Loss of $m^1acp^3\Psi$ ribosomal RNA modification is a major feature of cancer

Artem Babaian[1,2]*, Katharina Rothe[1,2], Dylan Girodat[3], Igor Minia[4], Sara Djondovic[1], Miha Milek[4], Sandra E. Spencer Miko[5], Hans-Joachim Wieden[3], Markus Landthaler[4,6], Gregg Morin[1,5], and Dixie L. Mager[1,2].

1. Department of Medical Genetics, University of British Columbia, Vancouver, BC, Canada

2. Terry Fox Laboratory, BC Cancer, Vancouver, BC, Canada

3. Alberta RNA Research and Training Institute, University of Lethbridge, Lethbridge, AB, Canada

4. Max Delbrück Center for Molecular Medicine Berlin in the Helmholtz Association, Berlin Institute for Medical Systems Biology, Berlin, Germany

5. Michael Smith Genome Sciences Centre, British Columbia Cancer Agency, Vancouver, BC, Canada

6. IRI Life Sciences, Institut für Biologie, Humbold Universität, Berlin, Germany

Characters: 20,567

***** Corresponding author

Dr. Artem Babaian, Ph.D.
Terry Fox Laboratory
BC Cancer
675 West 10th Avenue
Vancouver, BC, V5Z 1L3, Canada
Email: ababaian@bccrc.ca

Summary

The ribosome is an RNA-protein complex essential for translation in all domains of life. The structural and catalytic core of the ribosome is its ribosomal RNA (rRNA). While mutations in ribosomal protein (RP) genes are known drivers of oncogenesis, oncogenic rRNA variants have remained elusive. We discovered a cancer-specific single nucleotide variation in 18S rRNA at nucleotide 1248.U in up to 45.9% of colorectal carcinoma (CRC) patients and present across >22 cancer types. This is the site of a unique hyper-modified base, 1-methyl-3-α-amino-α-carboxyl-propyl pseudouridine ($m^1acp^3\Psi$), a >1 billion years conserved RNA modification at the ribosome's peptidyl decoding-site. A sub-set of CRC tumors we term 'hypo-$m^1acp^3\Psi$', show sub-stoichiometric $m^1acp^3\Psi$-modification unlike normal control tissues. A $m^1acp^3\Psi$ knockout model and hypo-$m^1acp^3\Psi$ patient tumors share a translational signature, characterized by highly abundant ribosomal proteins. Thus, $m^1acp^3\Psi$-deficient rRNA forms an uncharacterized class of 'onco-ribosome' which may serve as a chemotherapeutic target for treating cancer patients.

Keywords

Ribosome, Ribosomal RNA variation, RNA modification, Colorectal Carcinoma, Ribosomal Heterogeneity, Onco-ribosome

Introduction

The ribosome is a massive ribonucleoprotein particle (RNP) responsible for the transformation of genetic information encoded as nucleic acids into functional proteins encoded as amino acids. Unlike most RNPs, it is ribosomal RNA (rRNA), and not ribosomal proteins (RPs) that form the most ancient and catalytic core of the complex (Cech, 2000). rRNA is further functionalized by a constellation of at least 14 distinct chemical modifications across 200+ sites, (Taoka et al., 2016) clustering around active sites of the ribosome (Sloan et al., 2017), yet the function of many rRNA modifications remains unclear.

The human ribosome contains >80 RPs and four rRNAs, totaling ~80% of cellular RNA. During the initial human genome sequencing project, ribosomal DNA (rDNA) loci were systematically excluded from the reference genome (Lander et al., 2001); given that a reference sequence of the rRNA gene, *RNA45S*, was available and the 80-800 rDNA copies were believed to be homogeneous (Elder and Turner, 1995). Although there was early evidence for rDNA polymorphism in humans (Kuo et al., 1996; Leffers and Andersen, 1993). Thus, as technology and sequencing consortium projects revolutionized genomics and transcriptomics, our understanding of rDNA variation has lagged.

rDNA sequence variation at the intra- and inter-individual level has been documented in multiple species including humans (Bik et al., 2013; Rabanal et al., 2017; Babaian, 2017; Kim et al., 2018; Parks et al., 2018), but the functional implications of rDNA variation remain elusive. Mutation of RP genes and ribosome biogenesis factors can cause a class of diseases termed ribosomopathies, including Diamond Blackfan anemia (DBA) (Nakhoul et al., 2014), and some cancers (Gourdazi and Lindstrom). It has been hypothesized that cancer cells contain a functionally specialized class of ribosomes to facilitate rapid protein synthesis, termed the “onco-ribosomes” (Shi and Barna, 2015; Dinman, 2016). Cancer genomics has supported this notion with the identification of several oncogenic driver mutations in RP genes (Gourdazi and Lindstrom, 2016), the best characterized of which are RPL10 (uL16) p.R98S in T-cell acute lymphoblastic leukemia (Girardi et al., 2018; Kampen et al., 2019) and RPS15 (uS19) C’ terminal mutations in chronic lymphocytic leukemia (Bretones et al., 2018). In addition, germline mutations such as in DBA patients and in *RPS20* can cause heredity cancers including colorectal carcinoma (CRC) (Vlachos et al., 2012; Nieminen et al., 2014).

As RP mutations have been implicated in tumorigenesis, we hypothesized that rRNA variation or mutation is a cancer driver. To map functional rRNA sequence variation, we considered tumorigenesis

as a natural experiment in which polymorphic and mutant rRNA alleles undergo selective evolutionary change in frequency within each patient. We discovered a surprising 18S rRNA single nucleotide alteration at the decoding core of the ribosomal peptidyl (P)-site, affecting up to 45.9% of CRC patients, making this the most frequent ribosomal variant associated with cancer to date and potentially revolutionizing future chemotherapeutic strategies against this disease.

Results & Discussion

An unexpected rRNA variant in cancer: sub-stoichiometric modification of 18S.1248.m^{1}acp^{3}Ψ

In an initial screen for cancer-driver rRNA variants, we aligned RNA-seq reads from 66 colorectal carcinoma (CRC) tumors and patient-matched adjacent normal tissue to a single-copy reference rDNA. To test for allelic selection inconsistent with neutral drift, the patient-matched difference in expressed variant allele frequency (VAF) was measured for deviation from zero for each position of 18S and 28S (Fig. 1a). Non-reference reads at a single nucleotide deviated from neutrality, 18S:r.1248.U (p_{adj} = 3.81e-8). In this cohort, the 18S.1248.U reference allele is recurrently selected over non-U or 18S.1248V alleles in a striking 44.9% of CRC patients (Fig. 1a); in comparison, oncogenic *KRAS*^{G12} codon mutation occurs in only 36% of CRC patients (Tan and Du, 2012).

Surprisingly, at the DNA level, the respective nucleotide *RNA45S*:4908.T is invariable in humans and CRC (Babaian, 2017 Fig. S1a,b), and in the mature rRNA this uridine undergoes hyper-modification to 1-methyl-3-α-amino-α-carboxyl-propyl pseudouridine (m^{1}acp^{3}Ψ) (Fig. 1b) (Taoka et al., 2016). The m^{1}acp^{3}Ψ modification perturbs standard Watson-Crick base-pairing during cDNA synthesis by reverse transcriptase (RT) (Helm and Motorin, 2017), resulting in base misincorporation and enzyme stalling, which is read-out as a consistent 'modification signature' in RNA-seq (Fig. 1c, reviewed in (Helm and Motorin, 2017)). Restated, non-reference or variant reads at 18S.1248.U correspond to m^{1}acp^{3}Ψ modification and not a genetic variation. Thus, the increase in reference U sequence suggests that the m^{1}acp^{3}Ψ modification is incomplete or sub-stoichiometric in CRC tumors, which we term the 'hypo-m^{1}acp^{3}Ψ phenotype'. The 28S.1321.m^{1}A and 28S.4532.m^{3}U rRNA modifications also cause a 'modification signature' in RNA-seq (Helm and Motorin, 2017). These modifications do not decrease in CRC tumors or matched normal controls, excluding a non-specific rRNA modification effect (Fig. 1d).

Figure 1

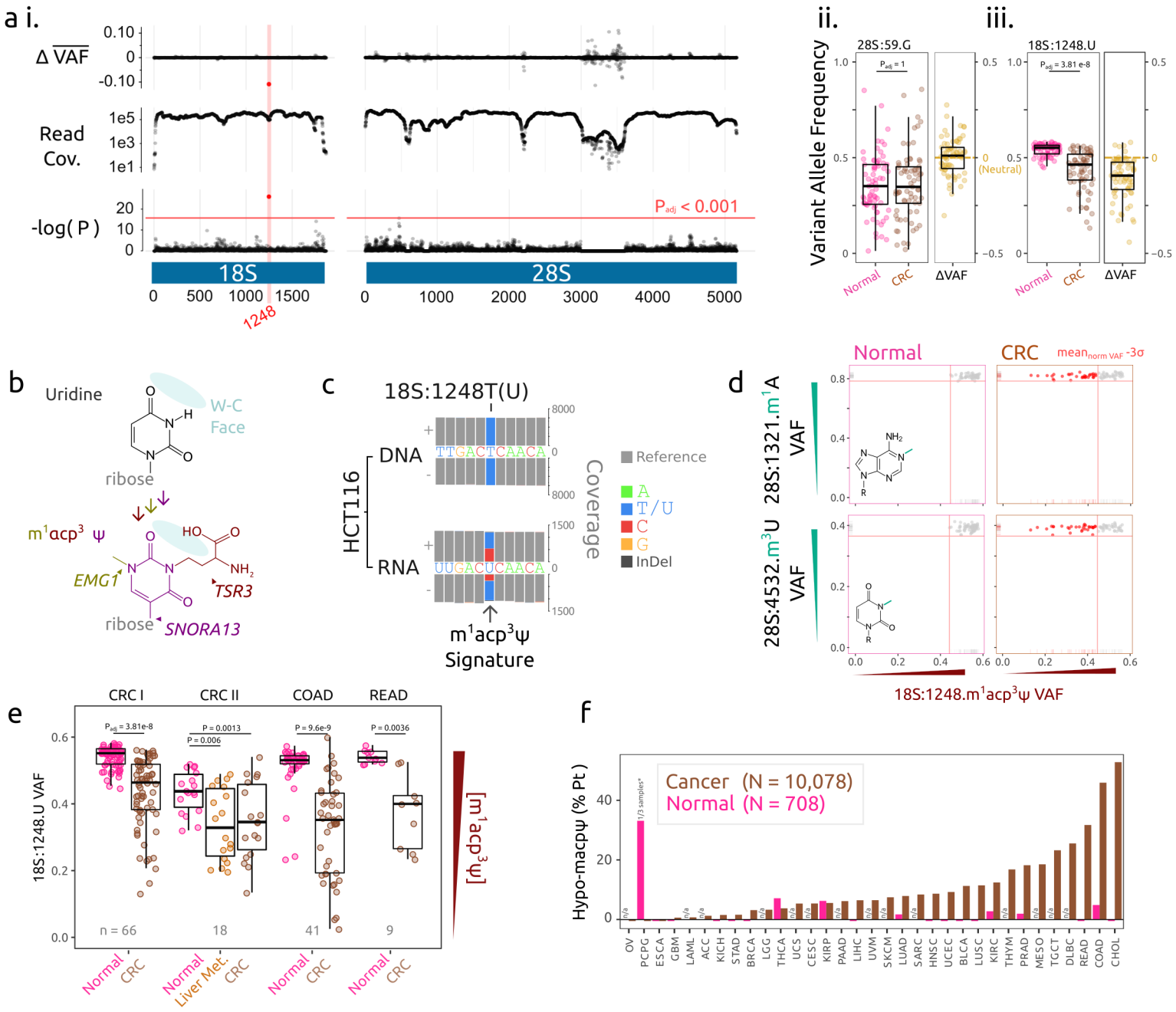

a i.
Δ VAF
Read Cov.
-log(P)
Padj < 0.001
18S
28S
1248
ii.
iii.
28S:59.G
18S:1248.U
Variant Allele Frequency
Normal
CRC
ΔVAF
(Neutral)
b
Uridine
W-C Face
ribose
m1acp3 ψ
EMG1
TSR3
SNORA13
c
18S:1248T(U)
HCT116
DNA
RNA
TTGACTCAACA
UUGACUCAACA
Coverage
Reference
A
T/U
C
G
InDel
m1acp3ψ Signature
d
Normal
CRC
meannorm VAF -3σ
28S:1321.m1A VAF
28S:4532.m3U VAF
18S:1248.m1acp3ψ VAF
e
CRC I
CRC II
COAD
READ
18S:1248.U VAF
n = 66
18
41
9
Normal
Liver Met.
CRC
[m1acp3ψ]
f
Hypo-macpψ (% Pt)
Cancer (N = 10,078)
Normal (N = 708)
OV
PCPG
ESCA
GBM
LAML
ACC
KICH
STAD
BRCA
LGG
THCA
UCS
CESC
KIRP
PAAD
LIHC
UVM
SKCM
LUAD
SARC
HNSC
UCEC
BLCA
LUSC
KIRC
THYM
PRAD
MESO
TGCT
DLBC
READ
COAD
CHOL

Figure 1: The hypo-m^{1}acp^{3}Ψ phenotype in cancer

a i, Screen for change in the average variant allele frequency (VAF) across *18S* and *28S* ribosomal RNA (rRNA) in colorectal cancer (CRC) RNA-seq compared to patient-matched normal epithelium controls (n = 69). Read coverage and quality drops at extreme GC-content (>90%) regions of 28S, these low-coverage regions were excluded from further analysis. **ii,** The common human rRNA polymorphism 28S:r.59G>A ranges from 0.05–0.93 DNA allele frequency (Babaian, 2017) and was expressed comparably in the normal epithelium between variant allele frequencies (VAF) of 0.01-0.86. Neither allele is directionally selected for during cancer evolution, consistent with neutral drift (p_{adj} = 1, t = -0.44). **iii,** 18S:r.1248.U is significantly enriched (p_{adj} = 3.81e-8, t = 8.33) for the reference U allele. **b,** The 18S:r.1248.U base normally undergoes enzymatic hyper-modification to 1-methyl-3-α-amino-α-carboxyl-propyl pseudouridine (m^{1}acp^{3}Ψ) in three steps; *SNORA13* guided pseudouridylation; EMG1 N1-methylation; and finally 3-amino-carboxyl-propylation by TSR3, which are not down-regulated in hypo-m^{1}acp^{3}Ψ tumors (Fig. S1d). **c,** Perturbation of the Watson-Crick face of the modified base results in a distinct nucleotide misincorporation signature by reverse transcriptase in first strand cDNA synthesis, which is read out on both the first (+) and second read-pair (-) of sequencing of matched HCT116 whole genome DNA- and RNA-seq (Cancer Cell Line Encyclopedia, (Ghandi et al., 2019)). **d,** Patient 18S:r.1248.m^{1}acp^{3}Ψ hypo-modification is defined as a decrease in VAF by three standard deviations (3σ) of the matched-normal samples. Hypo-m^{1}acp^{3}Ψ is not correlated with the loss of other rRNA modifications detectable by RNA-seq. **e,** The hypo-m^{1}acp^{3}Ψ phenotype is replicated in three additional, independent cohorts of CRC with patient-matched adjacent normal controls, including two cohorts from The Cancer Genome Atlas (TCGA), colorectal adenocarcinoma (COAD) and rectal adenocarcinoma (READ). **f,** Hypo-modification of 18S:r.1248.m^{1}acp^{3}Ψ is prevalent but not ubiquitous across the TCGA cancer cohorts (n = 10,078 patients) and largely absent from patient-matched normal controls (n = 708) (see: Fig. S1).

The hypo-m^{1}acp^{3}Ψ phenotype is reproducible at comparable frequency (27.8-45.9%) in three additional independent patient-matched CRC cohorts (Fig. 1e,f). Analysis of 10,036 cancer patients and 712 normal controls across an additional 31 cancer-patient cohorts reveals that hypo-m^{1}acp^{3}Ψ occurs at a significant frequency across a diverse set of cancers but is not pan-cancerous (global recurrence: 9.6% range: 0-52.8%) (Fig. 1f, S1).

To validate these findings, we designed a simple and rapid aminocarboxyl propyl reverse transcription (aRT)-PCR assay for measuring 18S.1248.$m^1acp^3\Psi$ modification. The aRT-PCR is reproducible and quantitative (Fig. S2). A critical technical limitation which also applies to RNA-seq is that different RT enzymes or RT reaction chemistries, differ the base misincorporation rates at 18S.1248.$m^1acp^3\Psi$ (Fig. S2d). Thus, cross-cohort or cross-experimental comparisons should be made cautiously (Fig. S2b-d). Indeed, batch-effects on VAF are seen within TCGA cohorts, yet hypo-$m^1acp^3\Psi$ replicates across batches further supporting that hypo-$m^1acp^3\Psi$ is occurring in tumors specifically (Fig. S1b).

We validated that the hypo-$m^1acp^3\Psi$ phenomenon also occurs in CRC cell lines assayed as a single batch and confirmed the results are not a sequencing artifact (Fig. S2e). To test if $m^1acp^3\Psi$-deficient rRNA incorporate into mature ribosomes, we isolated monosomes and polysomes and detected low $m^1acp^3\Psi$ modification levels in mono- and di-somes (Fig. S2f).

As the molecular, biological, and medical significance of 18S.1248.$m^1acp^3\Psi$ unfolds, it is obvious that genotyping technologies (such as sequencing or our aRT-PCR assay for $m^1acp^3\Psi$) and previous $m^1acp^3\Psi$ assays such as primer extension can be adapted as affordable and rapid diagnostic or prognostic assays.

18S:1248.$m^1acp^3\Psi$ is an ancient modification at the decoding core of the peptidyl-site

We next investigated the evolutionary and structural characteristics of 18S.1248.$m^1acp^3\Psi$ for functional insight. The 18S:1248.U base and $m^1acp^3\Psi$ modification are absolutely conserved across *Eukarya* at a residue located in the loop of universal helix 31 (Fig. S3). *TSR3* is the aminocarboxylpropyl transferase which deposits the acp_3 at 18S.1248.U (Meyer et al., 2016) and it only modifies this single rRNA position, in 100% of mature rRNA molecules in *Eukaryotes* (Taoka et al., 2018; Yang et al., 2016).

Structurally, 18S.1248.$m^1acp^3\Psi$ is solvent-exposed at the ribosomal P-site, immediately adjacent to the codon:anti-codon interface (Fig. 2). Cryo-EM structures (Li et al., 2019; Natchiar et al., 2017) and our molecular dynamics (MD) simulations implicate the $m_1acp_3\Psi$-modification in a direct interaction with P-site tRNA. The modification carboxyl-moiety forms a hydrogen bond with the universally conserved RPS16 p.R146 (Jindal et al., 2019), an interaction aided by the m^1 modification which stabilizes base rotation angle (Fig. S4c). Unlike *TSR3*, the m^1 transferase of 18S.1248, *EMG1*, is essential and its mutation causes Bowen-Conradi Syndrome (OMIM: 211180, Armistead et al., 2009; Wurm et al., 2010), supporting that 18S.1248 modification plays a crucial role in P-site function.

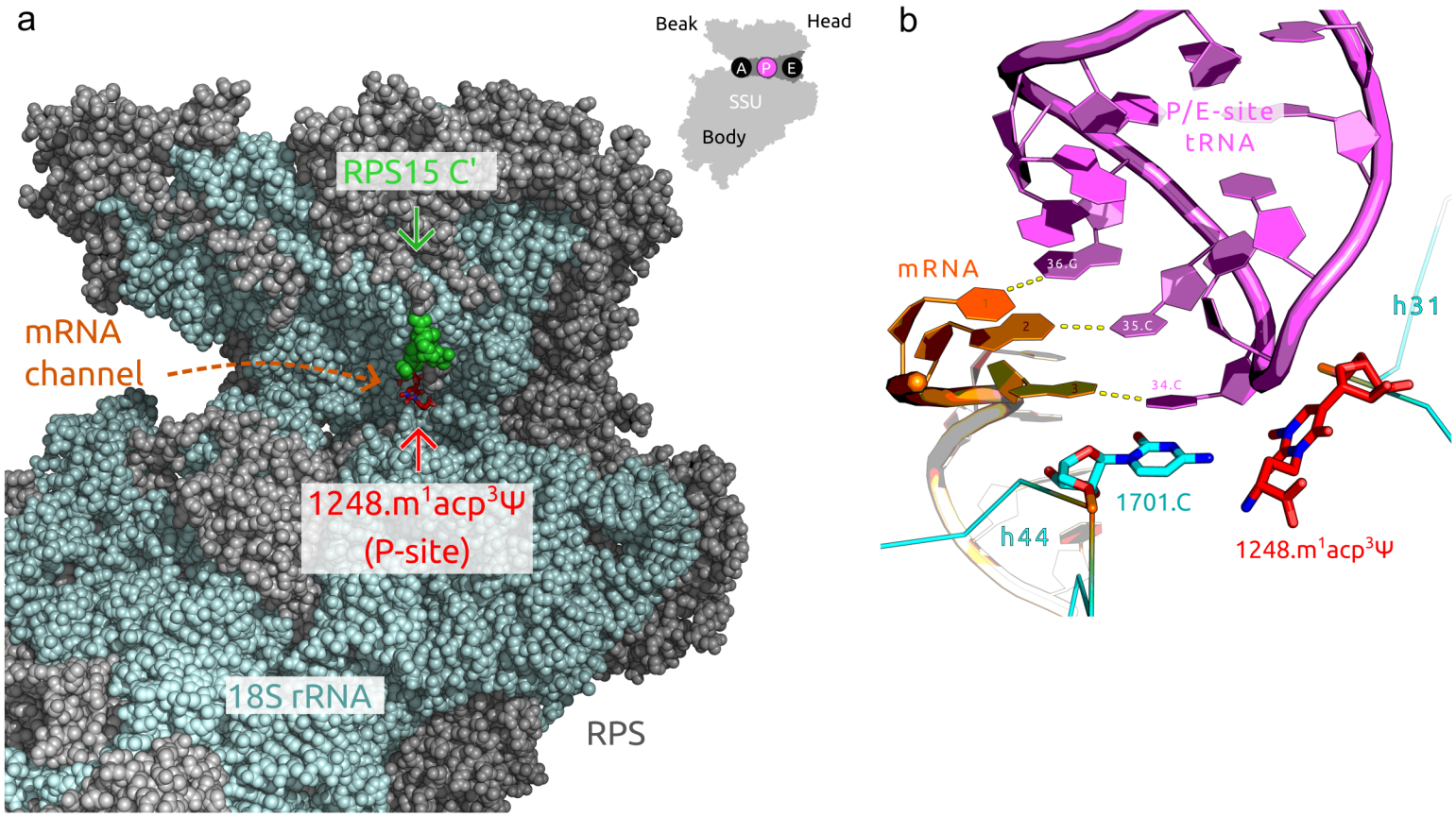
Start AUG-codon selection and translational initiation is a rate-limiting step in protein-synthesis and both occur at the P-site. Thus, the hypo-$m^1acp^3\Psi$ phenotype may demarcate a class of 'onco-ribosome' with deregulated translation. It is noteworthy that the two largest effect size RP cancer driver mutations also occur at the ribosomal P-site/tRNA interface, the RPL10 p.R98S at the peptidyl transfer site (Girardi et al., 2018; Kampen et al., 2019) and RPS15 C` tail mutations adjacent (<12Å) to 1248.$m^1acp^3\Psi$ (Fig. 2a) suggesting that the ribosomal P-site is a convergent multi-cancer oncogenic hot-spot.

Since the discovery of streptomycin in 1944, the ribosome has been the target of several important classes of drugs (McCoy et al., 2011). The pervasive and recurrent loss of a solvent-exposed and charged acp_3 modification at the decoding core of the small sub-unit raises the possibility that this pocket may be therapeutically exploited with ribosome targeting antibiotics or their derivatives as a new generation of chemotherapies.

Figure 2: 18S:1248.$m^1acp^3\Psi$ is located at the peptidyl decoding site

(next page) **a,** The mRNA channel of the human small sub-unit (SSU) cryo-EM structure with resolved base modifications (PDB: 6EK0 (Natchiar et al., 2017)). The 18S:1248.$m^1acp^3\Psi$ (red) nucleotide is on the loop of the universal helix 31, exposed to the mRNA channel at the center of the P-site. The CLL driver mutations in the RPS15 C` tail (green) are <12.8Å from 18S:1248.$m^1acp^3\Psi$. The minimal distance is likely shorter as the 10 C` terminal residues of RPS15 which extend into the P-site are labile and not modeled. **b,** Cryo-EM structure with a P/E-site tRNA, 18S.1248.Ψ and 18S.1701.C base stack with the ribose and base of the tRNA.34.C, respectively (PDB: 6OLE (Li et al., 2019)). 18S.1248.$m^1acp^3\Psi$ modification contributes to P-site decoding site stability via interaction with P-site tRNA and RPS16 (Fig. S4).

Figure 2


a
Beak
Head
A
P
E
SSU
Body
RPS15 C'
mRNA
channel
1248.m1acp3Ψ
(P-site)
18S rRNA
RPS
b
P/E-site
tRNA
mRNA
36.G
35.C
34.C
h31
1701.C
h44
1248.m1acp3Ψ

Loss of 18S.1248.m^{1}acp^{3}Ψ modification induces ribosomal protein mRNA translation

To delineate the function of 18S.1248.m^{1}acp^{3}Ψ, we generated *TSR3* knockout CRC cell lines (HCT116). Similar to yeast (Meyer et al., 2016), *TSR3* is non-essential and we isolated two *TSR3* homozygous knockouts (*TSR3*[KO 1,3]), a heterozygous knockout (*TSR3*[Het 2]), as well as three wildtype control clones (WT[1-3]). Knockouts were functionally confirmed by five independent m^{1}acp^{3}Ψ assays and editing sites validated by RNA-seq, with *TSR3*[Het 2] showing an intermediate or hypo-m^{1}acp^{3}Ψ phenotype (Fig. 3a, S2g, S5). The hypo-m^{1}acp^{3}Ψ seen in *TSR3*[Het 2] is more severe than seen in CRC cell lines (Fig. S2e), but comparable to what is observed in CRC patient tumors (Fig. S1a). SCARLET shows the presence of the Ψ and m^{1}Ψ precursor bases at 18S.1248 in *TSR3*[KO 1] cells (Fig. S2h). Together this demonstrates that loss of the acp$_3$ modification via *TSR3*[KO] is necessary for nucleotide misincorporation during RT as measured by RNA-seq, supporting hypo-m^{1}acp^{3}Ψ tumors contain sub-stoichiometric loss of the acp$_3$-moiety.

Morphologically, the HCT116 clones were indistinguishable and showed comparable, rapid growth, cell cycle timing, and loss of TSR3 did not alter global protein translation rates (Fig. 3b-d). To determine how the loss of 18S.1248.m^{1}acp^{3}Ψ modification alters translation we characterized the transcriptome (RNA-seq) and translatome (ribo-seq) of these cells and validated our findings by global mass spectrometry (Fig. S5-S7).

Transcriptionally, gene set enrichment analysis (GSEA) revealed that *TSR3*[KO]/[Het] (vs. WT) cells were dominated by a proliferative tumor expression signature, characterized by Rb-depletion/ elevated-E2F transcription factor activity (Fig. 3e, Fig. S8, table S3). Yet, *TSR3*[KO]/[Het] cells also have a paradoxical decrease in translation of the same E2F target genes (Fig. S8), stressing the specificity of using RNA-expression gene-set classifiers.

To determine how loss of 18S.1248.m^{1}acp^{3}Ψ modification alters translation, we contrasted translational efficiency between genotypes. *TSR3*[KO]/[Het] cells showed a remarkable enrichment (vs. WT) in the translation of RPs, such as RPS8 and RPL4, with an associated with a depletion of RP mRNA (Fig. 3f, S7d-g, S9), yet overall translation rates and polysome profiles of *TSR3*[KO]/[Het] cells was not altered (Fig. 3d, S9e).

To validate if this RP mRNA/protein signature is present in cancer patients, we analyzed the CPTAC-CRC cohort with tumor matched RNA-seq and proteomics data (Vasaikar et al., 2019). Similar to

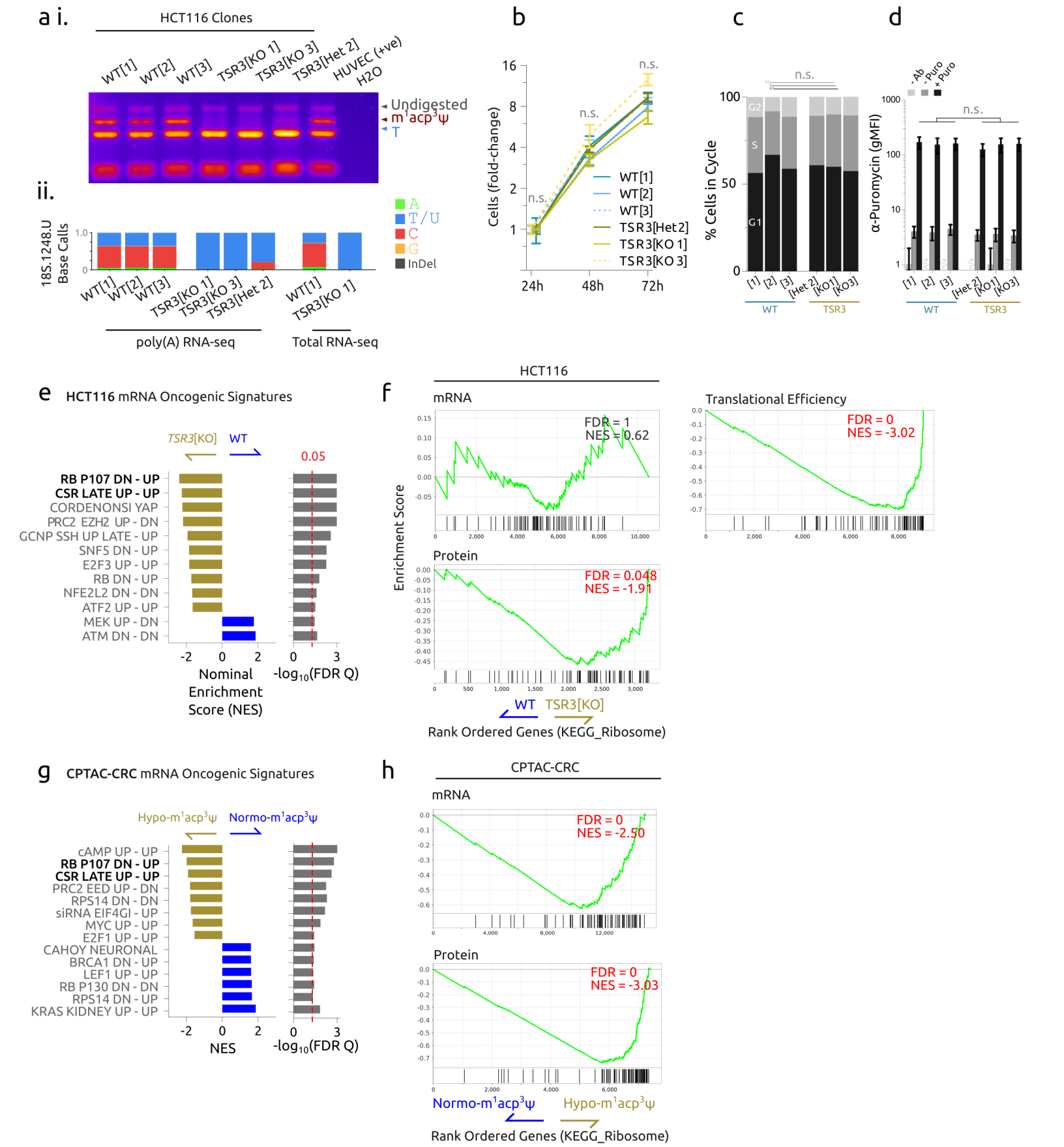
TSR3[KO]/[Het] cell lines, hypo-$m^1acp^3\Psi$ CRC tumors share the same E2F oncogenic gene signature and a proteomic increase in RPs relative to normo-$m^1acp^3\Psi$ CRC tumor controls (Fig. 3g,h).

There are two hypotheses with which to interpret the hypo-$m^1acp^3\Psi$ phenotype. The 'oncogenic driver hypothesis' is that $m^1acp^3\Psi$-deficient rRNA arise in tumorigenesis, and their dysregulated translation confers a selective advantage to the cancer, likely via high RP output. The recapitulation of the *TSR3*[KO]/[Het] multi-omic phenotype in hypo-$m^1acp^3\Psi$ CRC patients supports a causal model. Alternatively, $m^1acp^3\Psi$-deficiency arises in consequence to hyper-proliferation and high ribosomal biogenesis. Rapid cellular turn-over (oft associated with upregulated E2F1 activity) in turn results in 'incomplete modification' of rRNA. Under this model the consequences of $m^1acp^3\Psi$-deficient rRNA is near-neutral or tolerably detrimental to tumor fitness. Nevertheless, hypo-$m^1acp^3\Psi$ is a highly recurrent perturbation to the ancient peptidyl-decoding core and underlies a greater cancer-translational phenomenon.

Figure 3: The translational signature of $m^1acp^3\Psi$-deficient ribosomes

(next page) **a i,** Reverse transcription (RT)-PCR m^1acp^3-assay (see: methods and Fig. S2) and **ii,** RNA-seq measurement for nucleotide misincorporation at 18S:1248.$m^1acp^3\Psi$ in clones of the colorectal cancer HCT116 cell line with HUVEC as a normal positive control. **b,** The fold-change growth of HCT116 clone populations in culture, normalized to cells at 24 hours. **c**, Cell cycle timing analysis of log-growth HCT116 cell genotypes. **d,** Total protein translation rates measured by puromycin-incorporation and flow cytometric quantification. **e,** Summary of Gene Set Enrichment Analysis (GSEA) of RNA-seq comparing HCT116 WT[1-3] versus HCT116 *TSR3*[KO 1,3 / Het 2] clones. Only significant (False Discovery Rate Q-value < 0.05) gene sets in the Oncogenic Signature collection are shown. **f,** *TSR3*[KO/Het] cells GSEA shows no change in ribosomal protein (RP) mRNA abundance, while RP translational efficiency, and total protein increases. **g,** Summary of Oncogenic Signature GSEA comparing CPTAC colorectal carcinoma (CRC) tumors with normo-$m^1acp^3\Psi$ and hypo-$m^1acp^3\Psi$ modification. Gene sets common to HCT116 *TSR3*[KO]/[Het] and hypo-m^1acp^3 CPTAC-CRC patients are bolded and **h,** a similar increase in RP abundance is observed.

Figure 3


a i.
HCT116 Clones
WT[1]
WT[2]
WT[3]
TSR3[KO 1]
TSR3[KO 3]
TSR3[Het 2]
HUVEC (+ve)
H2O
Undigested
m1acp3ψ
T
ii.
18S.1248.U Base Calls
A
T/U
C
G
InDel
poly(A) RNA-seq
Total RNA-seq
b
Cells (fold-change)
n.s.
WT[1]
WT[2]
WT[3]
TSR3[Het2]
TSR3[KO 1]
TSR3[KO 3]
24h
48h
72h
c
% Cells in Cycle
G2
S
G1
WT
TSR3
d
α-Puromycin (gMFI)
- Ab
- Puro
+ Puro
e
HCT116 mRNA Oncogenic Signatures
TSR3[KO]
WT
0.05
RB P107 DN - UP
CSR LATE UP - UP
CORDENONSI YAP
PRC2 EZH2 UP - DN
GCNP SSH UP LATE - UP
SNF5 DN - UP
E2F3 UP - UP
RB DN - UP
NFE2L2 DN - DN
ATF2 UP - UP
MEK UP - DN
ATM DN - DN
Nominal Enrichment Score (NES)
-log10(FDR Q)
f
HCT116
mRNA
FDR = 1
NES = 0.62
Translational Efficiency
FDR = 0
NES = -3.02
Protein
FDR = 0.048
NES = -1.91
Enrichment Score
WT
TSR3[KO]
Rank Ordered Genes (KEGG_Ribosome)
g
CPTAC-CRC mRNA Oncogenic Signatures
Hypo-m1acp3ψ
Normo-m1acp3ψ
cAMP UP - UP
RB P107 DN - UP
CSR LATE UP - UP
PRC2 EED UP - DN
RPS14 DN - DN
siRNA EIF4GI - UP
MYC UP - UP
E2F1 UP - UP
CAHOY NEURONAL
BRCA1 DN - UP
LEF1 UP - UP
RB P130 DN - DN
RPS14 DN - UP
KRAS KIDNEY UP - UP
NES
-log10(FDR Q)
h
CPTAC-CRC
mRNA
FDR = 0
NES = -2.50
Protein
FDR = 0
NES = -3.03
Normo-m1acp3ψ
Hypo-m1acp3ψ
Rank Ordered Genes (KEGG_Ribosome)

Loss of $m^1acp^3\Psi$ ribosomal RNA modification is a major feature of cancer

Ribosomes are the fulcrum in the central dogma of molecular biology. Multi-omics studies have repeatedly highlighted the discordance between mRNA and protein abundance (Liu et al., 2016; Vogel and Marcotte, 2012), emphasizing the role of translational variability in physio-normal and pathological states. Several recent studies have begun to resolve the ribosome from a uniform assembly into a rich tapestry of functionally heterogeneous complexes making up distinct translational compartments in the cell (2015; Dinman, 2016; Slavov et al., 2015; Xue et al., 2015; Shi et al., 2017; Fujii et al., 2018). Specifically, several rRNA modifications such as pseudouridylation (Ruggero et al., 2003; Jack et al., 2011), ribose 2`-O-methylation (Marcel et al., 2013; Krogh et al., 2016) and, C^5-methyl cytosine (Heissenberger et al., 2019), have been implicated in altering the translational capacity of cancer cells (reviewed in (Jonkhout et al., 2017; Bastide and David, 2018). Together now with $m^1acp^3\Psi$ this supports a broader view of rRNA as a central oncogene, perturbed not genetically, but at the epigenetic modification level. The variable and sub-stochiometric loss of $m^1acp^3\Psi$ in cancer is noteworthy and suggests a sub-fraction of a cancer cell's ribosomes are affected. Further, the loss of 18S.1248 acp^3-modification altered only a sub-set of (RP) mRNA . Together this suggests that the molecular and possibly tumorigenic function of acp^3-modification is mediated by a distinct class of ribosome, although further research is needed to resolve this mechanism.

We have discovered a pervasive and cancer-specific 'onco-ribosome' marked by loss of rRNA $m^1acp^3\Psi$ modification. Enticingly, the cancer-specific $m^1acp^3\Psi$-deficient ribosomes are exceptionally recurrent and can be explored as a novel chemotherapeutic class.

Data and code availability

Sequencing data generated in this study is available on NCBI Short Read Archive (study: PRJNA602544). Proteomics data generated in this studyare available on PRIDE (accession pending). Electronic laboratory notebook for these experiments and analysis scripts are available at https://www.github.com/ababaian/Crown.

Acknowledgments

AB and KR thank Ada Maple Babaian for helpful discussion on the manuscript. This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Leukemia and Lymphoma Society of Canada to DLM. AB was supported by a National Science and Engineering Research Council (NSERC) Alexander Graham Bell Graduate Scholarship and a Roman Babicki Fellowship in Medical Research from the University of British Columbia. Sequence computation was provided by Amazon Web Services (AWS) under a Research Grant to AB. CRISPR-Cas9 reagents were provided by the Integrated DNA Technologies under the CRISPR-Challenge Prize to AB/DLM. Molecular dynamic computation was provided by Compute-Canada. DG was supported by an Alberta Innovates (Technology Futures) Graduate Student Scholarship. HJW was supported by an Alberta Innovates (Strategic Chairs Program SC60-T2) and NSERC Discovery Grant (RGPIN-2016-05199).

Data used in this publication were generated by The Cancer Genome Atlas (TCGA) Research Network (https://www.cancer.gov/tcga), National Cancer Institute Clinical Proteomic Tumor Analysis Consortium (CPTAC) and Genentech/gRED.

Author contributions

A.B. discovered hypo-$m^1acp^3\Psi$ led the study and. A.B. performed DNA- and RNA-seq analysis. A.B. and K.R. performed molecular and cell culture experiments. A.B. invented the aRT-PCR assay for $m^1acp^3\Psi$. S.D. performed primer extension and aRT-PCR optimization. D.G. performed molecular dynamic simulations and H.J.W. and A.B. helped analyze the data. I.M. prepared ribo-seq libraries, performed SCARLET validation and invented the aminocarboxylpropyl biotinylation method. A.B. and M.M processed and analyzed the ribo-seq data. S.E.S.M. performed the LC-MS/MS. H.J.W., M.L., G.M., and D.L.M. provided expert advice for experiments. All authors contributed to the design of the study. A.B. prepared the manuscript and figures.

Competing interests: The authors declare no competing interests.

Correspondence and requests for materials should be addressed to A.B.

Supplementary Information is available for this paper

References

Armistead, J., Khatkar, S., Meyer, B., Mark, B.L., Patel, N., Coghlan, G., Lamont, R.E., Liu, S., Wiechert, J., Cattini, P.A., et al. (2009). Mutation of a gene essential for ribosome biogenesis, EMG1, causes Bowen-Conradi syndrome. Am. J. Hum. Genet. *84*, 728–739.

Babaian, A. (2017). Intra- and Inter-individual genetic variation in human ribosomal RNAs. BioRxiv 118760.

Bastide, A., and David, A. (2018). The ribosome, (slow) beating heart of cancer (stem) cell. Oncogenesis *7*, 1–13.

Bernier, C.R., Petrov, A.S., Waterbury, C.C., Jett, J., Li, F., Freil, L.E., Xiong, X., Wang, L., Migliozzi, B.L.R., Hershkovits, E., et al. (2014). RiboVision suite for visualization and analysis of ribosomes. Faraday Discuss. *169*, 195–207.

Best, R.B., Zhu, X., Shim, J., Lopes, P.E.M., Mittal, J., Feig, M., and Mackerell, A.D. (2012). Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone φ, ψ and side-chain χ(1) and χ(2) dihedral angles. J. Chem. Theory Comput. *8*, 3257–3273.

Bik, H.M., Fournier, D., Sung, W., Bergeron, R.D., and Thomas, W.K. (2013). Intra-Genomic Variation in the Ribosomal Repeats of Nematodes. PLOS ONE *8*, e78230.

Bretones, G., Álvarez, M.G., Arango, J.R., Rodríguez, D., Nadeu, F., Prado, M.A., Valdés-Mas, R., Puente, D.A., Paulo, J.A., Delgado, J., et al. (2018). Altered patterns of global protein synthesis and translational fidelity in RPS15-mutated chronic lymphocytic leukemia. Blood *132*, 2375–2388.

Cancer Genome Atlas Research Network, Weinstein, J.N., Collisson, E.A., Mills, G.B., Shaw, K.R.M., Ozenberger, B.A., Ellrott, K., Shmulevich, I., Sander, C., and Stuart, J.M. (2013). The Cancer Genome Atlas Pan-Cancer analysis project. Nat. Genet. *45*, 1113–1120.

Cech, T.R. (2000). The Ribosome Is a Ribozyme. Science *289*, 878–879.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., et al. (2011). The variant call format and VCFtools. Bioinformatics *27*, 2156–2158.

Denning, E.J., Priyakumar, U.D., Nilsson, L., and Mackerell, A.D. (2011). Impact of 2’-hydroxyl sampling on the conformational properties of RNA: update of the CHARMM all-atom additive force field for RNA. J. Comput. Chem. *32*, 1929–1943.

Dinman, J.D. (2016). Pathways to Specialized Ribosomes: The Brussels Lecture. J. Mol. Biol. *428*, 2186–2194.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

Elder, J.F., and Turner, B.J. (1995). Concerted evolution of repetitive DNA sequences in eukaryotes. Q. Rev. Biol. *70*, 297–320.

Floor, S.N., and Doudna, J.A. (2016). Tunable protein synthesis by transcript isoforms in human cells. ELife *5*, e10921.

Frankish, A., Diekhans, M., Ferreira, A.-M., Johnson, R., Jungreis, I., Loveland, J., Mudge, J.M., Sisu, C., Wright, J., Armstrong, J., et al. (2019). GENCODE reference annotation for the human and mouse genomes. Nucleic Acids Res. *47*, D766–D773.

Friedman, S. (1972). Acylation of transfer ribonucleic acid with the N-hydroxysuccinimide ester of phenoxyacetic acid. Biochemistry (Mosc.) *11*, 3435–3443.

Fujii, K., Susanto, T.T., Saurabh, S., and Barna, M. (2018). Decoding the Function of Expansion Segments in Ribosomes. Mol. Cell *72*, 1013–1020.e6.

Ghandi, M., Huang, F.W., Jané-Valbuena, J., Kryukov, G.V., Lo, C.C., McDonald, E.R., Barretina, J., Gelfand, E.T., Bielski, C.M., Li, H., et al. (2019). Next-generation characterization of the Cancer Cell Line Encyclopedia. Nature *569*, 503–508.

Gillam, I., Blew, D., Warrington, R.C., von Tigerstrom, M., and Tener, G.M. (1968). A general procedure for the isolation of specific transfer ribonucleic acids. Biochemistry (Mosc.) *7*, 3459–3468.

Girardi, T., Vereecke, S., Sulima, S.O., Khan, Y., Fancello, L., Briggs, J.W., Schwab, C., de Beeck, J.O., Verbeeck, J., Royaert, J., et al. (2018). The T-cell leukemia-associated ribosomal RPL10 R98S mutation enhances JAK-STAT signaling. Leukemia *32*, 809–819.

Girodat, D., Mercier, E., Gzyl, K.E., and Wieden, H.-J. (2019). Elongation Factor Tu's Nucleotide Binding Is Governed by a Thermodynamic Landscape Unique among Bacterial Translation Factors. J. Am. Chem. Soc. *141*, 10236–10246.

Gourdazi, K.M., and Lindstrom, M.S. (2016). Role of ribosomal protein mutations in tumor development. Int. J. Oncol. *48*, 1313–1324.

Heissenberger, C., Liendl, L., Nagelreiter, F., Gonskikh, Y., Yang, G., Stelzer, E.M., Krammer, T.L., Micutkova, L., Vogt, S., Kreil, D.P., et al. (2019). Loss of the ribosomal RNA methyltransferase NSUN5 impairs global protein synthesis and normal growth. Nucleic Acids Res. *47*, 11807–11825.

Helm, M., and Motorin, Y. (2017). Detecting RNA modifications in the epitranscriptome: predict and validate. Nat. Rev. Genet. *18*, 275–291.

Hughes, C.S., Foehr, S., Garfield, D.A., Furlong, E.E., Steinmetz, L.M., and Krijgsveld, J. (2014). Ultrasensitive proteome analysis using paramagnetic bead technology. Mol. Syst. Biol. *10*, 757.

Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: visual molecular dynamics. J. Mol. Graph. *14*, 33–38, 27–28.

Ingolia, N.T., Brar, G.A., Rouskin, S., McGeachy, A.M., and Weissman, J.S. (2012). The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. Nat. Protoc. *7*, 1534–1550.

Jack, K., Bellodi, C., Landry, D.M., Niederer, R.O., Meskauskas, A., Musalgaonkar, S., Kopmar, N., Krasnykh, O., Dean, A.M., Thompson, S.R., et al. (2011). rRNA pseudouridylation defects affect ribosomal ligand binding and translational fidelity from yeast to human cells. Mol. Cell *44*, 660–666.

Jindal, S., Ghosh, A., Ismail, A., Singh, N., and Komar, A.A. (2019). Role of the uS9/yS16 C-terminal tail in translation initiation and elongation in Saccharomyces cerevisiae. Nucleic Acids Res. *47*, 806–823.

Jonkhout, N., Tran, J., Smith, M.A., Schonrock, N., Mattick, J.S., and Novoa, E.M. (2017). The RNA modification landscape in human disease. RNA *23*, 1754–1769.

Kampen, K.R., Sulima, S.O., Verbelen, B., Girardi, T., Vereecke, S., Rinaldi, G., Verbeeck, J., Op de Beeck, J., Uyttebroeck, A., Meijerink, J.P.P., et al. (2019). The ribosomal RPL10 R98S mutation drives IRES-dependent BCL-2 translation in T-ALL. Leukemia *33*, 319–332.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. *14*, R36.

Kim, J.-H., Dilthey, A.T., Nagaraja, R., Lee, H.-S., Koren, S., Dudekula, D., Wood III, W.H., Piao, Y., Ogurtsov, A.Y., Utani, K., et al. (2018). Variation in human chromosome 21 ribosomal RNA genes characterized by TAR cloning and long-read sequencing. Nucleic Acids Res. *46*, 6712–6725.

Krogh, N., Jansson, M.D., Häfner, S.J., Tehler, D., Birkedal, U., Christensen-Dalsgaard, M., Lund, A.H., and Nielsen, H. (2016). 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.

Kuo, B.A., Gonzalez, I.L., Gillespie, D.A., and Sylvester, J.E. (1996). Human ribosomal RNA variants from a single individual and their expression in different tissues. Nucleic Acids Res. *24*, 4817–4824.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al. (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860–921.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.

Lareau, L.F., Hite, D.H., Hogan, G.J., and Brown, P.O. (2014). Distinct stages of the translation elongation cycle revealed by sequencing ribosome-protected mRNA fragments. ELife *3*, e01257.

Lauria, F., Tebaldi, T., Bernabò, P., Groen, E.J.N., Gillingwater, T.H., and Viero, G. (2018). riboWaltz: Optimization of ribosome P-site positioning in ribosome profiling data. PLOS Comput. Biol. *14*, e1006169.

Leffers, H., and Andersen, A.H. (1993). The sequence of 28S ribosomal RNA varies within and between human cell lines. Nucleic Acids Res. *21*, 1449–1455.

Li, W., Ward, F.R., McClure, K.F., Chang, S.T.-L., Montabana, E., Liras, S., Dullea, R.G., and Cate, J.H.D. (2019). Structural basis for selective stalling of human ribosome nascent chain complexes by a drug-like molecule. Nat. Struct. Mol. Biol. *26*, 501–509.

Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J.P., and Tamayo, P. (2015). The Molecular Signatures Database Hallmark Gene Set Collection. Cell Syst. *1*, 417–425.

Lim, B., Mun, J., Kim, J.-H., Kim, C.W., Roh, S.A., Cho, D.-H., Kim, Y.S., Kim, S.-Y., and Kim, J.C. (2015). Genome-wide mutation profiles of colorectal tumors and associated liver metastases at the exome and transcriptome levels. Oncotarget *6*, 22179–22190.

Liu, N., Parisien, M., Dai, Q., Zheng, G., He, C., and Pan, T. (2013). Probing N6-methyladenosine RNA modification status at single nucleotide resolution in mRNA and long noncoding RNA. RNA *19*, 1848–1856.

Liu, Y., Beyer, A., and Aebersold, R. (2016). On the Dependency of Cellular Protein Levels on mRNA Abundance. Cell *165*, 535–550.

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

Machnicka, M.A., Milanowska, K., Osman Oglou, O., Purta, E., Kurkowska, M., Olchowik, A., Januszewski, W., Kalinowski, S., Dunin-Horkawicz, S., Rother, K.M., et al. (2013). MODOMICS: a database of RNA modification pathways--2013 update. Nucleic Acids Res. *41*, D262-267.

Marcel, V., Ghayad, S.E., Belin, S., Therizols, G., Morel, A.-P., Solano-Gonzàlez, E., Vendrell, J.A., Hacot, S., Mertani, H.C., Albaret, M.A., et al. (2013). p53 acts as a safeguard of translational control by regulating fibrillarin and rRNA methylation in cancer. Cancer Cell *24*, 318–330.

Masek, T., Vopalensky, V., Suchomelova, P., and Pospisek, M. (2005). Denaturing RNA electrophoresis in TAE agarose gels. Anal. Biochem. *336*, 46–50.

McCoy, L.S., Xie, Y., and Tor, Y. (2011). Antibiotics that target protein synthesis. Wiley Interdiscip. Rev. RNA *2*, 209–232.

Meyer, B., Wurm, J.P., Sharma, S., Immer, C., Pogoryelov, D., Kötter, P., Lafontaine, D.L.J., Wöhnert, J., and Entian, K.-D. (2016). Ribosome biogenesis factor Tsr3 is the aminocarboxypropyl transferase responsible for 18S rRNA hypermodification in yeast and humans. Nucleic Acids Res. *44*, 4304–4316.

Nakhoul, H., Ke, J., Zhou, X., Liao, W., Zeng, S.X., and Lu, H. (2014). Ribosomopathies: Mechanisms of Disease. Clin. Med. Insights Blood Disord. *7*, 7–16.

Natchiar, S.K., Myasnikov, A.G., Kratzat, H., Hazemann, I., and Klaholz, B.P. (2017). Visualization of chemical modifications in the human 80S ribosome structure. Nature *551*, 472–477.

Nieminen, T.T., O'Donohue, M.-F., Wu, Y., Lohi, H., Scherer, S.W., Paterson, A.D., Ellonen, P., Abdel-Rahman, W.M., Valo, S., Mecklin, J.-P., et al. (2014). Germline mutation of RPS20, encoding a ribosomal protein, causes predisposition to hereditary nonpolyposis colorectal carcinoma without DNA mismatch repair deficiency. Gastroenterology *147*, 595–598.e5.

Parks, M.M., Kurylo, C.M., Dass, R.A., Bojmar, L., Lyden, D., Vincent, C.T., and Blanchard, S.C. (2018). Variant ribosomal RNA alleles are conserved and exhibit tissue-specific expression. Sci. Adv. *4*, eaao0665.

Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T.-C., Mendell, J.T., and Salzberg, S.L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat. Biotechnol. *33*, 290–295.

Petrov, A.S., Bernier, C.R., Hsiao, C., Norris, A.M., Kovacs, N.A., Waterbury, C.C., Stepanov, V.G., Harvey, S.C., Fox, G.E., Wartell, R.M., et al. (2014). Evolution of the ribosome at atomic resolution. Proc. Natl. Acad. Sci. U. S. A. *111*, 10251–10256.

Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kalé, L., and Schulten, K. (2005). Scalable molecular dynamics with NAMD. J. Comput. Chem. *26*, 1781–1802.

Rabanal, F.A., Nizhynska, V., Mandáková, T., Novikova, P.Y., Lysak, M.A., Mott, R., and Nordborg, M. (2017). Unstable Inheritance of 45S rRNA Genes in Arabidopsis thaliana. G3 GenesGenomesGenetics *7*, 1201–1209.

Ruggero, D., Grisendi, S., Piazza, F., Rego, E., Mari, F., Rao, P.H., Cordon-Cardo, C., and Pandolfi, P.P. (2003). Dyskeratosis congenita and cancer in mice deficient in ribosomal RNA modification. Science *299*, 259–262.

Schmidt, E.K., Clavarino, G., Ceppi, M., and Pierre, P. (2009). SUnSET, a nonradioactive method to monitor protein synthesis. Nat. Methods *6*, 275–277.

Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nat. Methods *9*, 671–675.

Schofield, P., Hoffman, B.M., and Rich, A. (1970). Spin-labeling studies of aminoacyl transfer ribonucleic acid. Biochemistry (Mosc.) *9*, 2525–2533.

Schuster, C.F., and Bertram, R. (2014). Fluorescence based primer extension technique to determine transcriptional starting points and cleavage sites of RNases in vivo. J. Vis. Exp. JoVE e52134.

Seshagiri, S., Stawiski, E.W., Durinck, S., Modrusan, Z., Storm, E.E., Conboy, C.B., Chaudhuri, S., Guan, Y., Janakiraman, V., Jaiswal, B.S., et al. (2012). Recurrent R-spondin fusions in colon cancer. Nature *488*, 660–664.

Shi, Z., Fujii, K., Kovary, K.M., Genuth, N.R., Röst, H.L., Teruel, M.N., and Barna, M. (2017). Heterogeneous Ribosomes Preferentially Translate Distinct Subpools of mRNAs Genome-wide. Mol. Cell *67*, 71–83.e7.

Slavov, N., Semrau, S., Airoldi, E., Budnik, B., and van Oudenaarden, A. (2015). Differential Stoichiometry among Core Ribosomal Proteins. Cell Rep. *13*, 865–873.

Sloan, K.E., Warda, A.S., Sharma, S., Entian, K.-D., Lafontaine, D.L.J., and Bohnsack, M.T. (2017). Tuning the ribosome: The influence of rRNA modification on eukaryotic ribosome biogenesis and function. RNA Biol. *14*, 1138–1152.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. *102*, 15545–15550.

Tan, C., and Du, X. (2012). KRAS mutation testing in metastatic colorectal cancer. World J. Gastroenterol. WJG *18*, 5171–5180.

Taoka, M., Nobe, Y., Yamaki, Y., Yamauchi, Y., Ishikawa, H., Takahashi, N., Nakayama, H., and Isobe, T. (2016). The complete chemical structure of Saccharomyces cerevisiae rRNA: partial pseudouridylation of U2345 in 25S rRNA by snoRNA snR9. Nucleic Acids Res. *44*, 8951–8961.

Taoka, M., Nobe, Y., Yamaki, Y., Sato, K., Ishikawa, H., Izumikawa, K., Yamauchi, Y., Hirota, K., Nakayama, H., Takahashi, N., et al. (2018). Landscape of the complete RNA chemical modifications in the human 80S ribosome. Nucleic Acids Res. *46*, 9289–9298.

Vasaikar, S., Huang, C., Wang, X., Petyuk, V.A., Savage, S.R., Wen, B., Dou, Y., Zhang, Y., Shi, Z., Arshad, O.A., et al. (2019). Proteogenomic Analysis of Human Colon Cancer Reveals New Therapeutic Opportunities. Cell *177*, 1035–1049.e19.

Vlachos, A., Rosenberg, P.S., Atsidaftos, E., Alter, B.P., and Lipton, J.M. (2012). Incidence of neoplasia in Diamond Blackfan anemia: a report from the Diamond Blackfan Anemia Registry. Blood *119*, 3815–3819.

Vogel, C., and Marcotte, E.M. (2012). Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. Nat. Rev. Genet. *13*, 227–232.

Wurm, J.P., Meyer, B., Bahr, U., Held, M., Frolow, O., Kötter, P., Engels, J.W., Heckel, A., Karas, M., Entian, K.-D., et al. (2010). The ribosome assembly factor Nep1 responsible for Bowen-Conradi syndrome is a pseudouridine-N1-specific methyltransferase. Nucleic Acids Res. *38*, 2387–2398.

Xu, Y., Vanommeslaeghe, K., Aleksandrov, A., MacKerell, A.D., and Nilsson, L. (2016). Additive CHARMM force field for naturally occurring modified ribonucleotides. J. Comput. Chem. *37*, 896–912.

Xue, S., Tian, S., Fujii, K., Kladwang, W., Das, R., and Barna, M. (2015). RNA regulons in Hox 5′ UTRs confer ribosome specificity to gene regulation. Nature *517*, 33–38.

Yang, J., Sharma, S., Watzinger, P., Hartmann, J.D., Kötter, P., and Entian, K.-D. (2016). Mapping of Complete Set of Ribose and Base Modifications of Yeast rRNA by RP-HPLC and Mung Bean Nuclease Assay. PLoS ONE *11*.

Supplementary Figure 1: Detailed look at hypo-$m^1acp^3\Psi$ in the TCGA cohorts

a, The 18S.1248.U variant allele frequency (VAF) from 33 TCGA patient cohorts (study abbreviations in table S1). **b**, The rDNA position underlying *18S*.1248.U (*RNA45S*.4908.T) is invariable in TCGA-COAD. Matched tumor-normal whole exome sequencing (WXS) data from 438 TCGA-COAD patients was aligned to *hgr1*. 415 cancer and 413 normal samples contained rDNA coverage of *RNA45S*.4908.T. Three normal samples had a single variant read each at this position, consistent with sequencing error. **c**, Batch-specific shift in the average 18S.1248.U VAF in i, TCGA-COAD and ii, TCGA-DLBC libraries. Similar to seen in the aRT-PCR assay for $m^1acp^3\Psi$ (Figure S2), there are batch-effects with $m^1acp^3\Psi$ misincorporation, but the relative decrease in $m^1acp^3\Psi$-modification in CRC compared to normals is seen in across all batches. **d**, The gene expression of the $m^1acp^3\Psi$ modifying enzymes *TSR3* and *EMG1* is not decreased or lost across the TCGA cohorts.

Supplementary Figure 2: Aminocarboxyl propyl Reverse Transcription (aRT)-PCR assay for $m^1acp^3\Psi$ and rRNA modification in cell lines

a i, The 18S.1248.$m^1acp^3\Psi$ modification assay is based on the misincorporation of nucleotides in first strand complementary DNA (cDNA) strand synthesis by reverse transcriptase (RT). The cDNA is then PCR amplified and **ii,** the ratio of reference T and not-T (V = A, C, or G) is genotyped by the *HinFI* restriction enzyme cut-site which overlaps 18S.1248. **b,** The choice of RT-enzyme; SuperScript III (SSIII), SuperScript IV (SSIV), WarmStart RTx (WS RTx) or, UltraScript 2.0 (US 2.0), influences nucleotide misincorporation rates and the variant allele frequency (VAF) read-out of the m^1acp^3-assay, although VAF remains consistent across biological replicates of input RNA of the colorectal cancer (CRC) cell line HCT116 wildtype clone 1 (WT[1]), or HCT116 with *TSR3* gene knockout clone 1 (*TSR3*[KO 1]). **c,** PCR replicates of WT[1] and TSR3[KO 1] cDNA, shows consistent readout. **d,** HCT116 WT[1] and *TSR3*[KO 1] RNA was mixed at fixed weight ratios (µg total RNA) prior to RT to determine if the assay is quantitative for m_1acp_3 modification. **e,** The aRT-PCR assay applied to 11 CRC cell lines, primary human umbilical vein endothelial cells (HUVEC) as a normal control and, the blast-phase chronic myelogeous leukemia cell line K562 as a hypo-$m^1acp^3\Psi$ positive control. **f i,** Polysomal fractionation and **ii,** sub-fraction $m^1acp^3\Psi$ RT-PCR assay of the hypo-$m^1acp^3\Psi$ cell line K562. In cells containing a mixture of +/- $m^1acp^3\Psi$ modification, unmodified rRNA incorporates into mature ribosomes and is enriched in the lower-order mono- and di-somes. **g,** Primer extension assay for 18S.1248.$m^1acp^3\Psi$ modification in HCT116 WT[1] and *TSR3*[KO 1]. The helix 31 structural stop and

rRNA truncation via DNA cut oligo + RNase H treatment is used as internal load controls. **h,** SCARLET assay of HCT116 WT[1] and *TSR3*[KO 1] is not suitable for measuring $m^1acp^3\Psi$ as this base may interfere with probe hybridization, but precursor bases were detected and accumulated in the TSR3 knock-out cells. **i,** Total RNA treated with the amino-reactive ester, N-hydroxysuccinimidobiotin and detected with **i.** ethidium bromide staining as load control and **i.** anti-streptavidin-HRP antibody and. *TSR3*[KO 1] cells show loss of primary amine reactivity (such as $m^1\underline{a}cp^3\Psi$) in 18S rRNA but not in tRNA.

Supplementary Figure 3: 18S.1248.$m^1acp^3\Psi$ is absolutely conserved in *Eukarya*

a, Evolutionary conservation of 18S rRNA in the locus surrounding helix 31 in *Eukarya* and select *Archaea* and *Eubacteria* species (Bernier et al., 2014; Petrov et al., 2014). Magenta arrow indicates the position homologous to human18S:1248.U. **b-d,** The conserved secondary structure of helix 31 and its known modification sites (Machnicka et al., 2013; Petrov et al., 2014).

Supplementary Figure 4: Ribosomal molecular dynamics and modification modeling

a, The $m^1acp^3\Psi$ modification stabilizes the decoding peptidyl (P-) site via a hydrogen bond with the universally conserved RPS16 (uS9) p.R146 residue. Whole ribosome molecular dynamics simulations (MD) were ran for up to 50ns with 18S.1248.$m^1acp^3\Psi$, $acp^3\Psi$, $m^1\Psi$, Ψ or U base in an empty (**b,d,** PDB: 6EK0) or tRNA-occupied P-site (**c,e**, PDB: 6OLE). **b,i.** Root mean squared deviation (RMSD) of empty P-site MD atoms indicate simulations stabilize after 5ns (with up to ~6Å RMSD), up to 45ns (yellow highlight) was used for analysis. **ii,** Minimal distance between $m^1acp^3\Psi$ (3-carboxyl oxygen) or uridine (4-oxygen) and the closest guanidinium hydrogen of RPS16 p.R146 supports that m^1 and acp^3 are necessary to be within 3.5Å for hydrogen bonding and nucleotide stabilization. **c,** RMSD of P-site occupied simulations with **ii.** Root mean squared fluctuation (RMSF) of 18S rRNA adjacent to 18S.1248 and **iii,** within helix 31 shows 18S.1248.$m^1acp^3\Psi$ is comparable relative to other modification states. The average (**iv**) and standard deviation of χ-angle of the 18S.1248 nucleotide in each simulation shows acp^3 increases torsional variation of the base which is stabilized by the m^1 modification. **d**, 20ns simulation showing bonding between 18S.1248.$m^1acp^3\Psi$ and RPS16 p.R146 compared to **e,** the structure with mRNA and P-site tRNA. We postulate that 18S.1248 acp^3-modification is involved in coordinating RPS16 p.146 for P-site tRNA positioning and contributes to the stability of the decoding core. Statistical difference between χ-angle mean (Tukey HSD), and

standard deviation (Benjamini-Hochberg corrected F-tests) were used (* is $p < 0.05$, ** is $p < 0.001$ and *** is $p < 0.0001$).

Supplementary Figure 5: RNA-seq and *TSR3* editing site validation

HCT116 WT[1-3] versus *TSR3*[KO]/[Het] RNA-seq metrics. **a,** Differential mRNA expression of expressed (reads per million kilobase, $RPKM_{mRNA}$ >0.1) between WT[1-3] and *TSR3*[KO]/[Het] clones. **b,** Hierarchical clustering of libraries based on expressed genes. Globally, *TSR3*[Het 2] is more dissimilar to *TSR3*[KO 1,2] clones. **c,** MA-plot for mRNA expression highlighting genes in the '*KEGG_RIBOSOME'* and '*RB_P107_DN.V1_UP*' gene sets (see: Fig 3). **d,** Volcano plot for differentially expressed genes between WT and *TSR3*[KO]/[Het] with statistically significant genes highlighted in red (list is available in Table S5). **e,** *TSR3*[KO]/[Het] lines have decreased *TSR3* expression measured by RNA-seq (in reads per kilobase per million mapped reads, RPKM). **f,** *TSR3* gene structure and IGV screenshots showing the editing sites in (**g**) *TSR3*[KO 1,3], and (**h)** *TSR3*[Het 2].

Supplementary Figure 6: Ribo-seq

HCT116 WT[1-3] versus *TSR3*[KO]/[Het] ribo-seq metrics. **a, i,** As a quality control metric, the length distribution of mapped ribosome-protected fragments for each of the two WT[1-3], and three *TSR3*[KO]/[Het] biological replicates of ribo-seq libraries were plotted. The *TSR3*[KO 3] biological replicate 2 (r2) library had a bi-modal read-length distribution, peaking at 22 and 28 nt suggesting incomplete cycloheximide treatment (Lareau et al., 2014) thus, this library was excluded from downstream expression and positional analyses. **ii,** Short (21-23 nt) ribosome fragments coincide with ribosomes stalled in the rotated, post peptide-bond state (Lareau et al., 2014). *TSR3*[KO 1,3] libraries had fewer short fragments implying $m^1acp^3\Psi$-deficient ribosomes have a lower probability of being in the rotated transition state relative to WT ribosomes. **b,** Hierarchical clustering of libraries based on total translation recapitulates mRNA clustering. **c,** Metagene heatmap of P-site occupancy around start and stop codons. For each read optimal P-site offset was determined (13 nt from 5' end) and P-site occupancy computed. Log10-transformed P-site coverage is shown for all replicates, scaled by min and max number of reads per sample. A clear trinucleotide periodicity in the CDS of transcripts is observed. **d,** Differential translation of ribo-seq expressed ($RPKM_{Ribo}$ > 0.1) genes. **e,** MA-plot and **f,** volcano-plot for total translation, highlighting the ribosome and RB/E2F gene sets. **g,** The $\log_2$ mRNA fold-change and $\log_2$ translation fold-change (WT / *TSR3*[KO]/[Het]) and detail inlay, with each gene size-scaled by

mean $RPKM_{mRNA}$. RB/E2F gene sets are highlighted, showing RP genes are more efficiently translated in *TSR3*[KO]/[Het] clones (all points below diagonal in, see also Fig 3). **h,** The P-site periodicity within each library showed the majority of CDS ribosomes were in-frame, with no significant difference in frame-shifting upon $m^1acp^3\Psi$ perturbation. **i,** P-site occupancy was calculated over all expressed coding sequences (CDS). Globally, there was no significant difference between P-site occupancy per codon in WT[1-3] in *TSR3*[KO]/[Het] libraries. Since 18S.1248.$m^1acp^3\Psi$ is located at the P-site where initiation codon selection occurs, we tested if the initiation AUG codon was differentially occupied between any genotypes. *TSR3*[Het 2] and *TSR3*[KO 1], but not *TSR3*[KO 3] have elevated AUG occupancy relative to WT clones supporting a slower global initiation rate in those two samples. Tukey HSD test was used for testing a statistical difference between group means (* is $p < 0.05$, ** is $p < 0.001$ and *** is $p < 0.0001$).

Supplementary Figure 7: Liquid-Chromatography Tandem Mass Spetrometry (LC-MS/MS) validation

HCT116 WT[1-3] versus *TSR3*[KO]/[Het] LC-MS/MS results. **a,** After multiple-testing correction, no proteins were significantly different between genotypes, and no TSR3 peptides were found in the samples. **b-d**, Volcano plots for RNA-seq, ribo-seq, translational efficiency (Ribo-seq_{RPKM}/RNA-seq_{RPKM}) and proteomic data with genes in the `KEGG_RIBOSOME` gene-set highlighted. As a complimentary analysis to GSEA, a Fisher's Exact Test was performed to test for enrichment of `KEGG_RIBOSOME` genes either up- or down-regulated between genotypes, compared to all other genes.

Supplementary Figure 8: The RB/E2F transcriptional signature associated with *TSR*[KO]/[Het]

HCT116 WT[1-3] versus *TSR3*[KO]/[Het] Gene Set Enrichment analysis (GSEA) for the E2F Targets and oncogenic signature of *RB1*, *RBL1* (p107) knock-out. **a,** Rb proteins are repressors of the E2F transcription factors. The of E2F GSEA (**i**) was highly enriched upon *TSR3*[KO]/[Het] with (**ii**) *TSR3*[Het 2] showing the strongest enrichment. **b,** Gene set mRNA expression is enriched specifically in genes upregulated upon Rb1 and Rb1;p107 knockout (gene sets: *RB_DN.V1_UP*, *RB_P107_DN.V1_UP*). **c,** The translational output (ribo-seq signal) of genes upregulated in Rb1;p107 knockout remains increased but **d,** this gene set is translated less efficiently and, **e,** has no difference at the proteome level.

Supplementary Figure 9: Representative changes to translational efficiency

Ribo-seq traces for representative genes **a**, *EEF1A2* (eukaryotic translation elongation factor 1 alpha 2), **b**, *RPS8* (ribosomal protein S8), and **c**, *RPL4* (ribosomal protein L4) for HCT116 WT[1-3] versus *TSR3*[KO]/[Het]. **d**, Matched RNA expression (in reads per kilobase per million mapped reads, RPKM), Ribo-seq CDS expression (RPKM), translational efficiency (Ribo-seq$_{RPKM}$/RNA-seq$_{RPKM}$), and proteomic expression (median normalized expression values). **f**, Polysome profiles and relative area under the curve (AUC) intensity for sub-fractions (highlighted in gray) (inlay). Total absorbance was normalized to 100 artificial units in the working range of the gradient (between red ticks on x-axis).

Supplementary Table 1:

Accessions of DNA and RNA sequencing libraries used in this study. 18S.1248.U variant allele frequency (VAF) is provided for each sample.

Supplementary Table 2:

Primers, oligos, and guide RNA sequences.

Supplementary Table 3:

Significant (FDR q-value < 0.05) Gene Set Enrichment Analysis results for differential transcriptomics in HCT116 WT[1-3] versus *TSR3*[KO]/[Het] with the Hallmarks, C3 promoter motif, and C6 oncogenic signatures gene sets.

Supplementary Table 4:

Significant (FDR q-value < 0.05) Gene Set Enrichment Analysis results with the C2 gene-ontology and pathways gene sets for differential transcriptomics and translational efficiency in HCT116 WT[1-3] versus *TSR3*[KO]/[Het] cell line and transcriptomics and proteomics in normo-$m^1acp^3\Psi$ versus hypo-$m^1acp^3\Psi$ CPTAC-CRC tumors.

Supplementary Table 5:

Significant (adjust P-value < 0.05) genes deferentially expressed between WT[1-3], and *TSR3*[KO 1,3]/[Het 2] in the RNA-seq and Ribo-seq experiments.

Extended Methods

Ribosomal sequence alignment and variant allele frequency calculations

RNA-seq libraries used in this study were prepared via poly-A selection to enrich for the ~5% of mRNA from total RNA. Since rRNA is ~80% of cellular RNA it invariably 'contaminates' RNA-seq libraries. Typically, poly-(A) RNA-seq libraries contain 3.55% (+/- 0.685, 95% CI, CRC I cohort, N = 66) of total reads aligned to rDNA. Whole genome DNA-seq (WGS) libraries were processed from initial fastq files. The 876 TCGA-COAD whole exome DNA-seq (WXS) aligned libraries were downloaded and reads mapping to the contig "chrUn_GL000220v1" (which contains a complete *RNA45S* gene) were extracted with `*samtools*` into fastq files. 415/438 cancer samples and 413/438 normal samples had rDNA coverage at a total of 15,175x and 17,936x, respectively. A complete list of library accessions used in this study is available in Table S1 (Cancer Genome Atlas Research Network et al., 2013; Ghandi et al., 2019; Lim et al., 2015; Seshagiri et al., 2012).

Libraries were aligned to the *hgr1* reference rDNA sequence (Babaian, 2017) with *bowtie2* (v. 2.3.5.1, command: `*bowtie2 --very-sensitive-local -x hgr1 -1 <read1.fq.gz> -2 <read2.fq.gz>*`) (Langmead and Salzberg, 2012). For each cohort of libraries, a genomic variant call format (GVCF) was created with *bcftools* (v. 1.9, command: `*bcftools mpileup -f hgr1.fa –max-depth 10000 -A -min-BQ 30 -b <bam.file.list>*`) (Danecek et al., 2011).

GVCF files were processed in R by custom scripts to calculate variant allele frequency (VAF). VAF is defined as 1 – reference allele frequency (reference allele depth of coverage / total depth of coverage) (scripts available at https://www.github.com/ababaian/crown)

The threshold to define hypo-modification of an RNA base (including 18S.1248.$m^1acp^3\Psi$) was defined as three standard deviations below average VAF of the normal samples within the same cohort (false discovery rate = 0.00135) when available. Fixed formalin paraffin embedded (FFPE) libraries in TCGA were negative for 18S.1248.$m^1acp^3\Psi$, 28S.1321.m^1A and 28S.4532.m^3U modification signatures and excluded from further analysis. In the CPTAC-CRC cohort (normal RNA-seq is unavailable), hypo-$m^1acp^3\Psi$ and normo-$m^1acp^3\Psi$ was defined by the lower (<25%) and upper (>75%) quantiles of samples within a batch.

Transcriptome and translatome alignment, assembly and differential expression

RNA-seq reads were aligned to *hg38* (*GRCh38)* reference genome with *tophat2* (v.2.0.14) (Kim et al., 2013). Individual transcriptome assemblies for HCT116 [WT 1-3], [KO 1,2] and [Het 2] libraries were generated with *stringtie* (v 2.0) (Pertea et al., 2015), and then all merged together with the human *gencode basic* gene annotation (v. 31) (Frankish et al., 2019) ultimately yielding the *hct116_gencode.v31* reference gene set.

To generate a single-copy reference transcriptome for ribo-seq analysis of HCT116, isoform-specific quantification of gene expression was performed on the *hct116_gencode.v31* gene set with `*stringtie -G hct116_gencode.gtf*`. For each gene with non-zero expression (>10 unique reads), the one highest expression isoform (average expression from each clone) was chosen as the reference transcript for that gene.

For ribo-seq alignment, after read adapter trimming and alignment to *hgr1* as above, unmapped reads were aligned against a containment file containing human tRNA, mtDNA, snoRNA, snRNA and miRNA sequences. Reads remaining unmapped were then aligned to *hg38* and the *hct116_transcriptome* with *STAR* aligner (v. 2.5.2b, command: `*STAR --genomeDir hg38 --readFilesIn <input.fq> --sjdbFileChrStartEnd hg38/sjdbList.out.tab --outFilterMultimapNmax 10 --outFilterMismatchNmax 5 --outFilterMatchNmin 15 --alignSJoverhangMin 5 --seedSearchStartLmax 20 --outSJfilterOverhangMin 30 8 8 8 --quantMode TranscriptomeSAM*`) (Dobin et al., 2013). Transcriptome aligned Ribo-seq data was analyzed in *R* (v. 3.5.1) using the *riboWaltz* package (v. 1.1.0) (Lauria et al., 2018).

Gene-level expression and total translation was quantified with the *DEseq2* (Love et al., 2014) R package using *hg38* aligned bam files and the *hct116_gencode.v31* reference gene set. Ribo-seq Gene expression (RPKM) was calculated based only on reads mapping in-frame to a gene CDS, measure translating ribosomes. Translational efficiency was calculated per genotype as log2(Ribo-seq $Gene_{RPKM}$ / RNA-seq $Gene_{RPKM}$).

Gene expression and translation differences were calculated by Gene Set Enrichment Analysis (*GSEA*, v.4.0.0) (Subramanian et al., 2005) with `*-permute gene_set -nperm 5000*` and standard parameters. Transcriptomic GSEA was performed using the MSigDB (Liberzon et al., 2015) (v 7.0): hallmark, C2 pathways, C3 motif search, and C6 oncogenic signatures gene sets. Translatomic and proteomic GSEA was performed with C5 Gene Ontology (GO) gene set.

All bioinformatic analyses were scripted for reproducibility and are available at https://www.github.com/ababaian/crown, RNA-seq and Ribo-seq data can be directly viewed in the UCSC genome browser at https://genome.ucsc.edu/s/Artem%20Babaian/HCT116_TSR3%2DKO.

HCT116 cell culture and TSR3 knockout

The colorectal carcinoma cell line HCT116 (CCL-247, ATCC, Manassas, VA) was cultured in complete media [DMEM media (#36250, STEMCELL Technologies, Vancouver, Canada) supplemented with 10% fetal bovine serum (F1051, Invitrogen, Waltham, MA)].

To generate *TSR3* knockouts, 10^5 HCT116 cells were transfected with 10 nmol of one of three *TSR3* targeting Alt-R CRISPR-Cas9 ribonucleoproteins or non-targeting controls (table S2) by manufacturer's protocol (1081059, Integrated DNA Technologies (IDT), Coralville, IA). After 24 hours, single cells from each treatment group were isolated by limiting dilution and confirmed to be 1 cell /well by microscopy. Single cell clones were expanded to $5x10^5$ cells at which point half the culture was frozen (culture media + 10% DMSO) and half were processed for RNA. *TSR3* knockouts were genotyped by RNA-seq and functional knockout was confirmed by three independent $m^1acp^3\Psi$ assays (figure 3, S2).

Cell lines and clonal isolates were tested to be free of mycoplasma contamination by DAPI staining and microscopy and with LookOut Mycoplasma Detection Kit (MP0035, MilliporeSigma, Burlington, MA) by manufacturer’s protocol.

Cell cycle and puromycin incorporation assays

HCT116 cells were seeded at $2x10^5$ cells per 6-well plate and allowed to grow until log-phase of growth at ~50-75% confluency. To detach cells, cells were washed twice with PBS, detached with 0.5 ml 0.25% trypsin EDTA (07901, STEMCELL) for 5 minutes at RT, followed by inactivation with 2.5 ml of complete media and cells pelleted at 300 x g for 5 minutes at 4°C. Pellets were washed twice with cold PBS.

For cell cycle analysis, cells were fixed overnight at 4°C in 70% EtOH solution. Fixed cells were washed in PBS and resuspended in 500 ul of staining solution [Propidium Iodide (50 ug/ml), RNAse A (0.1 mg/ml), and 0.05% Triton X-100 in PBS], and incubated 30 minutes at RT in the dark followed by a wash step. Flow cytometry was performed on a FACScalibur with *CellQuest* acquisition software (BD Biosciences) acquiring at least 10,000 cells per sample and analyzed with *FlowJo* (v 10.0.8). Cell cycle gating was performed on single-cells using the automated ‘Watson Model’.

Puromycin-incorporation assay was performed as previously described (Schmidt et al., 2009). Briefly, cells were pulsed with either complete media only or complete media containing 10 ug/mL Puromycin (73342, STEMCELL) for 10 minutes in cell incubator. Cells were washed twice in pre-warmed PBS and chased for 50 minutes in incubator. Cells were then detached, washed and fixed in 2% paraformaldehyde in PBS for 15 minutes at RT. Cells were then washed twice in wash solution (5% FBS, and 0.1% Saponin in PBS). $1x10^5$ cells were stained in 100 ul of 1:100 anti-puromycin antibody (12D10) conjugated to Alexa Flour 488 (MAEBE343-AF488, MilliporeSigma) in wash solution for 30 minutes on ice followed by two final washes in wash solution and resuspension in 100 ul PBS for flow cytometry to acquire at least 5,000 cells per sample. Unstained controls were performed on puromycin treated cells. Puromycin staining intensity correlated with natural variation in cell size (forward scatter signal, FSC), thus the 488nm staining signal intensity (anti-puromycin) was normalized by cell size (normalized fluorescence = 250 * anti-puromycin / FSC). The geometric mean florescent intensity (gMFI) and standard deviation for each sample is reported.

RNA isolation

Cells for RNA extraction were lysed directly in TRIzol reagent (15596-018, Invitrogen), spun 5 min at 12,000 x g to pellet fat and nuclear DNA and then frozen at -80°C. RNA extraction was carried out by manufacturer's protocol. RNA quality was assessed via 2% denaturing RNA agarose gel electrophoresis (heat treated, 95°C for 5 minutes in 1.5x formamide loading buffer (Masek et al., 2005)) and concentration/purity assessed by spectrophotometer (NanoDrop 2000, ThermoFisher, Waltham, MA). RNA quality for RNA-seq library preparation had a >9.9 RIN score measured by Bioanalyzer 2100 (Agilent, Santa Clara, CA).

Primer extension for 18S.1248.$m^1acp^3\Psi$ modification

Primer extension was performed with 1 µg of total RNA, incubated with 2 pmol of *PE_1248_BLOCK* (IDT) primer and 2U of RNase H (18021-014, Invitrogen) or mock enzyme treatment at 37°C for 20 min followed by heat inactivation at 65°C for 10 min. SuperScript III reverse transcriptase (18080044, lot #2042663, Invitrogen) and the fluorophore labeled *PE_1248_FAM* primer were added for primer annealing and RT (1h at 50°C) as described by Schuster and Bertram (Schuster and Bertram, 2014). Labeled cDNAs were re-suspended in 1.5x formamide loading buffer and heated to 95°C for 3 min to eliminate secondary structures (Masek et al., 2005). Samples were separated on a 2% agarose gel at 114 V for 3h at 4°C or on a 12.5% polyacrylamide gel at 45 mA for 2.5h in 1x TBE. After migration, the

gel was visualized with the Typhoon FLA 9500 laser scanner (FAM filter, 50 µm pixel and 450 V unless otherwise noted, GE Healthcare, Chicago, IL).

Aminocarboxyl propyl Reverse Transcription (aRT)-PCR for 18S.1248.m^{1}acp^{3}Ψ modification

The aminocarboxylpropyl reverse transcription (aRT)-PCR assay for 1248.m^{1}acp^{3}Ψ modification was performed with 1 ug of DNase treated (AM1907, lot #00733051, Invitrogen) RNA after total RNA quality was assessed by denaturing agarose gel electrophoresis (Masek et al., 2005). RT reaction was carried out with SuperScript III (Invitrogen), SuperScript IV (18090010, lot #00721480, Invitrogen), UltraScript 2.0 (PB30.31-10, lot #PB130614-01-5, PCR Biosystems, Wayne PA) and WarmStart RTx (M0380L, lot #0061705, NEB) by each manufacturer's protocol with minor modifications. aRT reactions were carried out with a random hexamer primer only, and not poly(T) oligos. cDNAs were diluted five-fold and used as template for PCR (30 cycles: 94°C for 30 s, 55°C for 30 s, 72°C for 30 s) with *macp_F1* and *macp_R1* primers (table S2). Amplicons were digested with HinFI (R0155S, New England BioLabs) (25°C for 5 sec, 37°C for 90 min, 80°C for 20 min). Samples were separated on a 2.25% agarose gel in 1x TBE at 200 V for 45 min at 4°C. After migration, the gel was post-stained in 1x GelRed (41003, Biotium, Fremont CA) for 30 min. Gels were visualized by UV transillumination, captured in gray scale with a digital camera and pseudo-colored in ImageJ (Schneider et al., 2012) (v 1.52h, Lookup table > Fire) which retains the original pixel intensity values but highlights band-intensity visualization.

SCARLET for 18S.1248.m^{1}acp^{3}Ψ modification

The validation of 18S.1248.m1acp3Ψ was performed by site-specific cleavage and radioactive-labeling followed by ligation-assisted extraction and thin-layer chromatography (SCARLET) as previously described (Liu et al., 2013) with minor modifications.

Briefly, 1 µg of total RNA from HCT116 WT or *TSR3*[KO] was annealed with 3 pmol of chimera oligo (oligo sequences in Table S2) in 2.5 µl 30 mM Tris-HCl pH 7.5, by heating at 95 °C for 1 min. The sample was incubated at 44 °C for 1 h in total volume of 5 µl 0.4× T4 PNK buffer (B0201S, New England Biolabs) supplemented with 1 U/µl RNase H (Epicentre, Madison, WI) and 1 U/µl FastAP (EF0651, Thermo Fisher Scientific), followed by incubation at 75 °C for 5 min. 1 µl of radioactive labeling buffer (1× T4 PNK buffer, 2.5 U/µl T4 PNK, 4 µCi/µl [γ-^{32}P]ATP) was added to the RNA and incubated at 37 °C for 1 h and then at 75 °C for 5 min. The mixture was then annealed with 4 pmol splint oligo and 4 pmol 116-mer DNA oligo by heating at 75 °C for 3 min followed by addition of 2.9

µl ligation buffer (1.4× PNK buffer, 0.2 mM ATP, 57% DMSO, 5 U/µl T4 DNA ligase) and incubation for 3.5 h at 37 °C. The ligation was stopped by adding an equal volume of RNA urea loading buffer (8 M urea, 100 mM EDTA, 0.025% (w/v) bromphenol blue) and then was digested by incubated with 1 µl RNase T1/A mixture (EN0551, Thermo Fisher Scientific) at 37 °C for 16 h. The sample was loaded on 10% denaturing urea-PAGE to isolate the ligation product, which was desalted by ethanol precipitation. The pellet was resuspended in 9 µl sodium acetate/acetic acid buffer pH 4.6, supplemented with 2mM ZnCl2 and 5 units of nuclease P1, reaction was incubated) at 37 °C for 4 h and then was spotted on a TLC plate for separation. The result was visualized on a Typhoon FLA 9500 phosphor imager (GE Healthcare).

Biotinylation of 18S.1248.m^{1}acp^{3}Ψ modification

We have exploited reactivity of the m^{1}acp^{3}Ψ base with N-hydroxysuccinimide esters which was initially described in early 1970's (Gillam et al., 1968; Schofield et al., 1970; Friedman, 1972). Briefly, 10 µg of total RNA from HCT116 WT or *TSR3*[KO] in 47 µl of reaction buffer (100 mM phosphate buffer pH 8.0, 150 mM NaCl, 0.5 M EDTA pH 8.0) were denatured for 5 min at 70 ºC and placed on ice. After that 1µl of 100mM NHS-dPEG12-biotin (QBD10198-50MG, Sigma Aldrich) in DMSO was added. Reaction was incubated on ice for 2 h, every 40 minutes 1 µl of 100 mM of NHS-dPEG12-biotin was added (3 times in total). After treatment RNA was purified with Zymo-Spin IIC column, resolved on denaturing agarose gel and blotted on Hyperbond N+ (Amersham) membrane. RNA was UV crosslinked twice with 1200 µJ followed by incubation in blocking solution (phosphate buffered saline [PBS], pH 7.5, 10% SDS, 1 mM EDTA) for 20 min. The membrane was probed with 1:10000 dilution of 1 mg/ml streptavidin-horseradish peroxidase (Pierce) in blocking solution for 15 min, washed six times in PBS containing decreasing concentrations of SDS (10%, 1%, and 0.1% SDS, applied twice each) for 5 min and finally the biotin signal was visualized by ECL detection reagent (10 min exposure time).

RNA-seq

RNA-seq library preparation and sequencing was performed by the BC Cancer Genome Sciences Centre, Vancouver, Canada. Briefly, 75-bp stranded and paired-end poly-(A) RNA-seq libraries were prepared with NEBNext poly(A) mRNA magnetic isolation module (E7490L, New England BioLabs (NEB), Ipswich, MA), Maxima H minus First Strand cDNA synthesis kit (K1652, Thermo-Fisher), and NEBNext Ultra II directional RNA second strand synthesis (E7771, NEB). The total RNA-seq libraries

were prepared in parallel but without poly-(A) selection and only 2x PCR cycles (for adapter ligation). Libraries were sequenced on a HiSeq 2500 (Illumina, San Diego, CA).

Ribo-seq

Ribosome foot printing (ribo-seq) was performed as previously described (Ingolia et al., 2012) with minor modifications. For cell harvesting, the culture medium was aspirated, cells were washed twice with ice-cold PBS supplied with 100 µg/ml cycloheximide and plates were flash-frozen in liquid nitrogen. For cell lysis, the plates were placed on wet ice and 400 µl of mammalian polysome buffer (MPB) [20 mM Tris-HCl, pH 7.4, 150 mM NaCl, 5 mM $MgCl_2$, with 1 mM DTT and 100 µg/ml cycloheximide, 1% (vol/vol) Triton X-100, 25 U/ml Turbo DNase (AM2238, Invitrogen) was dripped onto the plates. Cells were scraped, the lysate was collected to fresh 1.5 ml tube, passed ten times through a 26-gauge needle, cleared by centrifugation at 20,000 x g for 10 min, flash-frozen in liquid nitrogen and stored at −80 °C until further use. For isolation of ribosome-protected RNA fragments, 240 µl of the lysate was digested with 6 µl of RNase I (AM2294, 100 U/µl, Invitrogen) at room temperature with rotation. After 45 min 8 µl of SUPERase-In (20 U/µl, AM2694, Invitrogen) was added to reaction and passed through MicroSpin S-400 HR columns (27-5140-01, GE Healthcare) equilibrated with mammalian polysome buffer. RNA was extracted from the flow-through using Trizol LS (10296-010, Invitrogen) followed by depletion of ribosomal RNA fragments with the RiboZero Kit (MRZH11124, Illumina). Ribosome-protected RNA fragments were loaded onto denaturing 17% urea-PAGE gel (EC-829, National Diagnostics) and gel area ranging from 27 nt to 30 nt, defined by corresponding RNA markers, was cut out. Purified RNA fragments were subjected to library generation using 3′ adapter 4N-RA3, 5′ adapter OR5-4N, RT primer RTP and PCR primers RP1 (forward primer) and RPI1-15 (reverse primers, containing barcodes). Libraries were sequenced on a HiSeq 4000 device (Illumina).

Polysome Fractionation

Polysome fractionation was performed as previously described (Floor and Doudna, 2016), with minor modifications. Media was removed from 100 mm dish with ~10^7 cells and washed with ice-cold ddH_2O containing 100 µM CHX. All subsequent steps were performed chilled at 4°C or on ice. After ddH_2O aspiration, cells were incubated for 30 min in 450 µL of hypotonic lysis buffer [0.1x polysome base buffer (PBB), 150 mM KCl, 20 mM Tris-HCl pH 7.4, 15 mM $MgCl_2$ in ddH_2O; with 1% Triton-X 100 and 1x protease inhibitor (4693132001, MilliporeSigma). After confirming >95% free nuclei with a

hemocytometer, nuclei were pelleted by centrifugation at 1,800 x g for 5 minutes. Cytoplasmic fraction was separated from mitochondria by centrifugation at 10,000 x g for 5 minutes. 300 µL cytoplasmic lysate was layered atop at 7-45% sucrose gradient (Gradient Master, BioComp, Fredericton, Canada) in 1x PBB. Gradients were ultra-centrifuged at 221,600 x g for 2 hours at 4°C (SW-41Ti rotor, 331362, Beckman Coulter, Brea, CA). Gradients were fractionated (Piston Fractionator, BioComp) into 20 x 300 µL fractions with in-line UV-scanning at 254 nm. Fractions were immediately frozen at -20°C for subsequent RNA extraction.

For comparative polysome profiles, the total input per gradient was normalized by taking the area under the curve (AUC) from the beginning of the SSU peak to the end of the last resolved polysome peak and normalizing the are to 100 artificial units. To normalize for sedimentary variation between gradients, the distance of the SSU, LSU, monosome, 2-, 3-, 4-some peaks across all gradients were averaged, and a simple linear regression was applied with the `*ln*` function in *R* (v 3.5.1) to regress peaks to a mean distance. Raw and processed data and analysis scripts are available in the online electronic notebook.

Ribosomal molecular dynamics simulations

All molecular dynamics (MD) simulations were performed as previously described (Girodat et al., 2019). In brief, 80S ribosome models were derived from available human Cryo-EM structures with a resolved 18S.1248.$m^1acp^3\Psi$ (PDB: 6EKO for E-site tRNA and 6OLE for A/P and P/E tRNA) (Li et al., 2019; Natchiar et al., 2017).

For simulations lacking complete $m^1acp^3\Psi$ modifications, the base was converted to each respective precursor. Each system was protonated with the *psfgen* package in VMD 1.9.3, and only e-nitrogen for histidine were protonated (Humphrey et al., 1996). Each system was solvated with a 10Å TIP3P water box with a concentration of 7mM $MgCl_2$ and 100mM KCl using the *solvate* and *autoionize* packages, respectively (Humphrey et al., 1996). All minimizations and MD simulations were performed with *NAMD* 2.1.2 using *CHARMM 36* standard parameters (Best et al., 2012; Denning et al., 2011; Phillips et al., 2005) and modified nucleic acid parameters from (Xu et al., 2016).

Each system underwent a steepest descent minimization of water for 10,000 steps then water and ions for 100,000 steps twice followed by minimization of nucleic acid and protein for 50,000 steps and finally the whole system for 100,000 steps. After minimization all systems were equilibrated to 300 and 350 K for 150 ps. Coordinates of the 350 K equilibration in conjunction with velocities from the 300 K

equilibration were used as initial parameters for the MD simulation. Each system was simulated for ~20ns. RMSD, RMSF, dihedral angel, and distance measurements were performed with VMD 1.9.3 (Humphrey et al., 1996).

Tandem Mass Tag labeling and liquid-chromatography and tandem mass spectrometry (TMT LC-MS/MS)

HCT116 [WT] and TSR3[KO/Het] cells were grown to 80% confluence in a 100mm dish, washed in PBS, detached, and washed twice in PBS. Cell suspensions were aliquot into technical triplicates and pelleted before snap freezing. One of each HCT116 [WT] and duplicates of HCT116 *TSR3*[KO/Het] clones were processed for LC-MS/MS. Pellets were lysed in 200 µL lysis buffer [4M Guanidine HCl (G4505-500G, Sigma), 40 mM 2-chloroacetamide (22790, Sigma), 10mM tris(2-carboxyethyl) phosphine (C4706-2G, Sigma), 5 mM Ethylenediaminetetraacetic acid (AM9260G, Invitrogen), 1x phosSTOP inhibitor (4906845001, Sigma), 1x cOmplete protease inhibitor cocktail, EDTA free (4693132001, Sigma), 50 mM HEPES (H3784-100G, Sigma)].

The cells were lysed by bead beating using Lysing D matrix tubes (116913100, Cedarlane Laboratories) in a FastPrep 24 instrument (116005500, MP Biomedical) at 8 M/s for 30 seconds, resting 90 seconds, and a second round of bead beating at 8 M/s for 30 seconds. Supernatant was transferred into 1.5 mL lo-bind tubes (022363204, Eppendorf) Protein were solubilized and denatured by heating at 95 °C for 15 minutes with shaking at 1000 rpm in a thermomixer (05-400-205, Fisher Scientific).

All experiments used a 1:1 combination of two different types of carboxylate-functionalized beads, both with a hydrophilic surface (45152105050350 and 65152105050350, Sera-Mag Speed beads, GE Life Sciences).The magnetic particles are an average diameter of 1µm. The beads were stored at 4 °C at a concentration of 10 µg/µL. Magnetic racks used in all experiments were designed and manufactured in-house as previously described (Hughes et al., 2014).

The Sera-Mag SP3 1:1 bead mix was rinsed once with 200 uL HPLC water (270733, Sigma,) and diluted to a final working concentration of 20 µg/µL. To each sample, 200µg of bead mixture was added and mixed by pipetting to generate a homogeneous solution. To induce protein binding to the beads, ethanol (34852, Sigma) was added to achieve a final concentration of 50% (v/v). Bead-protein solutions were mixed to ensure a homogeneous distribution of the beads and incubated for 5 minutes at 24 °C with 1000 rpm shaking in a thermomixer. After incubation, tubes were placed on a magnetic rack

for 2 minutes then the supernatant was removed and discarded. The beads were rinsed three times with 200 µL of freshly prepared 80% ethanol in water, and the supernatant was discarded each time. Rinsed beads were reconstituted in aqueous buffer (100µL, 50 mM HEPES pH 8.0) containing a 1:50 (µg:µg) enzyme to protein amount of trypsin/LysC mix (V5071, Promega). Mixtures were incubated for 16 hours at 37°C with 1000 rpm shaking in a thermomixer. The supernatants were recovered using a magnetic rack and transferred to fresh 1.5mL lo-bind Eppendorf tubes.

Tandem Mass Tag (TMT) labels (5 mg per vial, 90406, Pierce) were reconstituted in 500 µLof HPLC acetonitrile (34851-4L, Sigma) and frozen at -80 °C. Prior to labeling, TMT labels were removed from the freezer and allowed to equilibrate at room temperature. A standard 13 cell line “supermix” digest mix was prepared as described above and included as an internal standard. Labeling was performed by addition of TMT label in two volumetrically equal steps of to achieve a 2:1 (µg:µg) TMT label to peptide final concentration with 30 minutes incubation at room temperature for each labeling step. Reactions were quenched by addition of an equal volume of 1 M glycine (G8898, Sigma) in water. Labeled peptides were concentrated in a SpeedVac centrifuge (781001010234 + RVT400-115, Labconco) to remove acetonitrile, combined, acidified to a concentration of 1% trifluoroacetic acid (T6508-100mL, Sigma,) and peptides were purified using a C18 SepPak (WAT054960, Waters). The SepPak was conditioned with 2 mL 0.1% trifluoroacetic acid in HPLC grade acetonitrile and equilibrated with 2 mL of 0.1% trifluoroacetic acid in HPLC water. The sample was loaded on the SepPak and rinsed with 3 mL of 0.1% formic acid (, 56302-50ML-F, Sigma) in HPLC water. Peptides were eluted using two 600 uL aliquots of 80% acetonitrile in HPLC water with 0.1% formic acid.

High-pH reversed phase analysis was performed on an Agilent 1100 HPLC system (G1364C, G1315A, G1313A, G1311A). Fractionation was performed on a Kinetix XB C18 column (2.1 x 150 mm, 1.7 µm core shell, 100 Å, Phenomenex, 00F-4498-AN). Elution was performed with a gradient of mobile phase A (water and 0.1% formic acid) to 7% B (acetonitrile and 0.1% formic acid) over 2 minutes, to 25% B over 94 minutes, to 40% over 17 minutes, with final elution (80% B) and equilibration (5% B) using a further 7 minutes all at a flow rate of 450 nL/min. Fractions were collected every minute across the elution window for a total of 48 fractions, which were concatenated to a final set of 12 (e.g. 1 + 13 + 25 + 37 = fraction 1). Fractions were dried in a SpeedVac centrifuge and reconstituted in 0.1% formic prior to MS analysis.

Analysis of TMT labeled peptide fractions was carried out on an Orbitrap Fusion Tribrid MS (IQLAAEGAAPFADBMBCX, Thermo Scientific). Samples were introduced using an Easy-nLC 1000 system (LC120, Thermo Scientific). Trapping columns were packed in 100 µm internal diameter, 360 um outer diameter fused silica capillary (1068150023, Molex) to a length of 25 mm with C18 beads (Reprosil-Pur, Dr. Maisch GmbH, 3 µm particle size, r13.b9.0001). The analytical column was packed with C18 (Reprosil-Pur, Dr. Maisch, 3 µm particle size) to a length of 15 cm in a 100 µm 100 µm internal diameter, 360 um outer diameter fused silica capillary with a laser-pulled electrospray tip. Trapping was carried out for a total volume of 10 µL at a pressure of 400 bar. After trapping, gradient elution of peptides was performed on and heated to 50 °C using AgileSLEEVE column ovens (AS1032 + AS1532H, Analytical Sales & Service). Elution was performed with a gradient of mobile phase A (water and 0.1% formic acid) to 8% B (acetonitrile and 0.1% formic acid) over 5 minutes, to 25% B over 88 minutes, to 40% over 20 minutes, with final elution (80% B) and equilibration (5% B) using a further 7 minutes at a flow rate of 375 nL/min.

Data acquisition on the Orbitrap Fusion (control software version 3.1.2412.17) was carried out using a data-dependent method in positive ion mode. Survey scans covering the mass range of 400 – 1200 were acquired in profile mode at a resolution of 120,000 (at *m/z* 200), with quadrupole isolation enabled, an S-Lens RF Level of 60%, a maximum fill time of 120 ms, an automatic gain control (AGC) target value of 4e5, and one microscan. For MS2 scan triggering, monoisotopic precursor selection was enabled for peptides, charge state filtering was limited to 2 – 4, undetermined charge states were included, and dynamic exclusion of masses selected one time was enabled for 15 seconds with a tolerance of 20 ppm. HCD fragmentation with a maximum fill time of 20 milliseconds, quadrupole isolation, an isolation window of 1.4 *m/z*, isolation offset off, a fixed collision energy of 40%, injection for all available parallelizable time turned OFF, and an AGC target value of 1.2e5 was performed. The Orbitrap was used as the detector in normal scan range mode with a resolution of 50,000 at 200 *m/z* and the first mass was 120 m/z. Centroided data was collected with one microscan. The total allowable cycle time was set to 3 seconds.

LC-MS/MS data analysis

Peptide detection and protein inference was performed using *Proteome Discoverer* (v 2.4). Searches were performed using *Sequest HT* and the human proteome from Uniprot plus the CRAP-ome. Tryptic cleavages with 2-6 missed cleavages and peptide length of 6-144 was used. The precursor mass

tolerance was 10 ppm, fragment mass tolerance was 0.05 Da and monoisotopic masses were used. Only b- and y-type ions were included in the search. Oxidation was allowed as a dynamic modification and the protein terminus was allowed to have acetylation, Met-loss, or Met-loss+acetylation as dynamic modifications. TMT reagents (+229.163 Da) were set as static modifications on the peptide terminus and lysine and carbamidomethyl was assigned as a static modification of cystine. Percolator was used for scoring with an FDR of 0.01, concatenated Target/Decoy Selection, and validation based on the q-Value. The maximum Delta Cn was 0.05 and the maximum rank was 0. Unique + Razor peptides were used for quantification. *R* (v 3.5.1) was used for data analysis. Peptides were filtered for uniqueness and values were median normalized, and proteins containing at least two unique peptides were included in downstream differential protein expression analysis. The technically replicated samples were merged (mean) and statistical protein expression analysis was performed in *R.*

Statistics

Statistical analysis was performed in *R* (v 3.5.1). Differences in variant allele frequency (VAF) between tumor and normal patient samples was two-tailed, paired Student's T-test with degrees of freedom one less than reported *n.* Bonferonni multiple-testing correction was applied when screening for changes across 18S and 28S nucleotides. Error bars on boxplots are quantiles. Differential gene expression and translation was tested with *DEseq2* (Love et al., 2014) with Benjamini-Hochberg multiple testing correction at an alpha of 0.05. Multi-group comparisons between HCT116 WT[1],[2],[3] and *TSR3*[KO 1],[KO 3],[Het2] ribo-seq were performed with one-way ANNOVA, followed by Tukey's Honestly Significant Difference (HSD) test if indicated. Genotypic differences between biological triplicates of WT and *TSR3*[KO] cells were tested with Student's T-test with 2 degrees of freedom and Benjamini-Hochberg multiple testing correction at an alpha of 0.05 where indicated.