome density over the CDS in the not significant (NS) set for HeLaS3^45CKO, along with a tendency for increased ribosome density around the translation start site, which is more pronounced for TL-UP transcripts (Fig. 2i). These observations are compatible with differences in either translation initiation or elongation rates.

Taken together, these findings indicate that the proportion of ribosomes methylated at 18S:C174 affects how efficiently different subsets of mRNAs are translated.

Features of translationally altered mRNAs. We began searching for features common among the mRNA transcripts comprising each respective regulated set that might explain their altered translation. Several studies have reported that other potential sources of ribosome specialization may influence the preference for 5′ cap- or Internal ribosome entry site (IRES)-dependent modes of translational initiation[22,23]. To assess whether Cm174 status could affect

Fig. 2 | Ribosome 2′-O-methylation at 18S:C174 regulates translation of specific mRNAs. a, Comparison of RMS scores in 18S rRNA from HeLaS3 and HeLaS3^45CKO. Points denote mean RMS score (*n* = 3 libraries from individual cultures) per site. Horizontal and vertical error bars indicate ±s.d. Sites exhibiting significant change are labeled with nucleotide position (magenta). **b**, RMS scores for 2′-O-me at 18S:C174 in HeLaS3 and HeLaS3^45CKO. Columns indicate mean RMS score for each condition of *n* = 3 libraries from individual cultures. Error bars represent ±s.d. *P* values (Welch's *t*-test, two-tailed) are shown. **c**, Polysome gradient profiles in HeLaS3 (WT, black) and HeLaS3^45CKO (45CKO, red) cells. Representative profiles (*n* = 3 independent experiments) shown. **d**, Global protein synthesis rates in HeLaS3 and HeLaS3^45CKO. Representative (*n* = 3 independent experiments) autoradiograph of PAGE gel with western blot for protein loading shown. **e**, Quantification of [35S]methionine/cysteine incorporation from *n* = 3 independent experiments, as in **d**. Columns indicate mean relative to control (HeLaS3). Error bars represent ± s.d. and *P* value (Welch's *t*-test, two-tailed) shown above brackets. **f**, Changes in mRNA expression and ribosome footprint levels from ribosome profiling data, comparing HeLaS3 and HeLaS3^45CKO cells (*n* = 3 libraries from individual cultures). Histogram (left) giving the number of mRNA transcripts significantly regulated in each category (Benjamini–Hochberg P_{adj} < 0.05). Scatter plots (right) with mRNA transcripts colored according to indicated regulation type. Number of sequenced transcripts analyzed given by *n*. Numbers of transcripts up- or downregulated in 45CKO relative to WT cells in the translation alone set also shown (TL-UP, TL-DN, respectively). **g**, Gene ontology analysis of mRNA transcripts in TL-DN set. Top 10 biological process GO categories (FDR < 0.05) are labeled. Number of genes overlapping with each GO category indicated by the color-scale gradient (count). **h**, As in **g** for TL-UP set. **i**, Metagene analysis of ribosome footprint coverage (RPKM normalized median of libraries *n* = 3) across mRNAs in each category for HeLaS3 (WT, black) and HeLaS3^45CKO (KO, red). NS, TL-DN, TL-UP represent transcripts showing not significantly changed, downregulated or upregulated translation, respectively. Coding regions (between START and STOP) for each transcript scaled to 100 nucleotides. 5′ and 3′ untranslated regions (±30 nucleotides) unscaled.

IRES-mediated translation, we performed gene-set enrichment analysis (GSEA) on the translationally regulated mRNAs using a list of high-confidence IRES-regulated genes, and found no signifi-

cant enrichments (Extended Data Fig. 4a). Additionally, in reporter assay experiments, no differences in the activities of two cellular IRES elements were observed between HeLaS3[45CKO] and parental
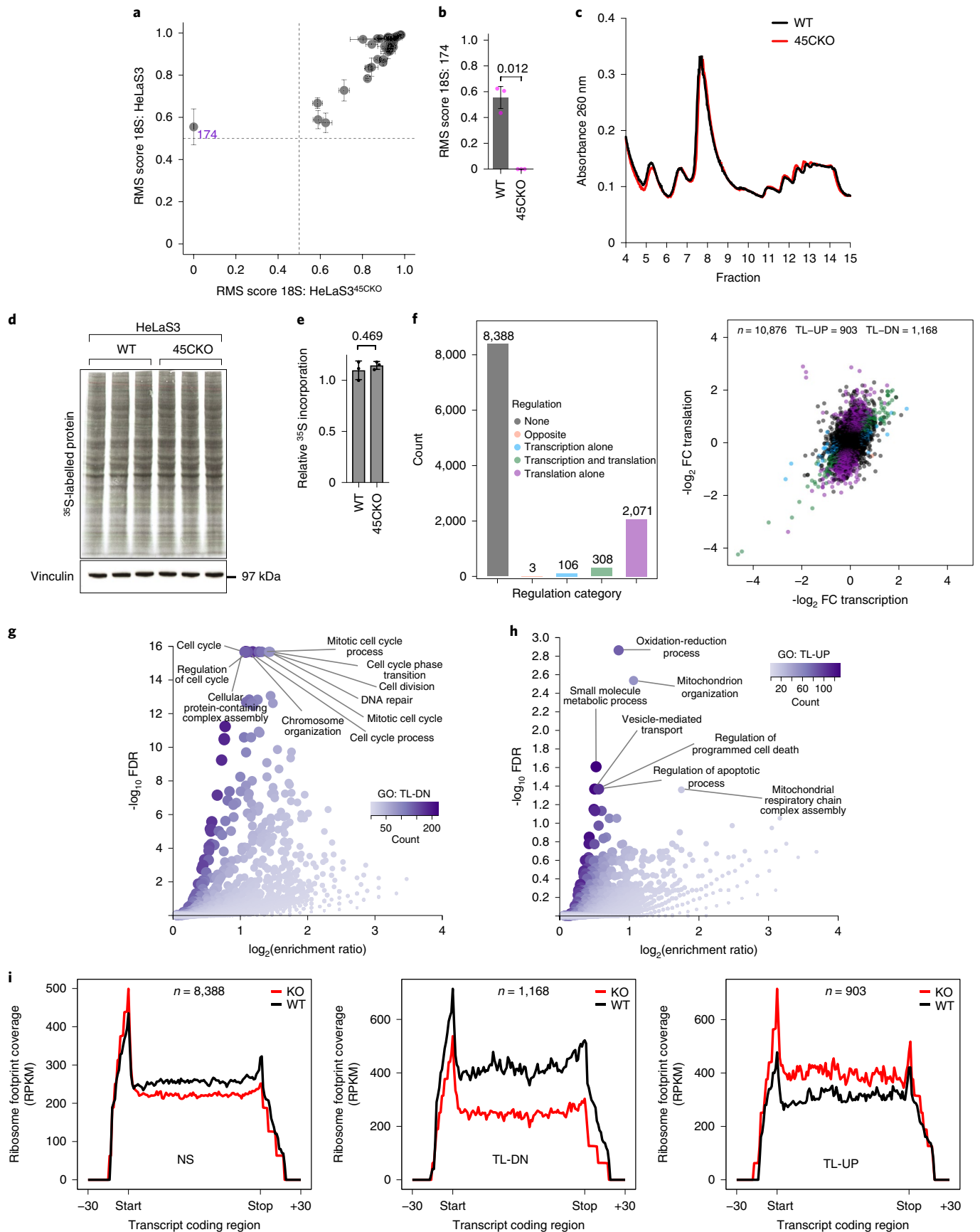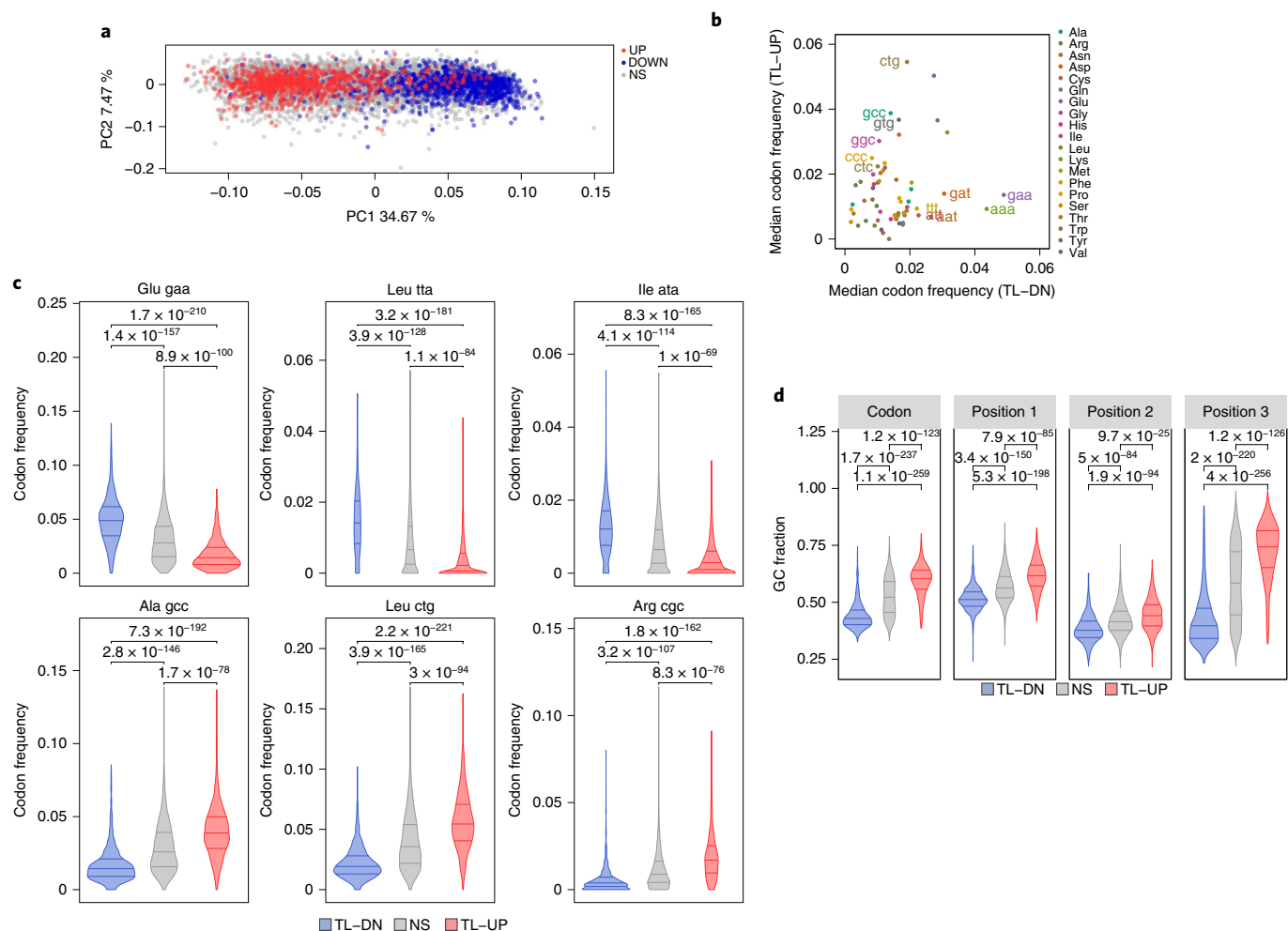
Fig. 3 | Translationally regulated mRNAs display distinct codon usage patterns. a, PCA of codon usage frequency in mRNA transcripts exhibiting increased (TL-UP, red), decreased (TL-DN, blue) and not significantly altered (NS, gray) translation in HeLaS3[45CKO] compared to HeLaS3 cells. **b,** Codon usage frequencies in mRNA transcripts exhibiting increased (TL-UP) or decreased (TL-DN) translation in HeLaS3[45CKO] cells. Median frequency of each codon is plotted in each case. Codons are color-coded by amino acid as indicated. Codons changing with $\log_2$(fold change) ($\log_2$ FC) >1, minimum frequency of 0.02 and $P_{adj} < 0.001$ (Bonferroni-corrected Wilcoxon tests, unpaired, two-sided) are annotated by codon sequence. **c,** Examples of individual codons displaying altered usage frequencies in each translation regulation category. Glutamic acid (gaa), alanine (gcc), leucine (tta), leucine (ctg), isoleucine (ata) and arginine (cgc) are shown. P_{adj} values (Bonferroni-corrected Wilcoxon tests, unpaired, two-sided) for comparisons are indicated by brackets. **d,** Fraction of G or C nucleotides present in codons comprising mRNA transcripts in each translation regulation category: downregulated (TL-DN, blue), not significantly regulated (NS, gray) or upregulated (TL-UP, red). Comparisons for entire codons (codon) or first, second and third codon positions (positions 1–3) are shown separately. P_{adj} values (Bonferroni-corrected Wilcoxon tests, unpaired, two-sided) for comparisons are shown above brackets.

cells (Extended Data Fig. 4b). Together, these results suggest that Cm174 has no major influence on IRES-dependent translation.

Next, we evaluated the codon compositions of mRNAs in the TL-DN, TL-UP and NS sets. Principal component analysis (PCA) indicated a strong variance in the use of elongation-related codons present in the CDSs of TL-DN and TL-UP transcripts (Fig. 3a). Comparison of the individual codon frequencies in these two sets revealed a striking two-pronged distribution pattern (Fig. 3b). In all, we identified 18 codons in TL-DN and 17 in TL-UP that were significantly enriched with $\log_2$(fold frequency changes) >1 (Fig. 3b,c and Supplementary Table 3). Codons enriched in TL-DN tend to be AU-rich, whereas those in TL-UP are more GC-rich, and this is especially apparent at the third nucleotide position (AT3 or GC3) (Fig. 3d).

These patterns apply only to translationally regulated transcripts, as no such differences in codon usage are apparent between the differentially expressed mRNA sets (Extended Data Fig. 4c–e).

Despite large energetic differences between GC- and AU-rich codon–anticodon duplexes, binding and accommodation rates of

aminoacyl-transfer RNAs (tRNAs) and the complex stabilities at the ribosomal A-site appear to be remarkably similar in vitro[24]. This is due to evolution of intricate mechanisms in both ribosomes and tRNAs that minimize differential thermodynamic contributions of codon–anticodon interactions to maintain translation fidelity. The complex ribosome structural rearrangements associated with conformational changes and activation of eukaryotic translation elongation factor 1 alpha 1 (EEF1A1, or EF-Tu in prokaryotes) during translation elongation constitutes one such example.

A revised energy-based interpretation of the genetic code was previously described[25]. Interestingly, codons with increased frequencies in TL-DN transcripts show a robust tendency to have thermodynamically weaker interactions with their cognate tRNAs, while those with increased frequencies in TL-UP tend to be stronger (Supplementary Table 3). Although the third codon position carries less weight in this energy-based theoretical model, recent work has linked GC3 and AT3 codons with optimal and suboptimal translation, as well as mRNA stability, respectively[26,27]. Here,
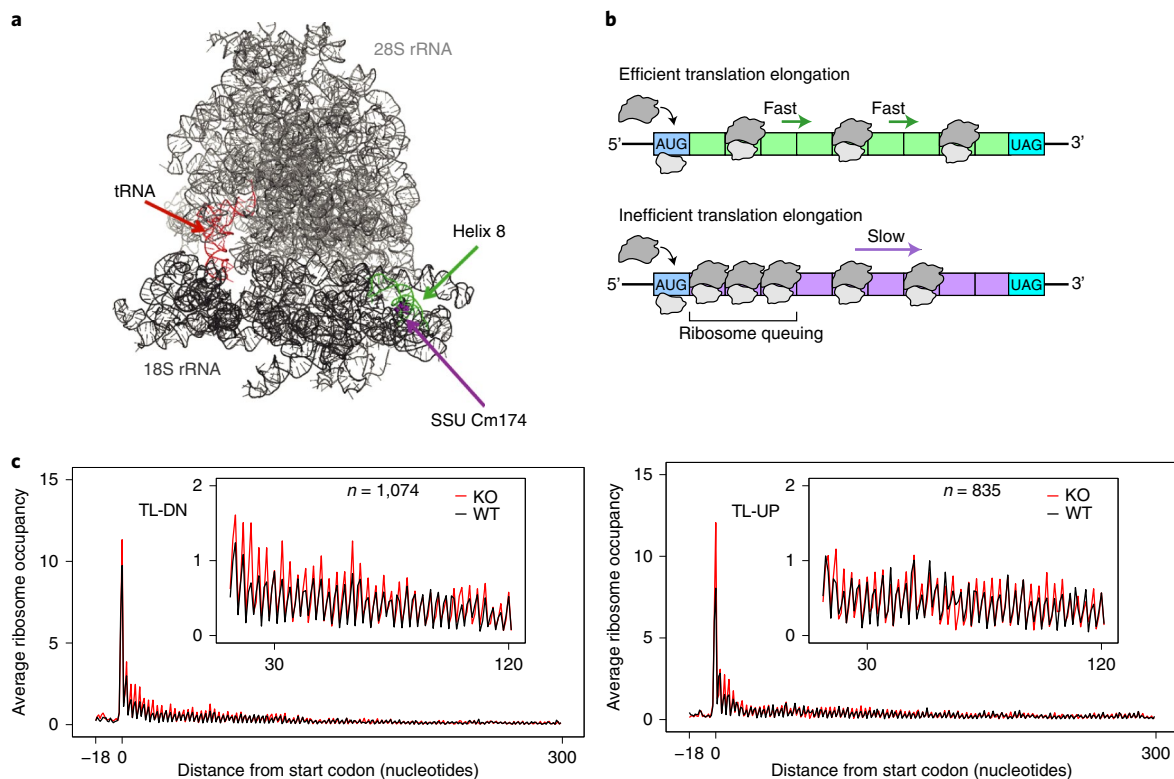
Fig. 4 | rRNA methylation at 18S:174 influences ribosome elongation rates on translationally regulated mRNAs. a, Structure of human ribosome indicating location of C174 (purple) at base of h8 (green) in human rRNA. PyMol graphic based on cryo-EM ribosome structure modeled by Khatter et al. 18S (dark gray) and 28S (light gray) rRNAs are indicated. E-site tRNA (red) is also shown for orientation. **b**, Schematic illustration of ribosome distribution in mRNAs where translation elongation is efficient or inefficient. Ribosomal subunits (dark gray and light gray) assemble at the start codon (AUG). Ribosomes are able to rapidly translate through transcripts where codons are optimal (top panel, green) or well translated by a particular specialized ribosome type. The same ribosome type may exhibit slower translation elongation rates over transcripts with nonoptimal codon composition (lower panel, purple). This could result in ribosome queuing back towards the start codon. **c**, Average ribosome occupancy in HeLaS3 (WT, black) and HeLaS3[45CKO] (KO, red) cells. Metagene analysis of mRNA transcripts displaying downregulated (TL-DN, right panel) or upregulated (TL-UP, left panel) translation in HeLaS3[45CKO] cells are shown. *n* denotes number of transcripts analyzed following removal of extreme outliers.

however, we do not observe significant changes in mRNA levels associated with altered stability between the differentially translationally regulated transcripts carrying GC3/AT3 codons. In addition to the general GC/AT bias, all codons we find enriched in the TL-UP set are GC3 codons, while all in TL-DN are AT3 (Fig. 3b and Supplementary Table 3).

Moreover, there is considerable overlap between the TL-DN- and TL-UP-enriched codons with corresponding anticodons previously shown to be utilized in translation of proliferation- or differentiation-associated genes, which appear to segregate according to GC/AT and also GC3/AT3 content[27,28] (Extended Data Fig. 4f).

This suggests that certain factors contributing to ribosome heterogeneity, in this case methylation at 18S:C174, can mediate differential translation of such codons by modulating the capacity of the ribosome to compensate for differences in codon–anticodon interactions, for instance.

Regulation of translation by rRNA 18S:C174 methylation. The observations above are particularly interesting when considering the location of Cm174 within the ribosome tertiary structure (Fig. 4a). 18S:C174 is situated at the base of an internal loop of helix 8 (h8), a highly flexible region of the small ribosomal subunit that forms part of the intersubunit bridge B8 (B8), together with helix 14 (h14)[29]. While this area is poorly resolved in human ribosome structures that capture dynamics of the translation elongation cycle[30], h8 is evolutionarily well conserved (Extended Data Fig. 4g), and in

bacteria is known to participate in substantial conformational changes during translation elongation[31,32]. Furthermore, in *Escherichia coli*, h8 is a hotspot for ribosomal ambiguity mutations (*ram*) that alter the helix structure and disrupt B8 (refs. [33,34]). *ram* mutations in h8 result in translation errors characterized by a decreased stringency of decoding and an increased rate of GTP hydrolysis by EF-Tu, which plays a central role in translation elongation[33–35]. It is therefore plausible that methylation status at C174 could affect ribosome function, albeit in a less drastic way than *ram* mutations, potentially by regulating the conformational flexibility of h8. We speculate that this in turn may modulate h8:h14 interaction, the dissociation of which is known to induce far-reaching ribosomal structural rearrangements that culminate in domain closure, EEF1A1/ EF-Tu activation and subsequent decoding[33].

It has also been proposed that lower translational efficiency, due to the presence of nonoptimal codons, can result in a back-queuing of ribosomes towards the translation start site[36]. We imagined that similar situations could arise due to regulation of ribosome function, resulting from differential 2′-O-methylation in this case, where altered translation of mRNAs with distinct codon compositions is a factor (Fig. 4b). With this in mind, we examined the average ribosome occupancy, which is considered a measure of ribosome dwell time or pausing[37], in both HeLaS3[45CKO] and parental cells. Transcripts comprising the TL-UP and TL-DN sets were compared (Fig. 4c). Curiously, the average ribosome occupancy rates over the first five codons appear higher in HeLaS3[45CKO] for both

sets. Over codons 6 to 40, TL-DN transcripts display higher average occupancy in HeLaS3[45CKO], whereas this difference is not observed for the TL-UP set. This apparent queuing of ribosomes on TL-DN transcripts in HeLaS3[45CKO] is consistent with slower decoding at AU-rich codons by ribosomes lacking Cm174, and that reduction in elongation rates can result in accumulation of ribosomes towards the translation start site.

Cm174 affects cellular phenotypes. Consistent with decreased translation of mRNAs linked to cell division processes, the proliferation rate and percentage of cells in S phase is reduced in HeLaS3[45CKO] (Extended Data Fig. 5a–c). To exclude the possibility that clonal artifacts might contribute to this growth phenotype, we reinstated SNORD45C expression in the knockout background (HeLaS3-GFP[45CKO+45C]). This completely rescued the decrease in cell proliferation, and, crucially, Cm174 levels were also restored (Fig. 5a–c and Extended Data Fig. 5d–f). Further analysis suggests the growth defect is due to SNORD45C and concomitant Cm174 loss, rather than any reduction in *RABGGTB* or SNORD45A expression, since these were not restored to wild-type (WT) levels in HeLaS3-GFP[45CKO+45C] cells (Extended Data Fig. 5g).

To assess if the increased translation of metabolic-related genes gave rise to any phenotypes, we analyzed respiration and key parameters of mitochondrial function. Cells lacking Cm174 displayed higher basal and ATP-coupled respiration, suggestive of altered metabolism (Fig. 5d and Extended Data Fig. 5h–j). Reintroduction of SNORD45C fully reversed the altered metabolic phenotype (Fig. 5d).

We then set out to ascertain whether the observed translational regulation resulted in changes at the protein level, which might explain these cellular phenotypes. Quantitative tandem-mass-tag MS was used to analyze the proteomes of HeLaS3 and HeLaS3[45CKO] cells (Fig. 6a, Extended Data Fig. 6a and Supplementary Table 4). Peptides derived from 7,893 distinct proteins were detected in both conditions and used for protein abundance (PA) measurements. A total of 936 proteins were identified with significantly reduced abundances (PA-DN) in HeLaS3[45CKO], and 631 with increased (PA-UP) abundances (Fig. 6a). Supporting the ribosome profiling

data, GO analysis of the PA-DN and PA-UP sets revealed striking similarities in the over-represented terms with those of the corresponding translationally regulated mRNA sets. Proteins in the PA-DN set were enriched for functions related to cell cycle and cell division, while those in PA-UP were associated predominantly with various metabolic processes and cellular transport (Fig. 6b,c).

We then compared the protein abundances measured in both wild-type and HeLaS3[45CKO] cells to ribosome occupancy or mRNA expression for equivalent protein-coding genes. The patterns observed were highly similar for the two cell lines and both correlations were statistically significant, although that between protein and ribosome occupancy was stronger (Fig. 6d,e and Extended Data

Fig. 5 | Ribosome methylation at 18S:174 affects cellular phenotypes.
a, Cell proliferation in HeLaS3[GFP] (WT-GFP, black), HeLaS3[45CKO-GFP] (45CKO-GFP, red) and HeLaS3[45CKO-GFP+45C] (45CKO-GFP + 45C, red, dashed) cells, measured by crystal violet assay at specified times. Points represent mean of independent experiments ($n = 3$). Error bars represent ±s.d. *P* values (Welch's *t*-test, two-tailed) indicated by brackets are shown for day 3. Significant changes in proliferation ($P < 0.05$) between 45CKO-GFP and WT-GFP or 45CKO-GFP + 45C were observed at all time points except at day 1. **b**, Comparison of RMS scores representing fraction of 18S rRNA molecules 2′-O-methylated at each site in HeLaS3[45CKO-GFP] and HeLaS3[45CKO-GFP+45C] cells (left). Points denote mean RMS score ($n = 3$ libraries from individual cultures) per site. Horizontal and vertical error bars indicate ±s.d. for each condition. Sites exhibiting significant change are labeled with nucleotide position (magenta). **c**, RMS scores for 2′-O-me at 18S:C174 in HeLaS3[45CKO-GFP+45C] and HeLaS3[45CKO-GFP] cells. Columns indicate mean RMS score for each condition of $n = 3$ libraries from individual cultures, points (magenta) denote each value separately. Error bars represent ±s.d. *P* value (Welch's *t*-test, two-tailed) is shown above the bracket. **d**, Respiration and key parameters of mitochondrial function in HeLaS3[GFP], HeLaS3[45CKO-GFP] and HeLaS3[45CKO-GFP+45C] cells. Basal respiration, ATP synthesis, proton leak and uncoupled respiration are measured by oxygen consumption rate (OCR) following injections with compounds modulating cellular respiration. Columns indicate mean OCR for each condition ($n = 5$ independent experiments), points (black) denote each value separately. *P* values (repeated measures one-way ANOVA with Tukey correction for multiple comparisons) are shown for each comparison, indicated by brackets. Error bars represent ±s.d.
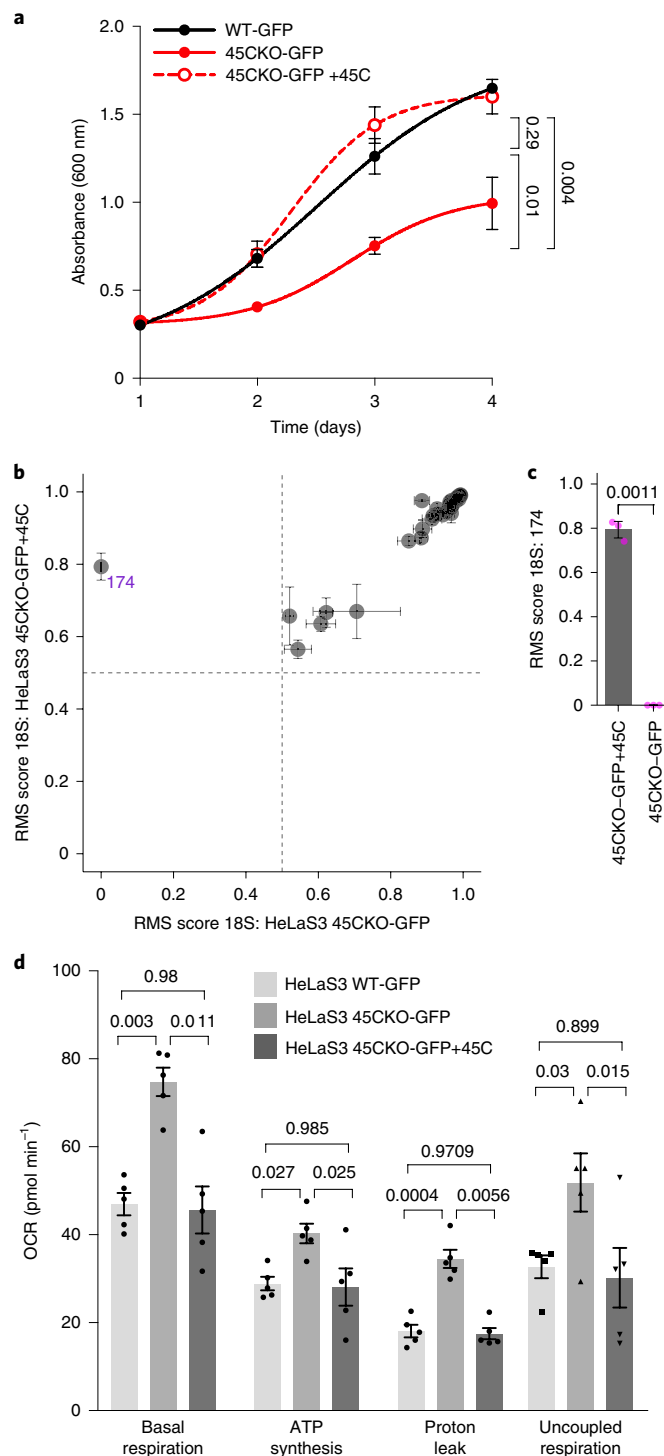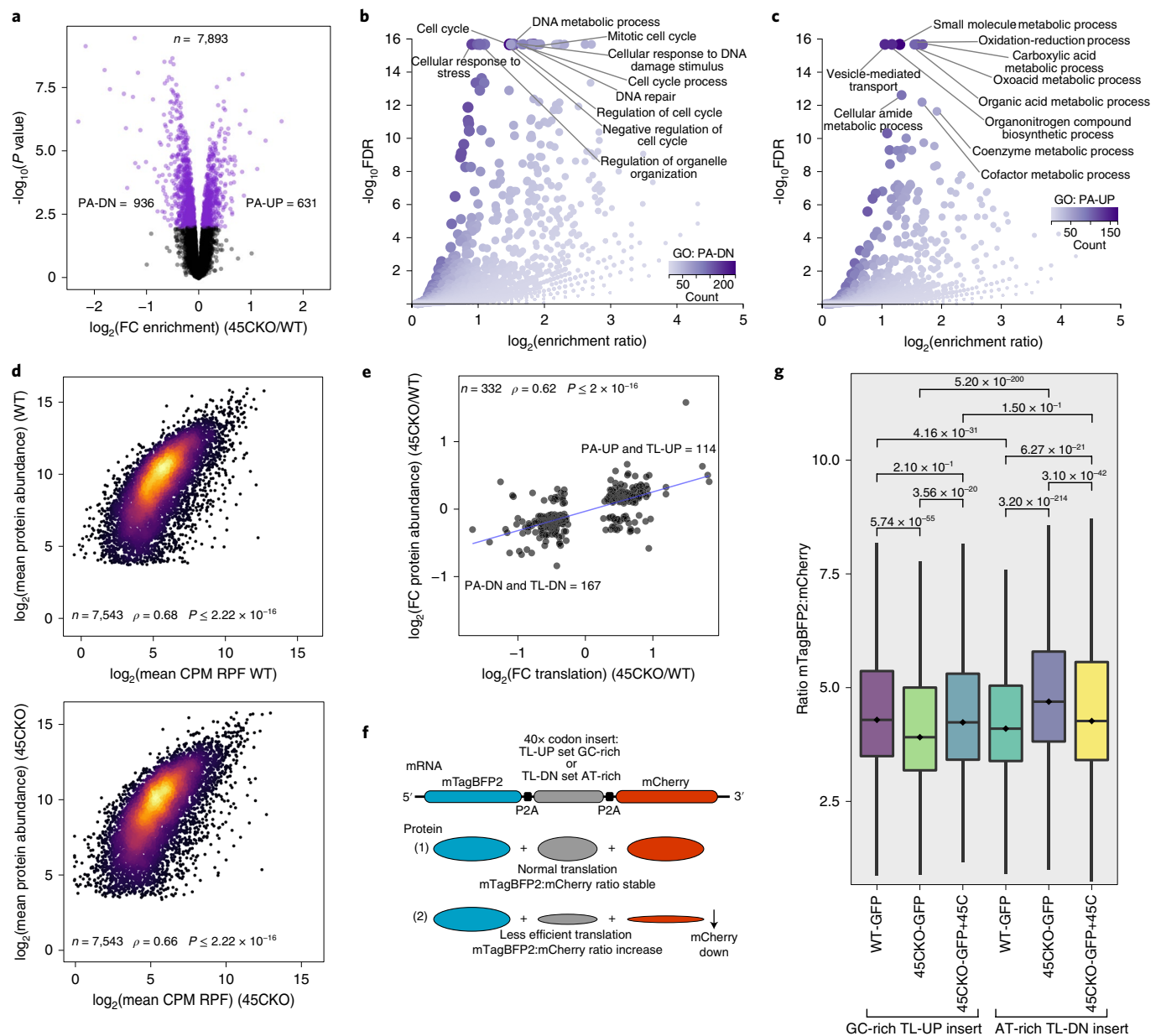
Fig. 6 | Ribosome methylation status at 18S:174 influences the cellular proteome and decoding of AT- and GC-rich codon sequences. a, Protein abundance changes in HeLaS3[45CKO] (45CKO) compared to HeLaS3 (WT). $\log_2$(abundance ratios) (45CKO/WT) and $-\log_{10}$(P values) shown ($n=4$ independent protein samples per condition). Those significantly changing (Benjamini–Hochberg $P_{adj} < 0.05$) are colored (magenta). n denotes number of proteins analyzed. Numbers of proteins enriched is indicated (PA-DN, PA-UP). **b**, GO analysis of proteins in the PA-DN set. Top 10 biological process categories (FDR < 0.05) are labeled. Number of genes overlapping with each category indicated by a color-scale gradient (count). **c**, As in **b**, for the PA-UP set. **d**, Correlation between mRNA translation and protein abundance for HeLaS3 (WT) (top) and HeLaS3[45CKO] (45CKO) (bottom). Points represent mean RPF count (CPM) per transcript ($n=3$ libraries from independent cultures, from ribosome profiling) and corresponding mean protein abundance for the same gene. Color scale indicates plotting density. n denotes number of genes/proteins. Spearman's ρ and associated P value shown. **e**, Correlation between mRNA translation and protein abundance for those genes with significant changes in both ribosome profiling and MS analysis of HeLaS3[45CKO] (45CKO) and HeLaS3 (WT). Points represent $\log_2$ FC translation and $\log_2$ FC protein abundance ratio. n denotes number of genes. The number of genes displaying concomitant $\log_2$ FC changes are indicated by PA-DN and TL-DN (down in both protein abundance and translation) and PA-UP and TL-UP (up in both protein abundance and translation). Spearman's ρ and associated P value shown. **f**, Diagram illustrating codon usage reporter assay. The mRNAs produced are depicted. Shown below are the relative amounts of the proteins produced in two possible situations: normal and less efficient translation through the AT- or GC-rich codon insert. **g**, High-content imaging assay of codon reporter cell lines (replicate 1), containing either AT- or GC-rich codon inserts. mTagBFP2:mCherry ratio after 24 h expression of the reporter cassette. Box plots: centerline is the median; hinges are first and third quartiles; whiskers are 1.5× interquartile range; outlier data points are omitted. Bonferroni P_{adj} values from Wilcoxon tests, two-tailed, are given. Number of double-positive mTagBFP2/mCherry cells assayed, from left to right: $n=10,321, 6,707, 3,518, 16,028, 8,262, 5,659$.

Fig. 6b,c). This affirms that, globally, the ribosome footprint measurements are strongly predictive of protein level, and are a more accurate indicator than mRNA levels alone.

Examining those hits identified by both proteomic and ribosome profiling analyses, we found 332 genes in common that were altered significantly in both datasets. While we were initially surprised not
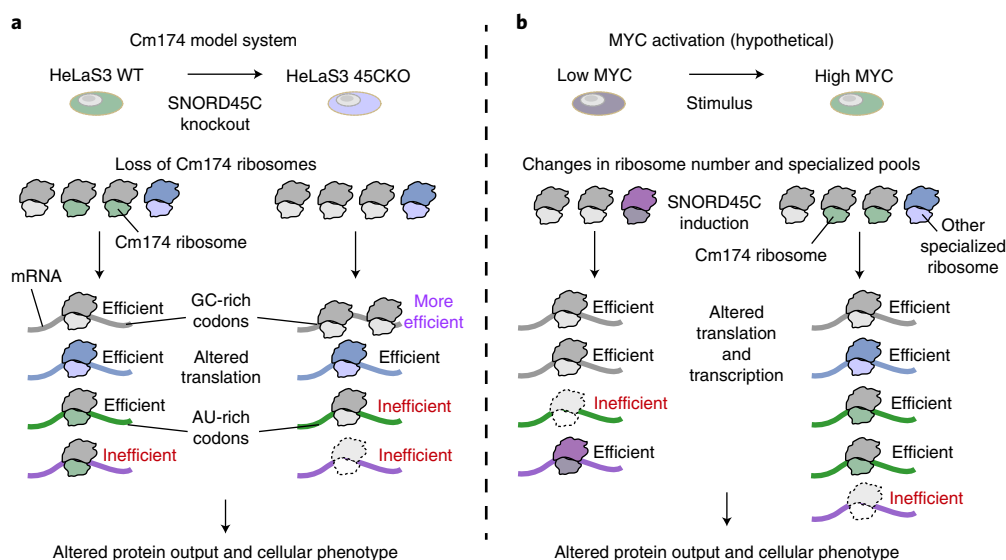
Fig. 7 | Diagram illustrating how translation of specific mRNAs can be altered by specialized ribosomes. a, Model based on Cm174 loss in HeLaS3[45CKO] cells. Deletion of SNORD45C results in loss of ribosomes modified at 18S:C174. Specific mRNAs are translated more or less efficiently as a result of changes in the relative abundance of ribosomes that differ in Cm174 rRNA 2′-O-methylation status. mRNA transcripts enriched in AU-rich codons are translated less efficiently in the absence of Cm174 ribosomes, while those enriched in GC-rich codons are translated more efficiently. This leads to differences in the cellular proteome and an altered phenotype. **b,** Hypothetical model of ribosome regulation following MYC activation in a biological context. Elevated MYC expression leads to increased ribosome biogenesis and induction of distinct specialized ribosome types, including a higher proportion of Cm174-modified ribosomes (due to increased expression of SNORD45C) and potentially differential production of other types. The transcriptome is also altered. Combined, the changes in the cellular ribosome pool along with differential mRNA expression, lead to differences in the proteome and cell phenotype.

to have detected an even larger overlap, even considering the differences between the two methodologies, it is plausible that some translational changes could be buffered by alterations in protein stability. Also, secondary effects may be detected by proteomics. The vast majority of the overlapping genes (281) underwent concomitant changes, resulting in a robust correlation (Fig. 6f). We reasoned that these high-confidence overlap sets of up- (PA-UP and TL-UP) and downregulated (PA-DN and TL-DN) genes, where Cm174-dependent changes in translation were also manifest at the protein level, should be able to explain the phenotypic differences between HeLaS3 and HeLaS3[45CKO] cells. We therefore again performed GO analysis and found that the enriched categories were highly similar to both the previous individual analyses of translational regulation and of protein abundance, which is consistent with reduced cell proliferation and altered metabolism in HeLaS3[45CKO] (Extended Data Fig. 6d,e). Furthermore, we examined the codon usage in the PA-UP and TL-UP and PA-DN and TL-DN sets, and once more detected clear codon usage bias associated with GC content in particular enriched codons (Extended Data Fig. 6e,f). This is highly similar to the analysis performed on TL-UP and TL-DN sets that we identified by ribosome profiling. Interestingly, for those genes where an opposite correlation was observed in the PA-UP and TL-UP and PA-DN and TL-DN comparison, the codon -usage bias was not apparent (Extended Data Fig. 6f).

To experimentally assess whether translation of AT- and GC-rich codons is specifically altered by the presence of Cm174 in ribosomes, we adapted a reporter assay approach previously developed to study ribosome stalling[38]. Instead of sequences promoting stalling, inserts composed of codons that were found to be significantly enriched in transcripts comprising either the TL-UP or TL-DN sets were placed between two fluorescent reporter genes (mTagBFP2, followed by mCherry) under the control of a doxycycline-inducible promoter (Fig. 6f). Once again, we utilized the HeLaS3-GFP[WT], HeLaS3-GFP[45CKO] and HeLaS3-GFP[45CKO+45C]

cell lines, stably integrating either of the two reporter constructs to create a panel of six AT- and GC-rich codon usage reporter cell lines: HeLaS3-GFP[WT-GC-UP]; HeLaS3-GFP[45CKO-GC-UP]; HeLaS3-GFP[45CKO+45C-GC-UP]; HeLaS3-GFP[WT-AT-DN]; HeLaS3-GFP[45CKO-AT-DN] and HeLaS3-GFP[45CKO+45C-AT-DN]. After expression of the reporter cassette was induced by doxycycline treatment for 24 hours, the ratio of production of mTagBFP2 to mCherry in each reporter cell line was measured for double-positive cells, with respect to background levels in untreated cells, by high-content imaging microscopy (Fig. 6g, Extended Data Fig. 6i and Supplementary Table 7). In this way it was possible gauge the influence of Cm174 on the ability of the ribosomes to translate through the AT- and GC-rich codon sequence inserts. The ratio of mTagBFP2:mCherry fluorescence intensity was reduced in HeLaS3-GFP[45CKO-GC-UP] cells compared to HeLaS3-GFP[WT-GC-UP], which is consistent with more efficient translation through the GC-rich codon insert in cells where ribosomes lack Cm174. In HeLaS3-GFP[45CKO+45C-GC-UP], where Cm174 is reinstated, the ratio is restored to a similar level to that in HeLaS3-GFP[WT-GC-UP] cells. Where we would predict less efficient translation through the AT-rich codon insert in HeLaS3-GFP[45CKO-AT-DN], the mTagBFP2:mCherry ratio was increased, indicating that ribosomes without Cm174 translate this less well than in HeLaS3-GFP[WT-AT-DN] cells. Again, this effect was rescued in HeLaS3-GFP[45CKO+45C-AT-DN] cells. In addition, the largest differences were observed between HeLaS3-GFP[45CKO-GC-UP] and HeLaS3-GFP[45CKO-AT-DN] cells, reinforcing the idea that transcripts enriched in AT- and GC-rich codons are differentially translated in the absence of Cm174.

In combination, these results show that absence of ribosomes methylated at 18S:C174 results in considerable phenotypic changes. Furthermore, they imply that these changes arise due to the impact of differential translation on the levels of specific proteins and that the presence of Cm174 on the ribosome can influence how efficiently mRNAs, enriched in AT- or GC-rich codons are translated.

Discussion

Several important conclusions can be drawn from our study, and these may conceptually enhance our understanding of how ribosome function can be regulated. Ribosome 2′-O-me sites can be generally grouped into two distinct categories, highly and fractionally methylated positions. There is little variation in 2′-O-me levels between different cell lines and conditions tested with regard to the highly methylated sites, suggesting that these are important for ribosome biogenesis and optimal function. In contrast, much greater variability is seen in the fractional set, indicating that methylation levels at these sites may be less constrained, making them better suited for regulatory functions.

The existence of fractionally methylated positions demonstrates the presence of distinct pools of differentially 2′-O-methylated ribosomes within specific cell populations, and reinforces the idea that ribosome heterogeneity exists at the level of RNA modifications. In addition, the observation that methylation profiles can differ substantially between cell types suggests a role for 2′-O-me in regulating translation to meet the specific requirements of each cell type. Interestingly, the sites found to exhibit the greatest degree of variation in human cell lines are not conserved in yeast, although they are present in murine ribosomes. This raises the possibility that regulation of ribosome function via differential 2′-O-me has manifested more recently in evolutionary terms. However, there are also 15 2′-O-methylated nucleotides present in yeast rRNA that have no obvious counterparts in human or mouse cells. Overall, the number of rRNA modifications has steadily increased during evolution, together with increased ribosome size and complexity[3]. While previous work in yeast demonstrated that depletion of clusters of ribosomal modifications had various effects on growth, ribosome biogenesis and translation efficiency, it was thought that individual 2′-O-me were of little consequence in terms of their impact on translation[11,39]. Other studies, however, suggest that disruption of single sites could potentially affect ribosome function, for example, knockouts of several individual snoRNAs displayed phenotypes with respect to proliferation and drug sensitivity in yeast[40].

Excitingly, we show that discrete rRNA 2′-O-me sites can actually respond dynamically to distinct upstream signaling pathways. Furthermore, as demonstrated in the case of Cm174, site-dependent rRNA 2′-O-me levels can influence translation of mRNAs associated with functionally distinct categories. Importantly, this resulted in altered cellular phenotypes.

The study of SNORD45C knockout cells, however, represents an artificial situation. Importantly, however, it facilitates dissection of Cm174 function from the pleiotropic effects of MYC expression (Fig. 7a), which results in increased ribosome biogenesis together with regulation of various translation factors, as well as considerable transcriptional changes[18]. In a biological setting, we expect that SNORD45C induction by MYC would occur in concert with an increase in ribosome number and other changes promoting or regulating global translation (Fig. 7b). In this sense, the production of specialized ribosomes can act concomitantly with changes in ribosome abundance, which alone may influence translation of distinct mRNAs[41]. As such, an additional newly synthesized ribosome pool, harboring Cm174, would act to more efficiently translate proliferation-related mRNAs and reinforce a MYC-activated program. Yet, the other pools of ribosomes lacking Cm174 would also still be present for some time, and efficiently translate other mRNA classes involved in maintenance of different cellular processes.

Our findings supplement the accumulating evidence that distinct variations resulting in heterogenous ribosomes exist[6]. These include sequence variability in rRNA alleles and, particularly, expansion segments[5], differential association and modification of ribosomal proteins[2,42], as well as other types of rRNA post-transcriptional modification[4,43]. Despite the abundant evidence of ribosome heterogeneity, so far only a few studies demonstrate that certain transcripts can be differentially translated due to the specific changes in ribosome composition[22,44].

It is intriguing to contemplate a role for rRNA modifications, or indeed other means of ribosome specialization, in regulating translation of specific mRNAs on the basis of their codon compositions. Indeed, there may be a role for Cm174 in fine-tuning ribosome structural rearrangements related to balancing energetic differences, and those intrinsic to GC3/AT3 codons, in tRNA–mRNA interactions during translation elongation. Further in-depth studies would be required to fully substantiate this, however.

The dynamics of the MYC-responsive 2′-O-me sites are not rapid, with changes occurring over time scales in line with turnover of the cellular ribosome pool[45]. Thus, this type of regulation may be more suited to tuning translation in long-term cellular programs, such as during cell differentiation. Also, the possibility exists that in diseases such as cancer, mechanisms generating differentially 2′-O-methylated ribosomes are positively selected, to produce ribosome populations favoring translation of disease-promoting mRNAs. It will be of great interest to determine the exact mechanisms by which different regulatory 2′O-me sites impact specific aspects of ribosome structure and function to elicit effects on translation. Likewise, elucidating the distinct consequences that the diverse signals converging on the ribosome have, and how they differently influence modification of its function, will prove exciting. Overall, our findings promote the idea that the core ribosome itself is a direct translational regulator, and they point to 2′-O-methylation of rRNA being a potentially powerful means to tune ribosome function and impact translation.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41594-021-00669-4.

References

1. Kurylo, C. M. et al. Endogenous rRNA sequence variation can regulate stress response gene expression and phenotype. *Cell Rep.* **25**, 236–248.e6 (2018).
2. Emmott, E., Jovanovic, M. & Slavov, N. Ribosome stoichiometry: from form to function. *Trends Biochem. Sci.* **44**, 95–109 (2019).
3. Krogh, N. et al. Profiling of 2′-O-Me in human rRNA reveals a subset of fractionally modified positions and provides evidence for ribosome heterogeneity. *Nucleic Acids Res.* **44**, 7884–7895 (2016).
4. Taoka, M. et al. Landscape of the complete RNA chemical modifications in the human 80S ribosome. *Nucleic Acids Res.* **520**, 640 (2018).
5. Parks, M. M. et al. Variant ribosomal RNA alleles are conserved and exhibit tissue-specific expression. *Sci. Adv.* **4**, eaao0665 (2018).
6. Gay, D. M., Lund, A. H. & Jansson, M. D. Translational control through ribosome heterogeneity and functional specialization. *Trends Biochem. Sci.* https://doi.org/10.1016/j.tibs.2021.07.001 (2021).
7. Ferretti, M. B. & Karbstein, K. Does functional specialization of ribosomes really exist? *RNA* **25**, 521–538 (2019).
8. Dinman, J. D. Pathways to specialized ribosomes: the Brussels lecture. *J. Mol. Biol.* **428**, 2186–2194 (2016).
9. Sloan, K. E. *et al.* Tuning the ribosome: the influence of rRNA modification on eukaryotic ribosome biogenesis and function. *RNA Biol.* https://doi.org/10.1080/15476286.2016.1259781 (2016).
10. Kiss-László, Z., Henry, Y. & Bachellerie, J. P. Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs. *Cell* **85**, 1077–1088 (1996).
11. Liang, X.-H., Liu, Q. & Fournier, M. J. rRNA modifications in an intersubunit bridge of the ribosome strongly affect both ribosome biogenesis and activity. *Mol. Cell* **28**, 965–977 (2007).
12. Polikanov, Y. S., Melnikov, S. V., Söll, D. & Steitz, T. A. Structural insights into the role of rRNA modifications in protein synthesis and ribosome assembly. *Nat. Struct. Mol. Biol.* **22**, 342–344 (2015).

13. Baudin-Baillieu, A. et al. Nucleotide modifications in three functionally important regions of the *Saccharomyces cerevisiae* ribosome affect translation accuracy. *Nucleic Acids Res.* **37**, 7665–7677 (2009).

14. Birkedal, U. et al. Profiling of ribose methylations in RNA by high-throughput sequencing. *Angew. Chem. Int. Ed. Engl.* **54**, 451–455 (2015).

15. Krogh, N. & Nielsen, H. Sequencing-based methods for detection and quantitation of ribose methylations in RNA. *Methods* **156**, 5–15 (2019).

16. Krogh, N. et al. Profiling of ribose methylations in ribosomal RNA from diffuse large B-cell lymphoma patients for evaluation of ribosomes as drug targets. *Nucleic Acids Res. Cancer* https://doi.org/10.1093/narcan/zcaa035 (2020).

17. Marcel, V. et al. Ribosomal RNA 2′ O-methylation as a novel layer of inter-tumour heterogeneity in breast cancer. *Nucleic Acids Res. Cancer* https://doi.org/10.1093/narcan/zcaa036 (2020).

18. van Riggelen, J., Yetil, A. & Felsher, D. W. MYC as a regulator of ribosome biogenesis and protein synthesis. *Nat. Rev.* **10**, 301–309 (2010).

19. Coller, H. A. et al. Expression analysis with oligonucleotide microarrays reveals that MYC regulates genes involved in growth, cell cycle, signaling, and adhesion. *Proc. Natl Acad. Sci. USA* **97**, 3260–3265 (2000).

20. Makarova, J. A. & Kramerov, D. A. Analysis of C/D box snoRNA genes in vertebrates: the number of copies decreases in placental mammals. *Genomics* **94**, 11–19 (2009).

21. Zeller, K. I. et al. Global mapping of c-Myc binding sites and target gene networks in human B cells. *Proc. Natl Acad. Sci. USA* **103**, 17834–17839 (2006).

22. Shi, Z. et al. Heterogeneous ribosomes preferentially translate distinct subpools of mRNAs genome-wide. *Mol. Cell* **67**, 71–83.e7 (2017).

23. Jack, K. et al. rRNA pseudouridylation defects affect ribosomal ligand binding and translational fidelity from yeast to human cells. *Mol. Cell* **44**, 660–666 (2011).

24. Wohlgemuth, I., Pohl, C. & Rodnina, M. V. Optimization of speed and accuracy of decoding in translation. *EMBO J.* **29**, 3701–3709 (2010).

25. Grosjean, H. & Westhof, E. An integrated, structure- and energy-based view of the genetic code. *Nucleic Acids Res.* **44**, 8020–8040 (2016).

26. Forrest, M. E. et al. Codon and amino acid content are associated with mRNA stability in mammalian cells. *PLoS ONE* **15**, e0228730 (2020).

27. Hia, F. et al. Codon bias confers stability to human mRNAs. *EMBO Rep.* **20**, e48220 (2019).

28. Gingold, H. et al. A dual program for translation regulation in cellular proliferation and differentiation. *Cell* **158**, 1281–1292 (2014).

29. Liu, Q. & Fredrick, K. Intersubunit bridges of the bacterial ribosome. *J. Mol. Biol.* **428**, 2146–2164 (2016).

30. Behrmann, E. et al. Structural snapshots of actively translating human ribosomes. *Cell* **161**, 845–857 (2015).

31. Villa, E. et al. Ribosome-induced changes in elongation factor Tu conformation control GTP hydrolysis. *Proc. Natl Acad. Sci. USA* **106**, 1063–1068 (2009).

32. Paci, M. & Fox, G. E. Centers of motion associated with EF-Tu binding to the ribosome. *RNA Biol.* **13**, 524–530 (2016).

33. Fagan, C. E. et al. Reorganization of an intersubunit bridge induced by disparate 16S ribosomal ambiguity mutations mimics an EF-Tu-bound state. *Proc. Natl Acad. Sci. USA* **110**, 9716–9721 (2013).

34. McClory, S. P., Leisring, J. M., Qin, D. & Fredrick, K. Missense suppressor mutations in 16S rRNA reveal the importance of helices h8 and h14 in aminoacyl-tRNA selection. *RNA* **16**, 1925–1934 (2010).

35. Hoffer, E. D., Maehigashi, T., Fredrick, K. & Dunham, C. M. Ribosomal ambiguity (*ram*) mutations promote the open (off) to closed (on) transition and thereby increase miscoding. *Nucleic Acids Res.* **47**, 1557–1563 (2019).

36. Hanson, G. & Coller, J. Codon optimality, bias and usage in translation and mRNA decay. *Nat. Rev.* **19**, 20–30 (2018).

37. Schuller, A. P., Wu, C. C.-C., Dever, T. E., Buskirk, A. R. & Green, R. eIF5A functions globally in translation elongation and termination. *Mol. Cell* **66**, 194–205.e5 (2017).

38. Juszkiewicz, S. & Hegde, R. S. Initiation of quality control during poly(A) translation requires site-specific ribosome ubiquitination. *Mol. Cell* **65**, 743–750.e4 (2017).

39. Liang, X.-H., Liu, Q. & Fournier, M. J. Loss of rRNA modifications in the decoding center of the ribosome impairs translation and strongly delays pre-rRNA processing. *RNA* **15**, 1716–1728 (2009).

40. Esguerra, J., Warringer, J. & Blomberg, A. Functional importance of individual rRNA 2′-O-ribose methylations revealed by high-resolution phenotyping. *RNA* **14**, 649–656 (2008).

41. Mills, E. W. & Green, R. Ribosomopathies: there's strength in numbers. *Science* **358**, eaan2755 (2017).

42. Imami, K. et al. Phosphorylation of the ribosomal protein RPL12/uL11 affects translation during mitosis. *Mol. Cell* **72**, 84–98.e9 (2018).

43. McMahon, M. et al. A single H/ACA small nucleolar RNA mediates tumor suppression downstream of oncogenic RAS. *Elife* **8**, E8603 (2019).

44. Leppek, K. et al. Gene- and species-specific Hox mRNA translation by ribosome expansion segments. *Mol. Cell* https://doi.org/10.1016/j.molcel.2020.10.023 (2020).

45. Wolf, S., Sameshima, M., Liebhaber, S. A. & Schlessinger, D. Regulation of ribosomal ribonucleic acid levels in growing, ³H-arrested, and crisis-phase WI-38 human diploid fibroblasts. *Biochemistry* **19**, 3484–3490 (1980).

Methods

Cell culture. The following cell lines were used: HeLaS3 cervical adenocarcinoma (American Type Culture Collection (ATCC), catalog no. CCL-2.2); HCT116 and HCT116[p53−/−] colorectal carcinoma[46] (obtained from the laboratory of B. Vogelstein); DLD1 colorectal carcinoma (ATCC, catalog no. CCL-221); BJ[hTERT] immortalized fibroblasts (obtained from the laboratory of K. Helin). HeLaS3 and BJ[hTERT] were the parental lines for HeLaS3[45CKO], HeLaS3-GFP[WT], HeLaS3-GFP[45CKO], HeLaS3-GFP[45CKO+45C] and BJ[MYC] cells, respectively. All cells were grown at 37 °C, in 5% CO_2, as adherent monolayer cultures on Nunclon polystyrene dishes/plates (Nunc). Cells were maintained in DMEM (HeLaS3 and derivatives; and BJ[MYC]) with GlutaMAX and pyruvate (ThermoFisher), McCoy's 5A (HCT116) or RPMI-1640 (DLD1). All media were supplemented with 10% fetal bovine serum.

Generation of BJ[MYC] cells and induction of MYC expression. The coding sequence of *MYC* (NM_002467.4) was cloned from complementary DNA (cDNA) derived from BJ[hTERT]. This was inserted into pLVX-TetOne-Puro (Takara Bio) using the In-Fusion Cloning system (Takara Bio) to give pLVX-TetOne-Puro-MYC. Lentivirus was produced in HEK293 cells (ATCC, catalog no. CRL-1573) according to the Lenti-X TetOne Inducible Expression Systems protocol (Takara Bio). BJ[hTERT] cells were transduced with virus and selected with 0.25 µg ml⁻¹ puromycin (Gibco) for 8 d. The resistant mixed clonal population was then expanded. Puromycin selection was removed at least 72 h before any experiments were performed. For MYC induction in BJ[MYC], cells were pulsed with doxycycline (100 ng ml⁻¹), to induce expression from the TET-responsive promoter, at 0 h and again at 48 h.

Generation of SNORD45C knockout cells. HeLaS3[45CKO] clones were generated using a CRISPR–Cas9 gene-editing approach. To remove the SNORD45C, sequences for sgRNAs were cloned into pX335 and pX458 plasmids[47]. Each plasmid was co-transfected into HeLaS3 cells using Lipofectamine 3000 (ThermoFisher). Single cells positive for GFP and crimson fluorescence were sorted into 96-well plates by FACS. After 2 weeks, clonal colonies were screened for successful deletion using PCR, verified by DNA sequencing and expanded.

Generation of SNORD45C overexpression and control cells. SNORD45C was cloned into intron1 from host gene *RPL23* replacing the endogenous snoRNA. The *RPL23* intron containing SNORD45C was then inserted into a two-exon EGFP sequence derived from pGINT (provided by C. Bellodi), which was used to replace the EGFP from the AAVS1-targeting vector pAAV-PuroCAG-EGFP, obtained from the Vallier laboratory[48]. To generate the overexpression HeLaS3-GFP[45CKO+45C] cells, the artificial EGFP-SNORD45C host gene was inserted into the AAVS1 locus of HeLaS3[45CKO] using zinc-finger nucleases (according to the protocol from Bertero et al.)[48]. To generate the HeLaS3-GFP[45CKO] and HeLaS3-GFP[WT] control cells, the two-exon EGFP without SNORD45C was similarly inserted in HeLaS3[45CKO] and HeLaS3, respectively. The rescue and control cells were FACS-sorted for EGFP expression and sequenced to ensure correct insertion. SNORD45C expression was verified by reverse transcription quantitative PCR (RT–qPCR).

RNA isolation. Large-scale and small-scale total RNA preparations were performed using QIAZOL (Qiagen) or the Direct-zol RNA MiniPrep kit (Zymo Research), respectively, according to the manufacturers' instructions.

RiboMeth-seq. RiboMeth-seq library construction and sequencing were performed as previously described[3,14]. Triplicate libraries were produced for each cell line or condition analyzed, and grown to ~70–80% confluence before collection. A portion of 4–5 µg of total RNA or purified 18S and 28S subunit rRNA (combined at 1:2.6 weight ratios to give an approximate 1:1 molar ratio) were used for input. Note that we have observed no major differences in RMS scores for canonical 2′-O-me sites between libraries constructed from total RNA (as in the latest version of the RMS protocol) or purified subunit rRNA. RNA was partially degraded in alkali at denaturing temperatures. The 20–40-nucleotide fragments were purified by PAGE and linkers added using a system relying on a modified *Arabidopsis* tRNA ligase joining 2′,3′-cyclic phosphate and 5′-phosphate ends. The libraries were sequenced on the Ion Proton platform using Ion PI Chip Kit v.3 (Life Technologies).

RiboMeth-seq data treatment. Data were analyzed as previously reported[3] using custom Python scripts (https://github.com/lundlab/RiboMeth-seq-public). Briefly, sequencing reads were mapped to a corrected human rRNA reference sequence. To facilitate comparison with other studies, we have used the human rRNA sequence numbering according to snoRNABase[49] throughout this study. An alignment table of these rRNA sequences is provided in ref. [3]. The RMS score represents the fraction of molecules methylated at each nucleotide position, calculated by comparing the number of read-end counts at the queried position to six flanking positions on either side. Quantifications were performed on 39 and 68 methylation sites in 18S and 28S rRNAs, respectively, for which both RMS plus MS evidence exist and are reliably detected in at least one of the cell lines examined in this study. For comparison plots of RMS signatures between two cell lines or conditions, for sites exhibiting significant change, in pairwise comparisons

($P < 0.05$, two-tailed unpaired Welch's *t*-test and ≥0.2 difference in RMS score) are labeled with nucleotide position, and *P* values are given in bar plots for specific sites. Heat map representations were produced using the heatmap.2 function in R by unsupervised clustering. RMS data have been deposited to GEO under accession code GSE153476. Note, we have previously published two of the ten RMS datasets included in this study, the data for HelaS3 (WT) and HCT116 (WT) cells[3] are accessible from Gene Expression Omnibus (GEO) GSE76393.

Evolutionary conservation of 2′-O-methylated rRNA nucleotides. Lists of validated 2′-O-methylated positions in 18S and 28S rRNAs, as described for human[3], yeast[14] and mouse[50], were collated and compared. Conserved sites at equivalent nucleotide positions were determined by examination of the respective location in secondary structures of rRNA from each organism. The equivalent sites were then categorized as being present or absent in each organism. For visualization and comparison, positions were rank-ordered according to the RMS score and unsupervised clustering analysis performed on multiple human cell lines. Heat map representations were produced using the heatmap.2 function in R.

rRNA MS analysis. rRNA fragments for MS analysis were isolated and subsequently digested with RNase A or T1, as previously described[14,51]. Digested RNA fragments were subjected to MALDI-TOF MS analyses on an Autoflex Speed MS instrument (Bruker Daltonics) using reflector and positive-ion modes.

Western blotting and antibodies. Western blotting was performed as previously described[52]. The following antibodies were used: MYC D84C12 rabbit (Cell Signaling Technology, catalog no. 5605); FBL (catalog no. sc-25397); p53 (catalog no. sc-126); MDM2 (catalog no. sc-56154); p16 CDKN2A (catalog no. sc-1661); GAPDH (catalog no. sc-25778) (Santa Cruz Biotechnology); p21 CDKN1A (BD Pharmingen, catalog no. 556431); vinculin (Sigma, catalog no. V9131).

siRNA knockdown of *MYC*. HeLaS3 cells were reverse transfected with either Silencer Select Negative Control No. 1 (ThermoFisher, catalog no. 4390843) or Silencer Select MYC siRNA (ThermoFisher, catalog no. 4392420, siRNA identity (ID), s9130) at 7.5 nM final concentration, using Lipofectamine RNAiMAX (ThermoFisher) and according to the manufacturer's protocols. Cells were collected 48 h post transfection and total RNA purified using Direct-zol RNA MiniPrep (Zymo Research), including DNase treatment, according to the manufacturer's protocol.

RT–qPCR. Reverse transcription reactions were performed using the TaqMan Reverse Transcription Reagents (ThermoFisher), according to the manufacturer's instructions. Quantitative real-time PCRs were performed using TaqMan Fast Universal PCR Master Mix (2×), no AmpErase UNG (ThermoFisher) or Fast SYBR Green Master Mix (ThermoFisher) on a StepOne Real-Time PCR System (ThermoFisher) using automatic threshold detection. The following TaqMan Gene Expression Assays were used: ACTB, Hs01060665_g1; CDKN1A, Hs00355782_m1; CDKN2A, Hs02902543_mH; FBL, Hs01070449_m1; MYC, Hs00153408_m1; NPM1, Hs02339479_g1 (ThermoFisher). Primer sequences used in SYBR green qPCR reactions are given in Supplementary Table 5. For individual experiments, reactions were performed in triplicate or duplicate per target gene for each condition. For biological replicates ($n = 3$), the mean Ct values from individual experiments were normalized to the control condition and standardized as previously described[53]. Statistical significance was tested using two-tailed unpaired Welch's *t*-test.

Polysome profiling. Cells at 70–80% confluency were incubated with 100 µg ml⁻¹ cycloheximide (Sigma Aldrich) for 3 min and collected by scraping in PBS containing 100 µg ml⁻¹ cycloheximide. Cells were lysed at 4 °C for 10 min in lysis buffer (20 mM Tris–HCl, 150 mM KCl, 5 mM $MgCl_2$) supplemented with 0.5% NP40 (Igepal, Sigma Aldrich, catalog no. CA-630), 2 mM DTT, 100 µg ml⁻¹ cycloheximide, protease inhibitor cocktail (cOmplete EDTA-free, Roche) and murine RNase inhibitor (NEB). The cell lysate was cleared by centrifugation at 12,000*g* for 15 min at 4 °C. Cleared lysates were normalized according to NanoDrop UV spectrophotometer (ThermoFisher) measurements and layered onto a 7–47% (w/v) linear sucrose gradient (Sigma BioUltra) in polysome buffer. Gradients were centrifuged at 35,000*g* for 3 h at 4 °C in an ultracentrifuge with the SW 40 Ti rotor head (Beckman). Fractions of 1 ml were collected from the top while continuously measuring the absorbance (A_{260}) using the BioLogicP system (BioRad). Absorbance spectra were plotted against time (from 4–16 mins) to generate the polysome profiles.

35S Metabolic labeling assay. For each of three biological replicates, HeLaS3 and HeLaS3[45CKO] were seeded at equal density in six-well plates in triplicate. After 24 hr media were replaced with methionine- and cysteine-deficient media, supplemented with 33 µCi of EasyTag EXPRESS Protein Labeling Mix (containing both [35S]L-methionine and [35S]L-cysteine, Perkin Elmer) and plates were returned to the incubator for 1 h. The cells were then washed and collected by scraping into 1 ml ice-cold PBS, pelleted, snap-frozen and stored at −80 °C until further use. Lysis, and quantification of 35S incorporation by densitometry, were performed

as previously described[54]. Vinculin protein levels were used to normalize ${}^{35}S$ incorporation measurements in each gel lane. Statistical significance was tested using two-tailed unpaired Welch's *t*-test.

Northern blotting analysis of rRNA processing. Total RNA (7.5 µg) was separated on a formaldehyde denaturing 1% agarose gel and transferred to a BrightStar-Plus membrane (Ambion) using capillary blotting. followed by UV cross-linking. The probes (10 pmol each) were radiolabelled with [γ-^{32}P]ATP using T4 PNK (Thermo Scientific) and hybridized to the membrane overnight in hybridization buffer (4× Denhardt's solution, 6× SSC, 0.1% SDS) at T_m of the probe of −10 °C. The membrane was subsequently washed four times in 3× SSC supplemented with 0.1% SDS, followed by exposure to a propidium iodide screen and scanned on a Typhoon scanner (GE Healthcare). Gel images were analyzed using Fiji software.

Ribosome profiling. Ribosome profiling was performed essentially as previously described[55], following the protocol given in TruSeq Ribo Profile Mammalian (Illumina), with minor modifications. Three individual replicates for each of the two cell lines were collected. A single 15-cm dish corresponding to one replicate was collected at a time. For each replicate, cell media was aspirated, and cells washed with ice-cold PBS. No cycloheximide pretreatment was performed. After thorough removal of PBS, the dish was fully immersed in liquid nitrogen and placed on dry ice. For cell lysis, 1 ml of 1× Mammalian Lysis Buffer (Illumina) containing 100 µg ml^{-1} cycloheximide was added dropwise to the dish, which was then placed on wet ice. Cells were then scraped off to the lower portion of the dish and allowed to thaw in the lysis buffer. Lysate was homogenized by pipetting and triturated ten times through a 25-gauge needle. The lysate was then transferred to a DNA LoBind 1.5-ml microfuge tube (Eppendorf) and incubated on ice for 5 min. The lysate was cleared by centrifugation at 20,000*g*, 4 °C for 10 min and the supernatant transferred to a fresh microfuge tube. Aliquots were prepared for each replicate, flash-frozen in liquid nitrogen and stored at −80 °C until further use. The steps detailed in TruSeq Ribo Profile Mammalian protocol (Illumina) were followed to generate total RNA and ribosome protected fragment (RPF) RNA-seq libraries corresponding to the three individual replicates from each of the two cell lines. For RPF libraries, following nuclease digestion, monosomes were purified using Illustra MicroSpin S-400 HR Columns. Ribo-Zero Gold Kit (Illumina) was used to deplete ribosomal rRNA. The libraries prepared from total RNA or RPF for both conditions were pooled and sequenced on a NextSeq 500 System (Illumina).

Ribosome profiling data analysis. The sequencing data were demultiplexed using Illumina bcl2fastq. Quality of the sequencing files was controlled with fastqc. Adapter sequences were removed with cutadapt. Reads derived from RPF and total RNA were aligned to human rRNA and tRNA sequences with bowtie2 (v.2.2.9) and the mapped reads discarded. The remaining reads were aligned to GRCh38. p12 (Ensembl v.97) with Spliced Transcripts Alignment to a Reference (STAR) software to both transcriptome and genome coordinates. Reads mapping to Human Genome Organisation (HUGO) approved genes were used for downstream analyses. RPF read lengths were analyzed for trinucleotide periodicity using Ribotaper[56]. RPF reads with lengths between 29 and 34 nucleotides were selected and the optimal P-site offset was defined as position 12 from 5′ read ends. RPF read alignment files were filtered with samtools to retain only 29–34-nucleotide read lengths, and no read length filtering was applied to the total RNA alignment files. The transcript coordinate alignment files obtained from STAR were converted into genomic coordinate alignments with RSEM. A single canonical transcript representing each protein-coding gene was selected from the GRCh38, v.97 Ensembl annotation file (Supplementary Table 6). FeatureCounts was used to generate counts of reads mapping to exons of these transcripts for both total RNA and RPF. RPF reads with ribosome P-site positions mapping within transcript coding region sequences (CDS) were again counted using FeatureCounts and, along with the mRNA exons mapped reads, used for further measurements of differential translation and mRNA expression. Ribosomal investigation and visualization to evaluate translation (RIVET)[57] was used for translation and expression analysis of the representative transcripts (similar results were obtained for gene-level analysis). No fold change cut-offs were directly applied so that more subtle changes in translation could also be detected. Regulated transcripts were therefore nominally identified by statistical significance. Translation regulation categories were defined according to RIVET on the basis of mRNA expression and ribosome occupancy, derived from normalized total RNA read counts or RPF read counts mapping to protein-coding mRNA transcripts, respectively (Supplementary Table 2). Plots from the resulting RIVET output files were generated using the ggplot2 package in R. The RNA sequencing data have been deposited to GEO under accession code GSE153476.

GO and GSEA. All GO analyses of ribosome profiling and proteomic data were performed using WebGestalt using the over-representation test against the GO biological process database[58]. GSEA was performed using GSEA software (v.4.0.3). Genes for analysis were preranked according to $\log_2$ FC in translation from RIVET analysis ($P_{adj} < 0.05$) and compared to a custom gene set (containing 112 distinct human genes) compiled from those previously described to contain

experimentally validated IRES elements, or that associated with polysomes following cap-dependent translation inhibition[59].

Metagene analysis. For metagene analyses, bam files containing exon-mapped reads for each library were converted to normalized reads per kilobase per million (RPKM) or counts per million reads (CPM) single-nucleotide resolution coverage bigwig files, with bamCoverage from the deepTools suite[60]. WiggleTools[61] (Ensembl) and wigToBigWig (Encode, kentUtils) were then used to merge these and create mean coverage files per condition. These were input to deepTools computeMatrix, together with an annotation file containing the exon coordinates for the selected mRNA transcripts. For RPF coverage over all transcripts, a count matrix was then generated for library RPKM RPF coverage over the coding regions (CDS), scaled to size 100 nt, flanked by unscaled regions before and after the translation start (TSS) and end (TES) sites. For further analysis, the scaled coverages of transcripts comprising the different translationally regulated categories were extracted from this matrix and median values at each position plotted. For average ribosome occupancy, CPM normalization was used and offset applied using bamCoverage, so as to use only the nucleotide position representing the ribosome P-site for each read as the signal (see 'Ribosome profiling data analysis', above). The P-site coverage files were input to computeMatrix and a count matrix generated for −30 to +330 or −330 to +30 nucleotides, relative to the CDS start or end site respectively for each transcript (unscaled). The resulting counts at each position were divided by the total RPF count in CDS for each corresponding transcript to give the average ribosome occupancy per nucleotide position in each transcript. The mean values at each equivalent nucleotide position relative to the translation start site were plotted after extreme outlier removal (>3x interquartile range), no smoothing was applied. For P-site CPM the same matrices were used, although here the counts at each position were summed at each nucleotide position. For plotting, extreme outliers (>3x interquartile range) were removed. Plots were produced using ggplot2 in R.

Cell proliferation assay by crystal violet staining. Cells were seeded in triplicate at equal density in 12-well plates. At each time point, cells were washed and then fixed with 4% paraformaldehyde. Subsequently, crystal violet staining was performed for 30 min at room temperature, followed by three washes in PBS. The staining was dissolved in 10% acetic acid and absorbance (at 600 nm) measured using a GloMaxMulti Detection plate reader (Promega).

EdU incorporation assay and flow cytometry analysis. For each of three biological replicates, HeLaS3 and HeLaS3^{45CKO} were seeded in 6-cm dishes in triplicate at densities resulting in equal confluence (approximately 70%) at collection. At 48 h later, cells were pulsed with 30 µM EdU for 40 min and then collected by trypsinization and washed 3× in PBS with 1% BSA. EdU detection was performed using the Click-iT EdU Alexa Fluor 647 Flow Cytometry Assay Kit (ThermoFisher) according to the manufacturer's instructions, and incorporation was assessed by flow cytometry analysis of 20,000 cells per condition. Statistical significance of cell cycle phase (G1, S, G2/M) data was determined using two-way analysis of variance (ANOVA) with Sidak correction for multiple comparisons. For propidium iodide staining, BJ^{MYC} cells were collected, washed, fixed in 70% ethanol, washed again and counted before incubation in FxCycle PI/RNase Staining Solution (ThermoFisher). Analysis of all FACS data was performed using the FlowJo software package (Tree Star).

Metabolic analysis of oxygen consumption rate. Cellular oxygen consumption rate (OCR) was measured using a Seahorse 96XF analyzer, utilizing the mitochondrial stress test according to the manufacturer's protocol. Cells were plated 24 h before the experiment. HeLaS3, HeLaS3-GFP^{WT} and HeLaS3${}^{45CKO+45C}$ were seeded at 10,000 cells per well, while HeLaS3^{45CKO} and HeLaS3-GFP^{45CKO} were seeded at 14,000 cells per well (seeding densities were optimized to ensure equal confluence after 24 h, and that OCR measurements were within the linear detection rate). At 1 h before the experiment, the cells were washed and media exchanged into Seahorse assay medium (DMEM with 3 mg l^{-1} phenol red and 5.5 mM glucose, pH 7.4) and kept at 37 °C. Four measurements were performed for each condition: baseline, oligomycin A (final concentration 1 µM), FCCP (final concentration 0.2 µM) and antimycin A/rotenone (final concentration 1.0 and 0.5 µM, respectively). Concentrations of oligomycin A, FCCP and antimycin A/rotenone were optimized with dose–response experiments to ensure optimal conditions for the assay. OCR measurements were analyzed similarly to previously described[62]. Each condition was performed with 16 technical replicates in each experiment, the data are presented as the mean of four independent experiments. Statistical significance was tested using repeated measures one-way ANOVA with Turkey's correction for multiple comparisons for experiments with three sample groups, or two-tailed unpaired Welch's *t*-test for two sample groups.

Reporter assays for IRES-dependent translation. HeLaS3 and HeLaS3^{45CKO} cells were seeded in a 96-well plate and 24 h later co-transfected with firefly luciferase reporter constructs pIRES-c-myc or pBIC-943-IGF1R[63] (IRES-dependent), along with pRLTK *Renilla* luciferase control (5′ cap-dependent), using Lipofectamine 3000 (ThermoFisher). Luciferase assays were performed 24 h post transfection.

In each experiment, firefly and *Renilla* luciferase activities were measured for quadruplicate transfections with the Dual-Glo Luciferase Assay System (Promega) using the GloMaxMulti Detection System (Promega). Ratios of firefly to *Renilla* luciferase activity were then calculated to assess 5′ cap- versus IRES-mediated translation initiation in each condition. Statistical significance was tested using two-tailed unpaired Welch's *t*-test.

Codon usage analysis. Coding sequences for transcripts mapped in ribosome profiling analysis were retrieved from Ensembl (GRCh38.p12, v.97). After exclusion of the start codon the codon usage frequency and GC content were calculated using the uco and GC functions from the seqinr R package. The codon usage frequencies or GC content were then grouped according to the regulation categories defined in the analysis of the ribosome profiling data and compared between conditions. PCA was executed with the R function prcomp, with the transcripts as observations and codons as variables. To evaluate which codons contributed the most to the different codon usage profiles, the mean and median of each codon was calculated for the above-mentioned categories. The $\log_2$ fold changes were calculated from the means of the different groups. P values for individual codons and GC content were determined by Wilcoxon test (unpaired, two-sided) and adjusted for multiple testing using the Bonferroni correction. Codons with $\log_2$(fold change in frequency) > 1, P_{adj} < 0.001, were selected for further analysis. Codons enriched in each translation category were then classified as 'proliferation', 'differentiation' or 'other' according to a previous study focused on tRNAs[28]. Codons were designated as having strong, weak or intermediate interactions with their cognate tRNAs, as defined in previous analysis[25] (Supplementary Table 3).

Proteomic MS analysis. HeLaS3 and HeLaS3[45CKO] cells were grown to ~70% confluency. A single 10-cm dish was collected for each of four biological replicates in each case. Cells were washed, collected by scraping into PBS, pelleted by centrifugation at 100*g* for 4 min and the supernatant removed. Samples were flash-frozen and stored at −80 °C until further use. Lysate preparation and digestion was similar to that previously described[64]. Briefly, cell pellets were lysed using 300 μl guanidinium hydrochloride lysis buffer (6 M GdnHCl, 100 mM Tris–HCl pH 8.5, 5 mM TCEP, 10 mM CAA). Samples were sonicated on high for 5× 30 s in a Bioruptor sonication water bath (Diagenode) at 4 °C. After determining sample protein concentration with Bradford reagent (Sigma), 10 μg of each was taken forward for digestion. Samples were diluted 1:3 with 10% acetonitrile, 50 mM HEPES pH 8.5 and LysC (MS grade, Wako) was added in a 1:50 (enzyme to protein) ratio, and samples were incubated at 37 °C for 4 h. Samples were further diluted to 1:10 with 10% acetonitrile, 50 mM HEPES pH 8.5 and trypsin (MS grade, Promega) was added in a 1:100 (enzyme:protein) ratio and samples were incubated overnight at 37 °C. Enzyme activity was quenched by adding 2% trifluoroacetic acid (TFA) to a final concentration of 1%. Before TMT labeling, the peptides were desalted on in-house packed C18 StageTips[65]. For each sample, two disks of C18 material (3M Empore) were packed in a 200-μl tip, and the C18 material activated with 40 μl of 100% methanol (HPLC grade, Sigma), then 40 μl of 80% acetonitrile, 0.1% formic acid. The tips were subsequently equilibrated twice with 40 μl of 1%TFA and 3% acetonitrile, after which 10 μg of sample was loaded using centrifugation at 4,000 r.p.m. After washing the tips twice with 100 μl of 0.1% formic acid, the peptides were eluted into clean 500-μl eppendorf tubes using 40% acetonitrile, 0.1% formic acid. The eluted peptides were concentrated in an Eppendorf Speedvac (ThermoFisher) and reconstituted in 50 mM HEPES (pH 8.5) for TMT labeling[66]. Labeling was done according to the manufacturer's instructions, and, subsequently, labeled peptides were mixed 1:1:1:1:1:1:1:1:1:1:1 (11-plex), acidified with 1% TFA and acetonitrile concentration brought down to <5% using 2% TFA. Before MS analysis, the peptides were fractionated using an offline ThermoFisher Ultimate 3000 liquid chromatography system using high pH fractionation (5 mM ammonium bicarbonate, pH 10) at 5 μl min⁻¹ flow rate. Peptides (30 μg portions) were separated over a 120-min gradient (5% to 35% acetonitrile) while collecting fractions every 120 s. The resulting 60 fractions were pooled into 44 final fractions and vacuum concentrated to dryness. Fractions were resuspended in 1% TFA, 2% acetonitrile for MS analysis. Analysis was performed on an Orbitrap Fusion mass spectrometer. For each fraction, peptides were loaded onto a 2-cm C18 trap column (ThermoFisher, catalog no. 164705) connected in-line to a 50-cm C18 reverse-phase analytical column (Thermo EasySpray ES803) using 100% buffer A (0.1% formic acid in water) at 750 bar, using the Thermo EasyLC 1200 HPLC system and the column oven operating at 45 °C. Peptides were eluted over a 60-min gradient ranging from 10 to 60% of 80% acetonitrile, 0.1% formic acid at 250 nl min⁻¹, and the Orbitrap Fusion instrument (ThermoFisher Scientific) was run in an SPS-MS3 top speed mode. Full MS spectra were collected at a resolution of 120,000, with an AGC target of 2×10^5 or maximum injection time of 50 ms and a scan range of 380–1,500 *m/z*. The MS² spectra were obtained in the ion trap operating at rapid speed, with an AGC target value of 1×10^4 or maximum injection time of 50 ms, a normalized CID collision energy of 35 and an intensity threshold of 5×10^3. Dynamic exclusion was set to 60 s, and ions with a charge state <2, >6 or unknown were excluded. From the resulting MS² scan, ten precursors were selected for SPS-MS3 analysis, fragmented with a normalized HCD collision energy of 65 and ions collected for a maximum of 86 ms or AGC target of 1×10^5. Resulting MS3 spectra were collected at 50,000 resolution and scan range of 100–500 for reporter ion quantification. MS performance was verified for consistency by running complex cell lysate quality control standards, and chromatography was monitored to check for reproducibility. All MS spectra were searched in Proteome Discoverer 2.4 (ThermoFisher) using the SEQUEST algorithm against the human proteome Uniprot database (containing its reversed complement and known contaminants). Spectral matches were filtered to false discovery rate (FDR) < 0.01 using the target-decoy strategy combined with linear discriminant analysis. Proteins were quantified only from peptides with an average reporter S/N threshold of 10, and co-isolation specificity of 0.75. Statistical analysis of protein abundance changes was performed using the DEqMS pipeline for TMT-labeled MS data (https://github.com/yafeng/DEqMS)[67], with q value <0.05. The $\log_2$ FC changes in protein abundance were determined to be significant at sca.adj.pval (Benjamini–Hochberg method adjusted DEqMS P values) <0.05 (Supplementary Table 4). The MS data have been deposited to the ProteomeXchange Consortium, via the PRIDE repository with the dataset identifier PXD019449.

Generation of codon usage reporter plasmids. The reporter cassette from plasmid pmGFP-P2A-K0-P2A-RFP[38] (Addgene, plasmid no. 105686) was amplified by PCR. The backbone of plasmid pLVX-TetOne-puro (Takara Bio) was amplified by PCR, so as to exclude the puromycin resistance gene, which was then replaced with the G418 resistance gene. The TetOne promoter cassette was amplified by PCR and all segments were assembled together with NEBuilder HiFi DNA Assembly kit (NEB). The resulting reporter plasmid pLVX-TetOne-G418-EGFP-mCherry was amplified by PCR with primer to exclude EGFP, which was replaced with mTagBFP2 to give pLVX-TetOne-G418-mTagBFP2-mCherry. Two separate 40-codon long inserts were designed comprising either eight AU-rich codons enriched in the TL-DN transcript set (AU-DN) or eight GC-rich codons (GC-UP) in the TL-UP set (Supplementary Tables 3 and 5). The order of specific codons was shuffled in each case to minimize secondary structure formation. AU-DN and GC-UP codon cassettes were ordered as gBlocks from Integrated DNA Technologies and inserted into the mTagBFP2-mCherry reporter plasmid, in frame, between the fluorescent reporter genes using the NEBuilder HiFi DNA Assembly kit, to give pLVX-TetOne-G418-AU-TL-DN and pLVX-TetOne-G418-GC-TL-UP plasmids. All assembly products were verified by sequencing.

Generation of codon usage reporter cell lines. Lentivirus was produced in HEK293 cells, using either the pLVX-TetOne-G418-AT-DN or pLVX-TetOne-G418-GC-UP plasmids, according to the Lenti-X TetOne Inducible Expression Systems protocol (Takara Bio). HeLaS3-GFP[WT], HeLaS3-GFP[45CKO] and HeLaS3-GFP[45CKO+45C] cells were then transduced with each of the codon usage reporter lentiviruses to derive six new AT- and GC-rich codon usage reporter cell lines: HeLaS3-GFP[WT-GC-UP]; HeLaS3-GFP[45CKO-GC-UP]; HeLaS3-GFP[45CKO+45C-GC-UP]; HeLaS3-GFP[WT-AU-DN]; HeLaS3-GFP[45CKO-AU-DN] and HeLaS3-GFP[45CKO+45C-AU-DN]. Cells carrying the codon usage reporter cassette were positively selected on the basis of G418 antibiotic resistance.

High-content imaging assays of codon reporter cell lines. For each of the six codon usage reporter cell lines, ten wells were seeded in 96-well plates. Media in nine of these wells contained doxycycline (100 ng ml⁻¹), with no doxycycline in the remaining well. After 24 h, fluorescent imaging was performed on live cells with an INCell2200 high-content imager (GE Healthcare), using a ×20 objective capturing 60 fields in each well. mTagBFP fluorescence was captured using the DAPI channel, GFP using the FITC channel and mCherry using the TexasRed channel. Image analysis was performed with Incell Analyzer 1000 workstation software (GE Healthcare). Cells were segmented on the basis of the mCherry signal using the Tophat segmentation method, and the mean intensity per cell of both the BFP and mCherry was measured. mCherry segmentation was used rather than mTagBFP, due to higher signal to background levels. For each cell line, thresholds for cells positive for mTagBFP and mCherry expression were defined using intensity measurements from the wells not treated with doxycycline. Further data analysis and plotting was performed using R. Cells that were positive for both mTagBFP and mCherry were selected, and the ratio of mTagBFP:mCherry calculated from the intensity measurements for individual cells per condition. Cells with outlying ratios (>1.5× interquartile range), probably resulting from imaging artifacts, were excluded from downstream analysis. Box plots of mTagBFP:mCherry ratios per cell for each condition were generated using the ggplot2 package in R. Statistical significance was tested using two-tailed unpaired Wilcoxon test with 95% confidence interval, and the resulting P values adjusted for multiple comparisons with the Bonferroni correction method. Statistical analysis was performed using the rstatix package in R.

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

Data availability

All RNA sequencing data from this study have been deposited in the GEO under series accession numbers GSE153476 (for RNA-seq data) and GSE76393

(for RiboMeth-seq data). The MS proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE repository with the dataset identifier PXD019449. Source data are provided with this paper.

Code availability

Python code for analysis of RiboMeth-seq data is available in the GitHub repository https://github.com/lundlab/RiboMeth-seq-public

References

46. Bunz, F. et al. Requirement for p53 and p21 to sustain G2 arrest after DNA damage. *Science* **282**, 1497–1501 (1998).
47. Cong, L. et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823 (2013).
48. Bertero, A. et al. Optimized inducible shRNA and CRISPR/Cas9 platforms for in vitro studies of human development using hPSCs. *Development* **143**, 4405–4418 (2016).
49. Lestrade, L. snoRNA-LBME-db, a comprehensive database of human H/ACA and C/D box snoRNAs. *Nucleic Acids Res.* **34**, D158–D162 (2006).
50. Hebras, J., Krogh, N., Marty, V., Nielsen, H. & Cavaillé, J. Developmental changes of rRNA ribose methylations in the mouse. *RNA Biol.* **17**, 150–164 (2020).
51. Douthwaite, S. & Kirpekar, F. Identifying modifications in RNA by MALDI mass spectrometry. *Methods Enzymol.* **425**, 3–20 (2007).
52. Jansson, M. D., Damas, N. D., Lees, M., Jacobsen, A. & Lund, A. H. miR-339-5p regulates the p53 tumor-suppressor pathway by targeting MDM2. *Oncogene* **34**, 1908–1918 (2015).
53. Willems, E., Leyns, L. & Vandesompele, J. Standardization of real-time PCR gene expression data from independent biological replicates. *Anal. Biochem.* **379**, 127–129 (2008).
54. Cunningham, J. T., Pourdehnad, M., Stumpf, C. R. & Ruggero, D. Investigating Myc-dependent translational regulation in normal and cancer cells. *Methods Mol. Biol.* **1012**, 201–212 (2013).
55. Ingolia, N. T., Ghaemmaghami, S., Newman, J. R. S. & Weissman, J. S. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* **324**, 218–223 (2009).
56. Calviello, L. et al. Detecting actively translated open reading frames in ribosome profiling data. *Nat. Methods* **13**, 165–170 (2016).
57. Ernlund, A. W., Schneider, R. J & Ruggles, K. V. RIVET: comprehensive graphic user interface for analysis and exploration of genome-wide translatomics data. *BMC Genomics* **19**, 809 (2018).
58. Liao, Y., Wang, J., Jaehnig, E. J., Shi, Z. & Zhang, B. WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res.* **47**, W199–W205 (2019).
59. Weingarten-Gabbay, S. et al. Comparative genetics. Systematic discovery of cap-independent translation sequences in human and viral genomes. *Science* **351**, aad4939 (2016).
60. Ramírez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).
61. Zerbino, D. R., Johnson, N., Juettemann, T., Wilder, S. P. & Flicek, P. WiggleTools: parallel processing of large collections of genome-wide datasets for visualization and statistical analysis. *Bioinformatics* **30**, 1008–1009 (2014).
62. Zhang, J. et al. Measuring energy metabolism in cultured cells, including human pluripotent stem cells and differentiated cells. *Nat. Protoc.* **7**, 1068–1085 (2012).
63. Belin, S. et al. Dysregulation of ribosome biogenesis and translational capacity is associated with tumor progression of human breast cancer cells. *PLoS ONE* **4**, e7147 (2009).
64. Kulak, N. A., Pichler, G., Paron, I., Nagaraj, N. & Mann, M. Minimal, encapsulated proteomic-sample processing applied to copy-number estimation in eukaryotic cells. *Nat. Methods* **11**, 319–324 (2014).
65. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896–1906 (2007).
66. Thompson, A. et al. Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal. Chem.* **75**, 1895–1904 (2003).
67. Zhu, Y. et al. DEqMS: a method for accurate variance estimation in differential protein expression analysis. *Mol. Cell. Proteom.* **19**, 1047–1057 (2020).

Acknowledgements

Author contributions

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41594-021-00669-4.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41594-021-00669-4.

Correspondence and requests for materials should be addressed to Martin D. Jansson or Anders H. Lund.

Peer review information *Nature Structural & Molecular Biology* thanks U. Thomas Meier and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available. Anke Sparmann and Carolina Perdigoto were the primary editors on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Reprints and permissions information is available at www.nature.com/reprints.
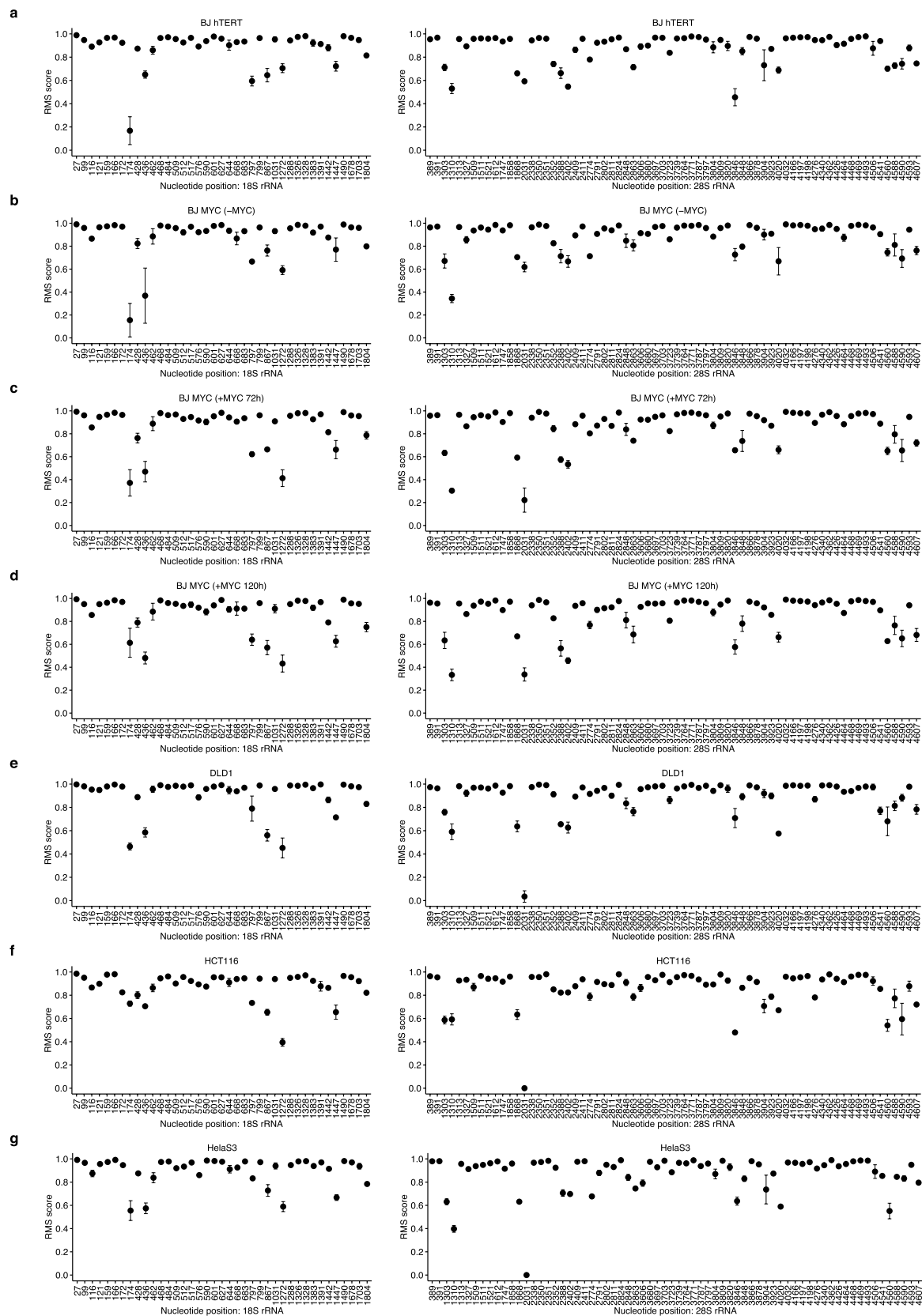
Extended Data Fig. 1 | Detection and quantification of rRNA 2'-O-me across multiple cell types. a-g, RiboMeth-seq (RMS) scores representing fraction of 2'-O-methylation at each site in 18S (left) and 28S rRNAs (right) present in total RNA purified from the indicated cell lines. Nucleotide positions in respective rRNAs are indicated. Points represent mean RMS scores of $n = 3$ sequenced libraries from individual cell cultures. Error bars represent ± s.d.. Note that (f) and (g) are derived from datasets we published previously (Krogh and Jansson *et al.*), included here for comparison.
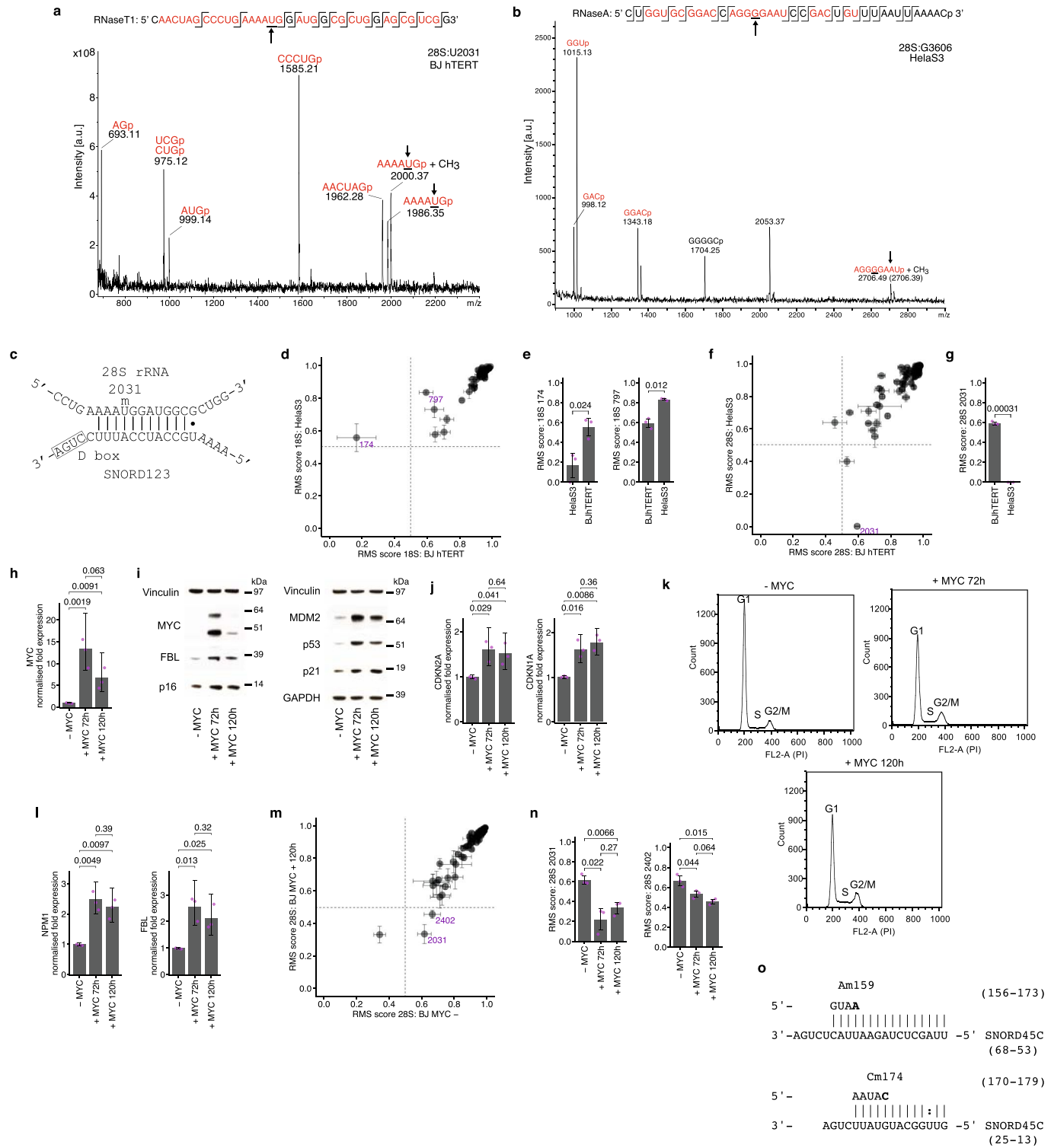
Extended Data Fig. 2 | See next page for caption.

**Extended Data Fig. 2 | Variation in 2'-O-me levels sites exists at specific sites in cell model systems. a**, MS analysis of 28S:U2031 from fragment generated by RNase T1 digestion. Both methylated and unmethylated fragments are present in the spectra, as expected from the RMS score. The sequence of the fragment is shown above the spectra, with an indication of the expected RNase cleavage sites. All identified fragments are labelled in red. Arrows mark 28S:U2031 position and spectra peaks. Observed mass is shown below each fragment. **b**, As (a) for MS analysis of 28S:G3606 from fragment generated by RNase A digestion. Theoretical mass (in brackets) is shown for fragment containing Gm3606. **c**, Schematic showing base pairing of SNORD123 and target region in 28S rRNA containing U2031. **d**, Comparison of RiboMeth-seq (RMS) scores representing fraction of 18S rRNA molecules 2'-O-methylated at each site in 18S rRNA from HelaS3 and BJ[hTERT] cells. Points denote mean RMS score ($n=3$ libraries from individual cultures) per site. Horizontal and vertical error bars indicate ±s.d. for each condition. Sites exhibiting significant change are labelled with nucleotide position (magenta). **e**, RMS scores for 2'-O-me at 18S:174 and 18S:G683 in HelaS3 and BJ[hTERT] cells. Columns indicate mean RMS score for each condition of $n=3$ libraries from individual cultures, points (magenta) denote each value separately. Error bars represent ±s.d. *P* values (Welch's *t*-test, two-tailed) are shown above brackets. **f**, As (d), for 28S rRNA. **g**, As (e), for 2'-O-me at 28S:U2031. **h**, Expression of *MYC* mRNA in BJ[MYC] cells. RT-qPCR analysis where MYC expression is not induced, or induced for 72 or 120 h as indicated. Columns indicate mean ($n=3$ independant experiments) fold-change relative to the control condition, normalized to ACTB expression. Error bars indicate 95% confidence intervals. *P* values (Welch's *t*-test, two-tailed) are shown for pairwise comparisons, indicated by brackets. **i**, Western blot analysis in BJ[MYC] cells, where MYC expression is not induced, or induced for 72 or 120 h as indicated. Protein levels of MYC, FBL, p53, MDM2, p21 and p16 are shown. Vinculin and GAPDH are included as loading controls. Positions of molecular weight markers (kDa) are indicated. Data is representative of $n=2$ independent experiments. **j**, As (h), for expression of cell-cycle inhibitors *CDKN2A* (p16) and *CDKN1A* (p21) in BJ[MYC] cells. **k**, Cell cycle profiles of BJ[MYC] cells where MYC expression is not induced, or induced for 72 or 120 h as indicated. FACS analysis of propidium iodide (PI) stained cells, showing proportion of cells (count) against PI intensity (FL-2A). Representative data ($n=2$ independent experiments) is shown. **l**, As (h), for expression of MYC target genes *NPM1* and *FBL* in BJ[MYC] cells. **m**, As (d), for 28S rRNA from BJ[MYC] cells, where MYC expression is induced for 120 h or not induced. **n**, As (e), for 2'-O-me at 28S:U203 and 28S:U2402 in BJ[MYC] cells, without MYC induction or with MYC induction for 72 or 120 h. **o**, Sequence alignment of SNORD45C guide regions with putative target sequences in 18S rRNA. Position of methylation sites at A159 (top) and C174 (bottom) is indicated by *.
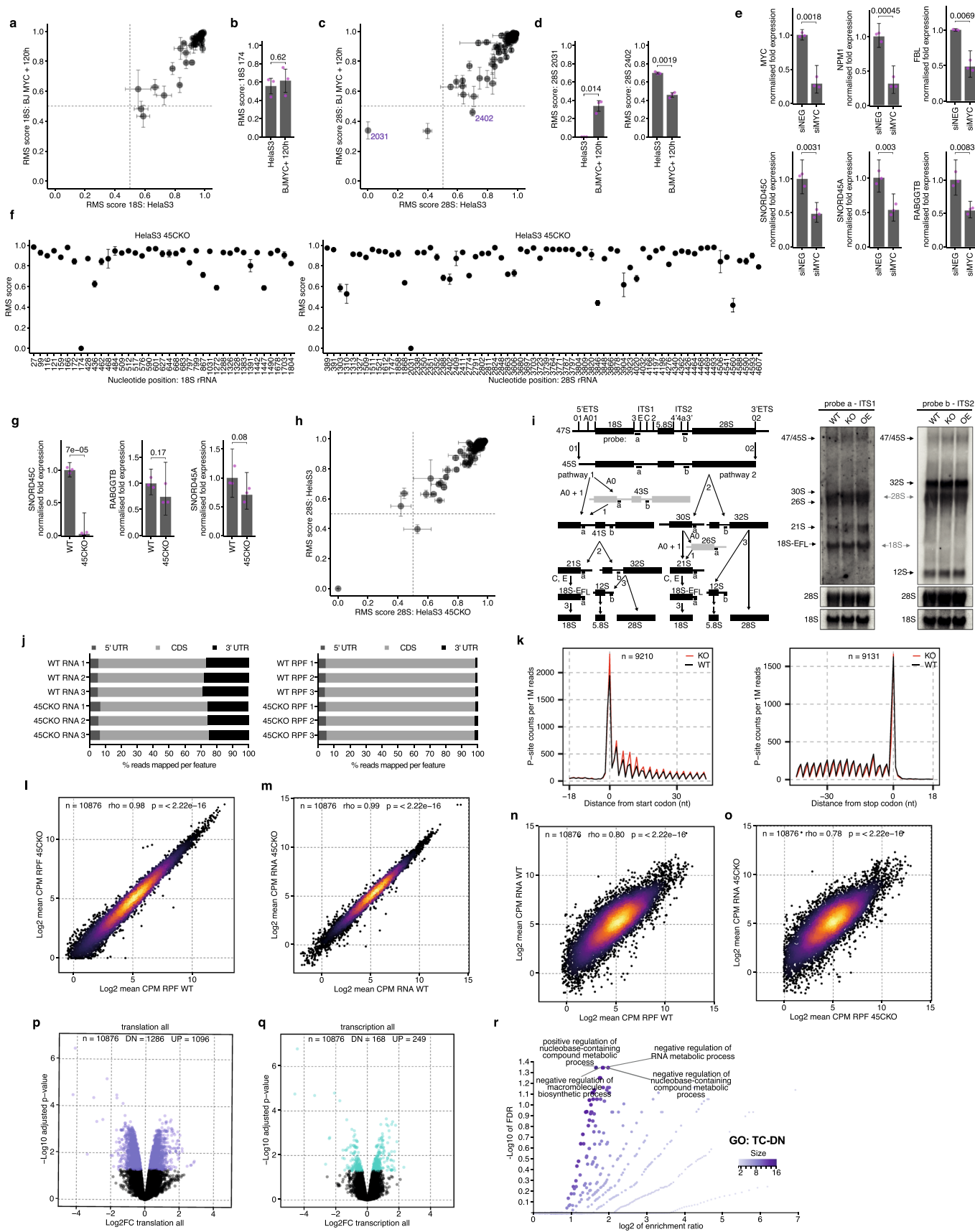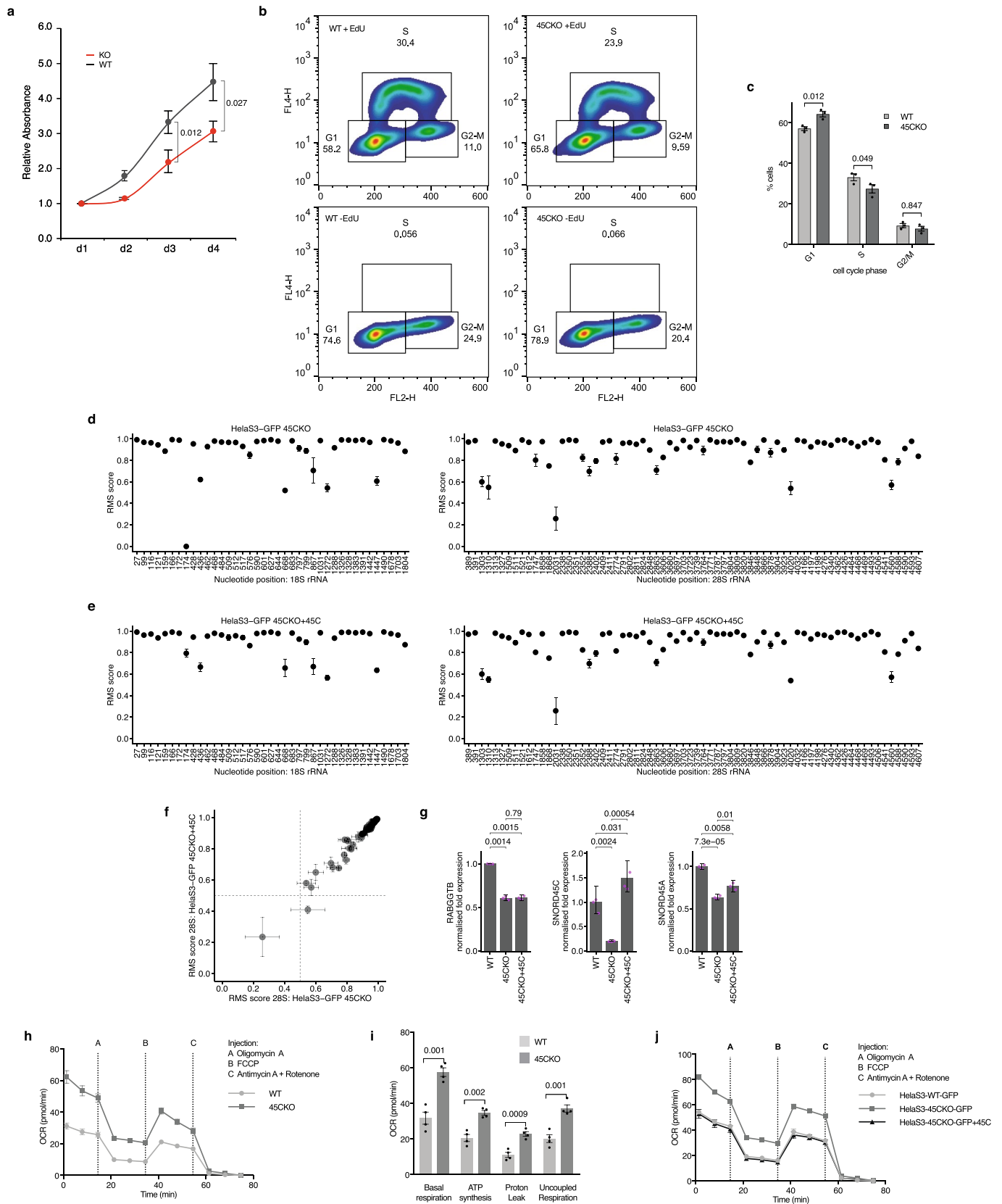
Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Development and analysis of a model to elucidate the role of Cm174 in ribosome regulation. a, Comparison of RiboMeth-seq (RMS) scores representing fraction of 18S rRNA molecules 2′-O-methylated at each site in HelaS3 and BJ^{MYC} cells, where MYC expression is induced for 120 h or not induced. Points denote mean RMS score ($n = 3$ libraries from individual cultures) per site. Horizontal and vertical error bars indicate ±s.d. for each condition. Sites exhibiting significant change are labelled with nucleotide position (magenta). **b**, RMS scores for 2′-O-me at 18S:174 in HelaS3 and BJ^{MYC} cells, where MYC expression is induced for 120 h or not induced. Columns indicate mean RMS score for each condition of $n = 3$ libraries from individual cultures, points (magenta) denote each value separately. Error bars represent ±s.d. P values (Welch's t-test, two-tailed) are shown above brackets. **c**, As (a), for 28S rRNA. **d**, As (B) for 2′-O-me at 28S:U2031 and 28S:U2402. **e**, siRNA-mediated knock-down of *MYC* in HelaS3 (WT) cells. Expression of *MYC*, and canonical target genes *NPM1* and *FBL*, as well as SNORD45C, SNORD45A and host gene *RABGGTB*, as determined by RT-qPCR analysis. Columns indicate mean ($n = 3$ independent experiments) fold-change relative to the control condition, normalized to U6 expression. Error bars indicate 95% confidence intervals. P values (Welch's t-test, two-tailed) are shown for pairwise comparisons, indicated by brackets. **f**, RiboMeth-seq (RMS) scores representing fraction of 2′-O-methylation at each site in 18S (left) and 28S rRNAs (right) from HelaS3^{45CKO} cells. Nucleotide positions in respective rRNAs are indicated. RMS scores are means of $n = 3$ sequenced libraries from individual cell cultures. Error bars represent ± s.d. **g**, Expression of SNORD45C, SNORD45A and host gene *RABGGTB* in HelaS3 (WT) and HelaS3^{45CKO} (45KO) cells, as determined by RT-qPCR analysis. Columns indicate mean ($n = 3$ independent experiments) fold-change relative to the control condition, normalized to U6 expression. Error bars indicate 95% confidence intervals. P values (Welch's t-test, two-tailed) are shown for pairwise comparisons, indicated by brackets. Note, due to high sequence similarity between SNORD45C and -A, a residual signal for SNORD45C is apparent in 45KO. **h**, As (a), for 28S rRNA in HelaS3 and HelaS3^{45CKO} cells. No sites exhibit significant change. **i**, Analysis of rRNA biogenesis pathways in HelaS3^{GFP} (WT), HelaS3$^{45CKO-GFP}$ (KO) and HelaS3$^{45CKO-GFP+45C}$ (OE) cells, derived from parental HelaS3 and HelaS3^{45CKO} (please refer to section 'Cm174 affects cellular phenotypes' in the main text where they are formally introduced and described in more detail). Schematic (left) showing rRNA processing steps and location of probes 'a' and 'b'. Levels of pre-rRNA and processing intermediates as assessed by northern blot. Data shown is representative of $n = 2$ independent experiments. **j**, Percentage reads mapping to 5′ untranslated regions (5′ UTR), coding sequences (CDS) or 3′ untranslated regions (3′ UTR) of protein-coding genes, for each replicate sequencing library. Both total RNA (left) and ribosome protected fragment (right) derived reads are shown. **k**, Periodicity of ribosome protected fragment (RPF) reads mapped to mRNA transcripts. Metagene analysis shows normalized mean counts, at single-nucleotide resolution, representing ribosome P-site occupancy relative to start (left) and stop (right) codons from HelaS3 (WT, black) or HelaS3^{45CKO} (KO, red) libraries ($n = 3$ libraries from individual cultures). Number of transcripts analyzed after extreme outlier removal is given by 'n'. **l**, Correlation between reads mapped per mRNA transcript in HelaS3 (WT) and HelaS3^{45CKO} (KO) RPF libraries ($n = 3$ libraries from individual cultures). Normalized (CPM) mean counts are plotted. Color scale indicates plotting density. 'n' denotes number of transcripts analyzed. Spearman's rho and associated P value (algorithm AS 89) are shown. **m**, As (k), for reads mapping to mRNA from total RNA libraries. **n**, As (k), for correlation between reads mapping to mRNA transcripts from RPF and total RNA libraries in HelaS3 (WT) cells. **o**, As (m), for HelaS3^{45CKO} (KO) cells. **p**, Differential ribosome occupancy on mRNA transcripts in HelaS3^{45CKO} compared to HelaS3 cells. Log2 fold-change in read counts derived from analysis of RPF libraries (Log2FC RPF) and corresponding -Log10 of Benjamini-Hochberg P^{adj} values are plotted. Transcripts changing significantly between conditions are colored (purple). 'n' denotes total number of transcripts analyzed. The number of transcripts showing reduced (DN) or increased (UP) translation is also shown. **q**, As (o), for mRNA transcripts displaying altered expression level (Log2FC mRNA), as measured by analysis of read counts mapped to mRNA transcripts derived from total RNA libraries. Those changing significantly between conditions are colored (cyan). **r**, Gene ontology analysis of mRNA transcripts displaying elevated expression in HelaS3^{45CKO} cells. GO categories (FDR < 0.05) for biological process are labelled. Number of genes overlapping with each biological process GO category is indicated by the color scale gradient (count). No GO terms were found to be overrepresented for transcripts with reduced expression levels (at FDR < 0.05).
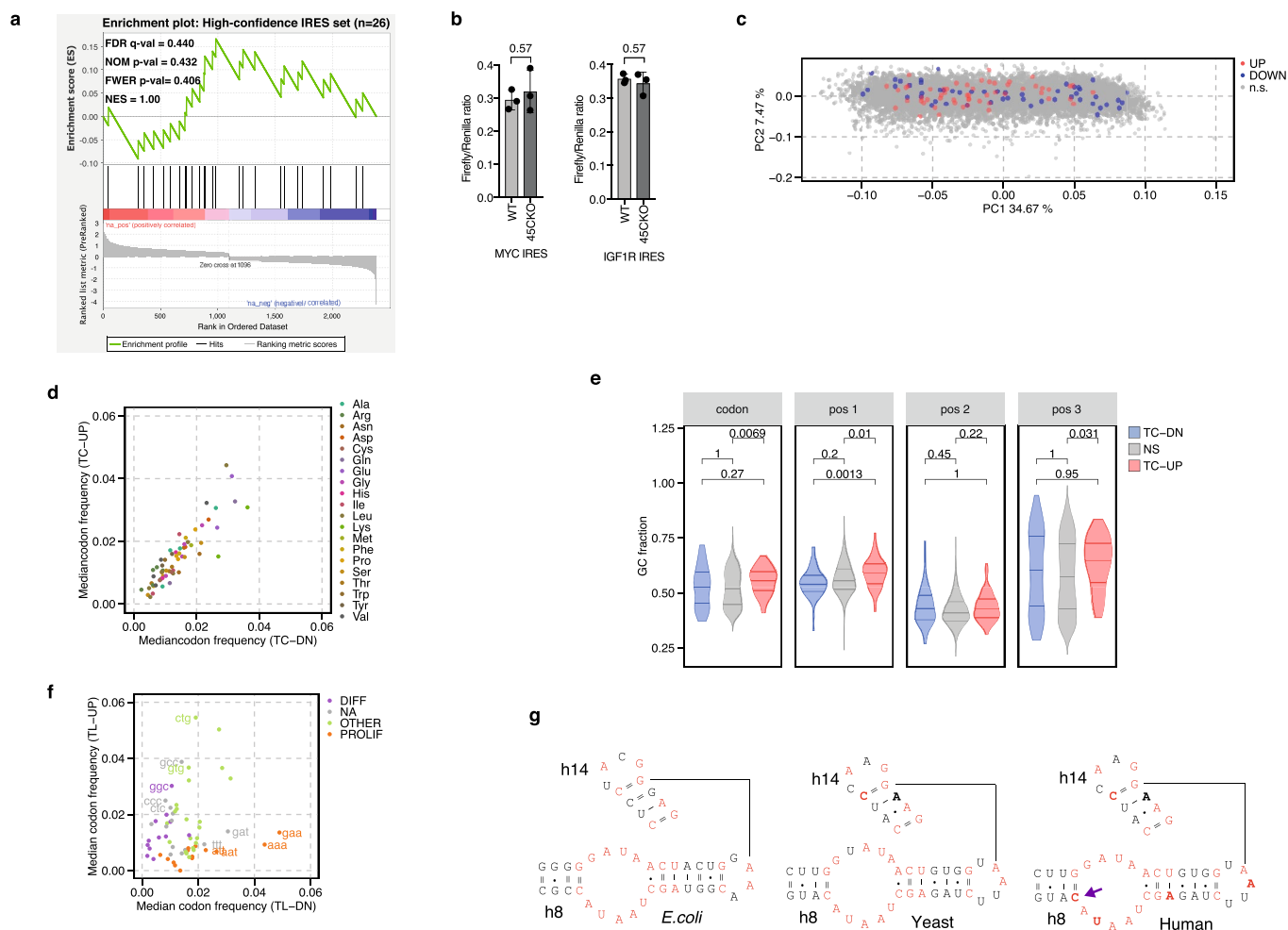
Extended Data Fig. 4 | mRNA codon composition determines translational regulation by Cm174 methylated ribosomes. a, Gene set enrichment analysis comparing genes that exhibit Cm174-dependent translational regulation to a high-confidence set harboring IRES elements. Number of genes in both sets is given by 'n'. Values obtained for FDR q-value, nominal (NOM) p-value, FWER p-value and normalized enrichment score (NES) are indicated. **b**, Luciferase reporter assays for IRES versus 5'cap-mediated translation initiation in HelaS3 (WT) compared to HelaS3[45CKO] (KO) cells. MYC and IGFR1 IRES elements were tested. Columns indicate means for each condition ($n = 3$ independent experiments), points (black) denote each value separately. P values (Welch's t-test, two-tailed) are shown, indicated by brackets. **c**, Principal component analysis of codon usage frequency in transcripts exhibiting increased (TC-UP, red), decreased (TC-DN, blue), and not-significantly altered (NS, grey) expression in HelaS3[45CKO] cells. **d**, Usage frequencies for each codon in mRNA transcripts exhibiting increased (TC-UP), decreased (TC-DN) expression in HelaS3[45CKO] cells. Median frequencies for all transcripts are plotted in each case. Codons are color coded by amino acid as indicated. Codons changing with Log2FC > 1, minimum frequency of 0.02 and P_{adj} value < 0.001 (Bonferroni corrected Wilcoxon tests, unpaired, two-sided) are annotated by codon sequence. **e**, Fraction of G or C nucleotides present in codons comprising mRNA transcripts in each mRNA expression regulation category, down-regulated (TC-DN, blue), not-significantly regulated (NS, grey) or Up-regulated (TC-UP, red). Comparisons for entire codons (codon) or first, second and third codon positions (pos 1-3) are shown separately. P_{adj} values for comparisons are shown above brackets. **f**, Codon usage frequencies in mRNA transcripts exhibiting increased (TL-UP) or decreased (TL-DN) translation in HelaS3[45CKO] cells. Median frequency of each codon is plotted in each case. Colors denote groups of codons enriched in mRNAs associated with indicated meta-GO categories defined in a previous study: DIFF (differentiation), PROLIF (proliferation), OTHER (not related to proliferation or differentiation), and NA (codons without a cognate anti-codon match). **g**, Conservation of h8 and h14. Schematic showing sequence and base-pairing interactions of h8 and h14 in E.coli, human and yeast 18S rRNA. Conserved residues (red) and 2'-O-methylated residues (bold font) are indicated. Position of Cm174 in human is highlighted by an arrow.

Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | Cm174 ribosome methylation influences cell proliferation and metabolism. a, Cell proliferation in HelaS3 (WT, black) and HelaS3^{45CKO} (KO, red) cells, measured by crystal violet assay at specified times. Points represent mean of $n=3$ independent experiments. Error bars represent $\pm$ s.d. P values (Welch's t-test, two-tailed), indicated by brackets are shown comparisons at days 3 and 4. **b**, Cell cycle phase analysis in HelaS3 (WT) and HelaS3^{45CKO} (45CKO) cells. Incorporation of EdU (+EdU), as determined by FACS. Cell seeding was optimized to ensure approximately equal densities at harvest, in order to reduce the effect of differing proliferation rates on cell cycle phase duration. Percentage of cells in gates representing cell cycle phases G1, S, G2/M is shown. Bottom panels show control samples not pulsed with EdU (-EdU). Data shown is representative of $n=3$ independent experiments. **c**, Quantification of cell cycle phase analysis, as described in (b). Columns represent mean percentage cells in each phase from $n=3$ independent experiments. Error bars represent $\pm$SEM. P values (two-way ANOVA with Sidak correction for multiple comparisons) are indicated by brackets. **d**, RiboMeth-seq (RMS) scores representing fraction of 2'-O-methylation at each site in 18S (left) and 28S rRNAs (right) HelaS3$^{GFP-45CKO}$ control cell line, where SNORD45C is not present within the EGFP intron. Nucleotide positions in respective rRNAs are indicated. Points denote mean RMS score ($n=3$ libraries from individual cultures) per site. Error bars represent $\pm$ s.d. **e**, As (d), for HelaS3$^{GFP-45CKO+45C}$ cell line, where SNORD45C is inserted inside the EGFP intron. **f**, Comparison of RiboMeth-seq (RMS) scores in 28S rRNA from HelaS3$^{GFP-45CKO}$ and HelaS3$^{GFP-45CKO+45C}$ cells. Points denote mean RMS score ($n=3$ libraries from individual cultures) per site. Horizontal and vertical error bars indicate $\pm$s.d. for each condition. No sites exhibit significant change. **g**, Expression of SNORD45C, SNORD45A and host gene *RABGGTB* in HelaS3^{GFP} (WT), HelaS3$^{GFP-45CKO}$ (45KO) and HelaS3$^{GFP-45CKO+45C}$ (45KO + 45C) cells, as determined by RT-qPCR analysis. Columns indicate mean ($n=3$ independant experiments) fold-change relative to the control condition, normalized to invariant SNORD46 expression. Error bars indicate 95% confidence intervals. P values (Welch's t-test, two-tailed) are shown for pairwise comparisons, indicated by brackets. **h**, Oxygen consumption rates (OCR) in HelaS3 (WT) compared to HelaS3^{45CKO} (45CKO) cells, measured by Seahorse XFe96 analyzer mito-stress test assay. Trace displays OCR over time, at basal levels and following injection of with compounds modulating cellular respiration (oligomycin, FCCP and a combination of rotenone and antimycin A) at the indicated time points. Representative data from $n=4$ independent experiments is shown. Error bars represent $\pm$ SEM. **i**, Respiration and key parameters of mitochondrial function in HelaS3 (WT) compared to HelaS3^{45CKO} (45CKO) cells. Quantification of OCR as described in (A) measuring basal respiration, ATP synthesis, proton leak and uncoupled respiration. Columns indicate mean OCR for each condition ($n=4$ independent experiments), points (black) denote each value separately. Error bars represent $\pm$ SEM. P values (Welch's t-test, two-tailed) for comparisons are shown, indicated by brackets. **j**, As (h), for HelaS3^{GFP} (WT-GFP), HelaS3$^{45CKO-GFP}$ (45CKO-GFP) and HelaS3$^{45CKO-GFP+45C}$ (45CKO-GFP+45C) cells. Representative data from $n=5$ independent experiments is shown.
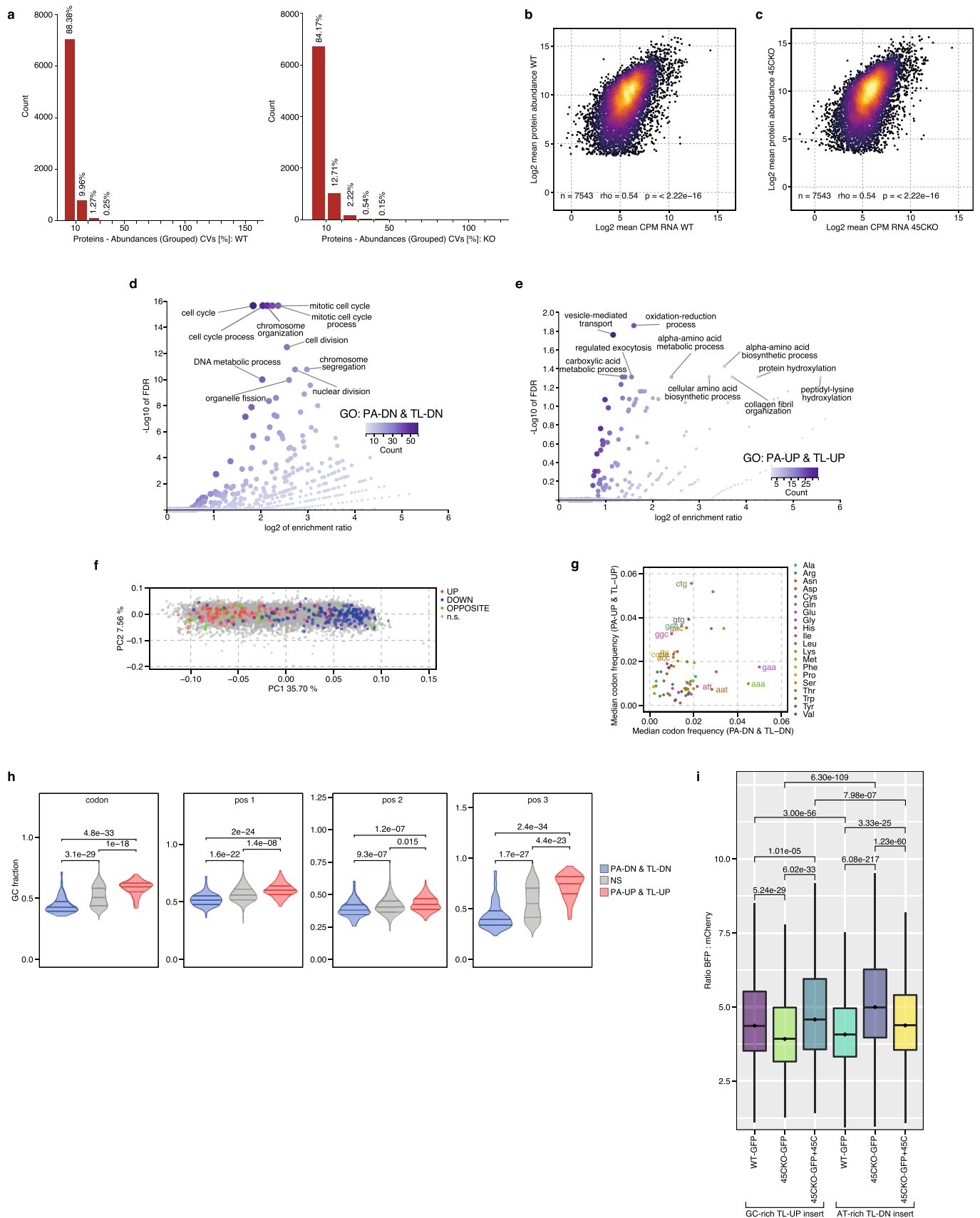
Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | Ribosome regulation by Cm174 alters the levels of proteins involved in distinct cellular processes and affects translation of codons with differing GC content. a, Overview of protein-level quantitation values between replicate mass spectrometry analyses. Coefficient of variation (CV) is below 20% for >96% of protein quantitation values in for both HelaS3 (WT, left) and HelaS3[45CKO] (KO, right) samples ($n=4$ per condition), indicating high reproducibility between replicates ($n=4$ independent protein samples per condition). **b**, Correlation between mRNA translation and protein abundance in HelaS3 (WT) cells. Points represent mean total RNA (RNA) read count per million mapped reads (CPM) for each transcript ($n=3$ libraries from independent cultures per condition, from ribosome profiling) and corresponding mean protein abundance for the same gene ($n=4$ independent protein samples per condition, from MS analysis). Color scale indicates plotting density. 'n' denotes number of genes/proteins. Spearman's *rho* and associated *P* value (algorithm AS 89) are shown. **c**, as (b), for HelaS3[45CKO] (45CKO) cells. **d**, Gene ontology analysis of overlap between PA-DN and TL-DN sets. Top 10 GO categories for biological process are labelled. Number of genes overlapping with each GO category is indicated by the color scale gradient (count). **e**, As (d), for overlap between PA-UP and TL-UP sets. **f**, Principal component analysis of codon usage frequency in transcripts exhibiting significant Log2FC changes in both proteomic and ribosome profiling analyses of HeLaS3 and HelaS3[45CKO] cells. Increased (PA-UP & TL-UP, red), decreased (PA-DN & TL-DN, blue), not-significantly altered (NS, grey), and opposite (magenta) change categories are indicated. **g**, Usage frequencies for each codon in transcripts included in PA-UP & TL-UP versus PA-DN & TL-DN sets. Median frequencies are plotted in each case. Codons are color coded by amino acid as indicated. Codons changing with Log2FC >1, minimum frequency of 0.02 and *P[adj]* value < 0.001 (Bonferroni corrected Wilcoxon tests, unpaired, two-sided) are annotated by codon sequence. **h**, Fraction of G or C nucleotides present in codons comprising mRNA transcripts in included in PA-DN & TL-DN (blue), PA-UP & TL-UP (red), or not-significantly altered (NS, grey) sets. Comparisons for entire codons (codon) or first, second and third codon positions (pos 1-3) are shown separately. *P[adj]* values (Bonferroni corrected Wilcoxon tests, unpaired, two-sided) for comparisons are given above brackets. **i**, High content imaging assay of codon reporter cell lines (replicate 1), containing either AT or GC rich codon inserts as indicated. mTagBFP2:mCherry ratio, after expression of the reporter cassette was induced for 24 h. The center of the box represents the median; the top and the bottom edges of the box represent the third and first quartiles, respectively. Whiskers represent minima and maxima, excluding outliers beyond 1.5× IQR. Bonferroni *P[adj]* values derived from Wilcoxon tests, two-tailed for each pairwise comparison indicated by brackets are given. See Supplementary Table 7 for all possible statistical comparisons. The number of double-positive mTagBFP2 and mCherry cells assayed in each condition, from left to right: $n=7667, 1828, 1839, 9081, 3545, 2914$.

nature portfolio

Corresponding author(s):   Martin D. Jansson
                           Anders H. Lund

Last updated by author(s):   Aug 19, 2021

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used
Data analysis	Open source software: R v3.6.1, Illumina bcl2fastq v2.20.0.422, fastqc v0.11.7, cutadapt v1.18, bowtie2 v2.2.9, STAR v2.5.1a, Ribotaper 1.3.1, FeatureCounts v1.5.1,Samtools 1.9,extract_reads.py, deepTools v3.3.1, wiggleTools v1.2.3, wigToBigWig v.4, RIVET, WebGestalt 2019, Gene Set Enrichment Analysis (GSEA) v4.0.3 (Broad Institute), SEQUEST (University of Washington), ImageJ v1.51j (NIH). Note: parameters are provided in Methods, or in public database submissions, where appropriate. RiboMeth-seq data analysed with custom code (https://github.com/lundlab/RiboMeth-seq-public). Commercial software: StepOne Software v2.0 (Applied Biosystems), FlowJo v10 (FlowJo, LLC), BD CellQuest Pro (BD Biosciences), Prism 7.0d (Graphpad), Proteome Discoverer 2.4 (ThermoFisher), Photoshop (Adobe), Excel (Microsoft), PyMol Molecular Graphics System, v2.4.1 (Schrödinger, LLC), InCell Analyzer 1000 (GE Healthcare).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

All RNA sequencing data from this study have been deposited in the Gene Expression Omnibus (GEO) under series accession numbers: GSE153476 (for RNA-seq

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Samples sizes are stated in Figure legends or Methods. No statistical methods were used to predetermine sample size. Samples sizes for each experiment are typical of those commonly used in the field. No previous study was used to determine the number of samples.
Data exclusions	No data was excluded.
Replication	All experiments were typically reproduced from three or more biological replicates, with a minimum of two where stated. Number of replicates is given in the figure legends and/or described in Methods. All RNA-seq experiments were performed from triplicate samples for each condition. Proteomic experiments were performed from quadruplicate samples for each condition.
Randomization	There was no randomization in the allocation of samples because proper controls were used. All samples (controls and treatments) were treated in the exact same manner for all the experiments.
Blinding	Researchers were not blinded during experiments and data analysis because proper controls were used.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	MYC D84C12 Rabbit mAb #5605 (Cell Signaling Technology) 1:1000, FBL sc-25397 1:1000; p53 sc-126 1:2000; MDM2 sc-56154 1:1000; p16 CDKN2A sc-1661 1:1000; GAPDH sc-25778 (Santa Cruz Biotechnology) 1:2000; p21 CDKN1A 556431 (BD pharmingen)1:1000; Vinculin V9131 (Sigma) 1:200000.
Validation	All antibodies were commercially purchased and validated for their respective application by the manufacturers, additionally all of the antibodies used for western blot have been used by the Lund lab previously except MYC D84C12 which is validated by Cell Signaling Technology for several human cell lines.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	HeLaS3 (ATCC CCL-2.2), DLD1 (ATCC CCL-221), HCT116 (ATCC CCL-247) and HCT116-p53-/- were previously obtained from the laboratory of Bert Vogelstein, BJ-hTERT were obtained from the laboratory of Kristian Helin, HEK293 (ATCC CRL-1573)
Authentication	None of the cell lines used were authenticated.

Mycoplasma contamination	Cell lines were routinely tested for mycoplasma contamination, no contamination was detected.
Commonly misidentified lines (See ICLAC register)	No commonly mis-identified cell lines were used according to ICLAC register.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	For each sample cells were seeded in 6cm dishes in triplicate at densities resulting in equal confluence (approximately 70%) at harvest. 48h later, cells were pulsed with 30µM EdU for 40min and then harvested by trypsinisation and washed 3x in PBS with 1% BSA. EdU detection was performed using Click-iT EdU Alexa Fluor 647 Flow Cytometry Assay Kit (ThermoFisher) according to manufacturer's instructions.
Instrument	BD FACSCalibur
Software	BD CellQuestTM Pro, FlowJo 10 (FlowJo, LLC)
Cell population abundance	At least 10,000 counts were recorded.
Gating strategy	Cells were gated on FSC/SSC to remove doublets, and subsequently relative to EdU647 and PI staining. Gating used was identical for each condition. Gating strategy for cell cycle phases is illustrated in Extended Data Fig. 5b.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.