Article

RNA codon expansion via programmable pseudouridine editing and decoding

https://doi.org/10.1038/s41586-025-09165-x

Received: 29 April 2024

Accepted: 15 May 2025

Published online: 25 June 2025

Check for updates

Jiangle Liu[1,2,3,11], Xueqing Yan[1,11], Hao Wu[1,3,11], Ziqin Ji[2], Ye Shan[2,3], Xinyan Wang[4], Yunfan Ran[5], Yichen Ma[1,3], Caitao Li[6], Yuchao Zhu[2,7], Ruichu Gu[1,8], Han Wen[8,9,10], Chengqi Yi[1,2,3,9] ✉ & Peng R. Chen[2,3,7] ✉

The incorporation of non-canonical amino acids (ncAAs) enables customized chemistry to tailor protein functions[1–3]. Genetic code expansion offers a general approach for ncAA encoding by reassigning stop codons as the 'blank' codon; however, it is not completely orthogonal to translation termination for cellular transcripts. Here, to generate more bona fide blank codons, we developed an RNA codon-expansion (RCE) strategy that introduces and decodes bioorthogonally assignable pseudouridine (Ψ) codons (ΨGA, ΨAA or ΨAG) on specified mRNA transcripts to incorporate ncAAs in mammalian cells. The RCE strategy comprises a programmable guide RNA[4], an engineered decoder tRNA, and aminoacyl-tRNA synthetase. We first developed the RCE(ΨGA) system, which incorporates functional ncAAs into proteins via the ΨGA codon, demonstrating a higher translatome-wide and proteomic specificity compared with the genetic code expansion system. We further expanded our strategy to produce the RCE(ΨAA) and RCE(ΨAG) systems, with all three Ψ codon:(Ψ codon)-tRNA^{Pyl} pairs exhibiting mutual orthogonality. Moreover, we demonstrated that the RCE system cooperates compatibly with the genetic code expansion strategy for dual ncAA encoding. In sum, the RCE method utilized Ψ as a post-transcriptional 'letter' to encode and decode RNA codons in specific mRNA transcripts, opening a new route for genetic alphabet expansion and site-specific ncAA incorporation in eukaryotic cells.

The rapidly growing repertoire of ncAAs represents a broad range of chemical structures and functions, whose incorporation enables the probing, dissection and modulation on various proteins of interest in living systems[1–3]. The genetic code expansion (GCE) strategy has been developed to reassign stop codons as a genetic blank codon[1–3] in order to specifically incorporate ncAAs into proteins. However, reassigning stop codons is not completely orthogonal to the translation termination of cellular transcripts, because this may result in off-target readthrough for endogenous stop codons in the translatome. Alternative strategies have been devised to create genetic blank codons in *Escherichia coli*, including: (1) orthogonal ribosome evolution for quadruplet codon recognition[5–7]; (2) genome engineering[8–11] to create genomically recoded organisms such as syn61 (ref. 12), *rE.coli-57* (refs. 10,13) and Ochre[14]; and (3) the development of unnatural base pairs with hydrophobic interactions[15,16]. However, transplanting these genetic blank codons into mammalian cells encounters challenges such as potential cytotoxicity, global translation byproducts and the complex eukaryotic genome[17]. Instead, two strategies have emerged to exploit translation processes. An artificial, membraneless organelle was created in mammalian cells to achieve orthogonal translation for the UAG stop codon[18,19]. In addition, RNA base editors have converted sense codons into stop codons[20]. Nevertheless, GCE utilizes genetic blank codons transcribed from DNA to RNA, which have been limited to stop codons in mammalian cells, and these interfere with translation termination of cellular transcripts.

Alongside the central dogma, the genetic information contained in RNA codons in the GCE strategy remains the same as in upstream DNA codons. However, given that diverse and abundant RNA modifications are independent of DNA codons and can directly influence the decoding process[21–23], we postulate that it is possible to create RNA-based blank codons outside of the canonical 64 genetic codons. Therefore, we proposed that generating the post-transcriptional letter as RNA-based blank codons might help to overcome the issues associated with leveraging existing genetic blank codons in the translatome. This could be realized by programming specific encoding and decoding processes of a targeted codon of interest in specified mRNA transcripts. For instance, as a representative RNA modification, programmable installation of pseudouridine (Ψ) on specified mRNAs via guide small nucleolar RNAs (gsnoRNAs)[4,24,25] could generate Ψ-modified nonsense codons (Ψ codons) with enhanced, near-cognate decoding, probably owing to Ψ-promoted base pairing, base stacking and strengthening sugar-phosphate backbone[26]. However, there is a current lack of

[1]The National Key Laboratory of Gene Function Studies and Manipulation, School of Life Sciences, Peking University, Beijing, China. [2]Synthetic and Functional Biomolecules Center, Beijing National Laboratory for Molecular Sciences, College of Chemistry and Molecular Engineering, Peking University, Beijing, China. [3]Peking-Tsinghua Center for Life Sciences, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing, China. [4]DP Technology, Beijing, China. [5]School of Medicine, Shanghai Jiao Tong University, Shanghai, China. [6]School of Life Science and Technology, ShanghaiTech University, Shanghai, China. [7]Key Laboratory of Bioorganic Chemistry and Molecular Engineering of Ministry of Education, Peking University, Beijing, China. [8]Institute for Advanced Algorithms Research, Shanghai, China. [9]Beijing Advanced Center of RNA Biology (BEACON), Peking University, Beijing, China. [10]AI for Science Institute, Beijing, China. [11]These authors contributed equally: Jiangle Liu, Xueqing Yan, Hao Wu. ✉e-mail: chengqi.yi@pku.edu.cn; pengchen@pku.edu.cn

specific decoder tRNAs that are capable of recognizing modified RNA codons from the 64 genetic codons, posing challenges for harnessing the modified RNA codons as orthogonal blank codons to expand the genetic alphabet.

Here we developed an RCE approach to assign modified RNA codons as new blank codons for ncAA incorporation in mammalian cells, which are independent of the endogenous codon assignment in the translatome. To decode the three generated Ψ codons (ΨGA, ΨAA and ΨAG) on specified mRNA transcripts, we generated three specific decoder tRNAs to three Ψ codons: (ΨGA)-tRNA^{Pyl}, (ΨAA)-tRNA^{Pyl} and (ΨAG)-tRNA^{Pyl}. These Ψ codon-tRNAs exhibited a robust preference for the corresponding Ψ codons compared with endogenous codons, which were retained in the translatome. RCE(ΨGA) achieved high specificity for ncAA incorporation in ribosome profiling and proteome analysis, preserving the UGA codon which represents around 52% of stop codons in the human genome. We demonstrated that the three pairs—ΨGA:(ΨGA)-tRNA^{Pyl}, ΨAA:(ΨAA)-tRNA^{Pyl} and ΨAG:(ΨAG)-tRNA^{Pyl}—are orthogonal to each other, enabling the specific incorporation of ncAAs carrying different side chains into mammalian proteins. Moreover, our RCE approach cooperated with the current GCE system, confirming the compatibility of different codon-expansion strategies for encoding dual ncAAs within cells. Overall, the RCE method enabled programmable encoding and decoding of modified RNA codons via independent Ψ codon:(Ψ codon)-tRNA^{Pyl} pairs, significantly improving the translatome-wide specificity of ncAA incorporation, allowing for the precise investigation and modulation of proteins of interest under various in vivo settings.

Rationale of RCE

RNA modifications are well-studied for their dynamic influences on the decoding process of the targeted codon during translation[21], independent of the central dogma. To harness RNA modifications as post-transcriptional letters and expand the genetic alphabet, we programmed the RCE using a two-step process: encoding and decoding (Fig. 1a). For encoding using RNA modifications, we utilized the programmable pseudouridylation tool RESTART[4] (RNA editing to specific transcripts for pseudouridine-mediated premature termination codon readthrough) to produce site-specific Ψ codons, converting the targeted uridine (U) to Ψ on specific mRNA transcripts. For decoding, we identified specific decoder tRNAs for the modified RNA codons by screening the wild-type and engineered tRNA^{Pyl} constructs. Therefore, using programmable encoding and decoding processes, we were able to assign the modified RNA Ψ codons as new blank codons for ncAA incorporation that operate orthogonally to canonical codon assignment for protein biogenesis throughout the translatome.

Programmable encoding of the ΨGA codon

To implement the RCE approach, we first ensured that the Ψ codon installation method was robust. As the proposed RCE approach includes ternary components containing the RESTART tool for Ψ installation, a non-canonical aminoacyl-tRNA synthetase (RS) and a decoder tRNA for Ψ codon decoding, we evaluated the efficiency and robustness of programmable and targeted RNA pseudouridylation as the initial step to create the Ψ codon[4,24] for the RCE system.

We constructed several dual-fluorescent mCherry–linker–GFP reporters, with a target UGA codon in the linker sequence (Supplementary Table 1). We then designed guide RNAs targeting the UGA codon, following the principle of RESTART[4] for Ψ installation. To quantitatively determine Ψ ratios, we used PRAISE, a Ψ detection method that utilizes sequencing[27] (Supplementary Fig. 2). We found high Ψ ratios for the targeted UGA with different flanking sequences in the reporters, ranging from 55% to 96% (average ratio approximately 71%). We also tested different stoichiometry between the gsnoRNAs and the corresponding reporters (such as 1:4, 1:2 and 1:1), and observed consistently high Ψ installation efficiency (Fig. 1b,c and Extended Data Fig. 1a–c). After we tested a series of tRNA dosages, we found that the tRNA dosage did not influence Ψ ratio (Fig. 1d,e and Extended Data Fig. 1d). Together, these results demonstrated that our designed gsnoRNAs operated robustly and efficiently for Ψ codon encoding during the RCE process.

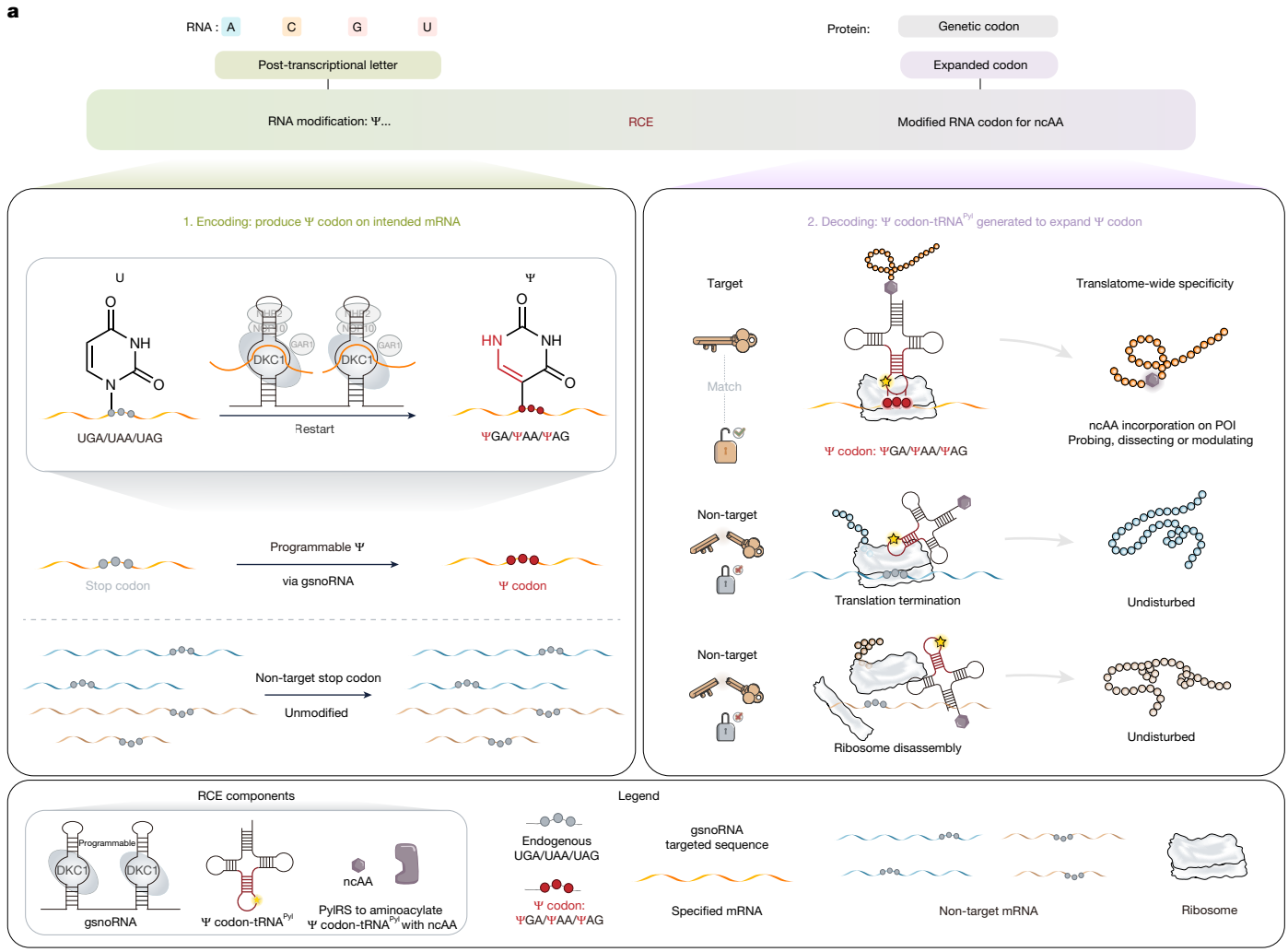
To further verify the encoding specificity, we determined the transcriptome-wide off-target Ψ sites[27] (Supplementary Table 2 and Supplementary Fig. 3). Because the guide RNAs were carefully designed to minimize off-targets, we identified only 9 off-target sites with a relatively low Ψ ratio (ranging from 10.7% to 25.2%) as well as a consensus sequence motif resembling the on-target sequence (Extended Data Fig. 2a–f). Of note, none of these off-target sites resides at endogenous stop codons (Extended Data Fig. 2g). In addition, we showed that these off-target Ψ sites had negligible impact on RNA expression and ribosome footprints (Extended Data Fig. 2h,i). In sum, we conclude that our programmable and targeted RNA pseudouridylation is efficient and specific for creating Ψ codons.

Generating a decoder tRNA for ΨGA

To make ΨGA a bioorthogonally assignable codon for reprogramming, a decoder tRNA is necessary to efficiently suppress ΨGA but not the UGA codon with customized ncAAs. This is difficult, not only because UGA represents the most prevalent stop codon (approximately 52% of stop codons) in the human genome, but also because UGA is naturally prone to readthrough among the 3 stop codons[28]. Initially, we evaluated whether wild-type tRNA^{Pyl}(UCA) could distinguish ΨGA from the cognate UGA codon. We selected three representative wild-type tRNA^{Pyl} molecules from distinct species: *Methanosarcina mazei* (*Mm*) tRNA^{Pyl}(UCA), *Methanomethylophilus alvus* (*Ma*) tRNA^{Pyl}(UCA) from class A $^{\Delta NPyl}$tRNAs, and *Methanomassiliicoccus intestinalis* (*Mi*) tRNA^{Pyl}(UCA) from class B $^{\Delta NPyl}$tRNAs[29]. Notably, minimal readthrough differences between ΨGA and UGA were observed for the three wild-type tRNA^{Pyl} decoder tRNAs.

To produce a decoder tRNA with a high ΨGA preference, we focused on the anticodon stem-loop (ASL) of tRNA and attempted to identify possible variant tRNAs with desired properties, relying on ASL mutations to influence stability, flexibility and proximal interactions of the tRNA^{Pyl} decoder tRNAs during decoding in the small subunit of the ribosome[30], thereby potentially swaying the decoding preference toward ΨGA. It is equally important that the ASL mutations do not affect the interaction of PylRS with tRNA^{Pyl} (refs. 31,32), so that we can generate tRNA^{Pyl} variants without affecting its aminoacylation activity.

To assess the possibilities, we screened 150 tRNA^{Pyl} variants for all combinations of single-nucleotide mutations within the ASL region of the three representative wild-type tRNA^{Pyl}(UCA) decoder tRNAs (Fig. 2a). We used the dual-fluorescent reporter Screen-TGA to measure the readthrough efficiency of tRNA variants (Supplementary Fig. 4a), and the fold change between the readthrough ratios of ΨGA and UGA as a measure for the ΨGA preference. We determined that more than half of the mutations at the ASL produced at least 1.5-fold ΨGA preference, with mutations at position 37 generally producing a high ΨGA preference (greater than twofold) (Fig. 2b and Supplementary Table 3). Screening these tRNA variants, we determined that although the exact identity of the 37th nucleotide had to be experimentally tested, position 37 improved the ΨGA preference in general (Fig. 2b). This finding mirrors previous structural insights, which revealed that nucleotide 37 consistently stacks between anticodon nucleotides, contributing to overall ASL stability and flexibility[30]. Therefore, from a series of tRNA variants responding to ΨGA, we selected *Mm*tRNA^{Pyl}(UCA)-37G owing to its higher specificity for the ΨGA codon as well as its efficiency for incorporating various ncAAs, and we named this decoder tRNA (ΨGA)-tRNA^{Pyl}.



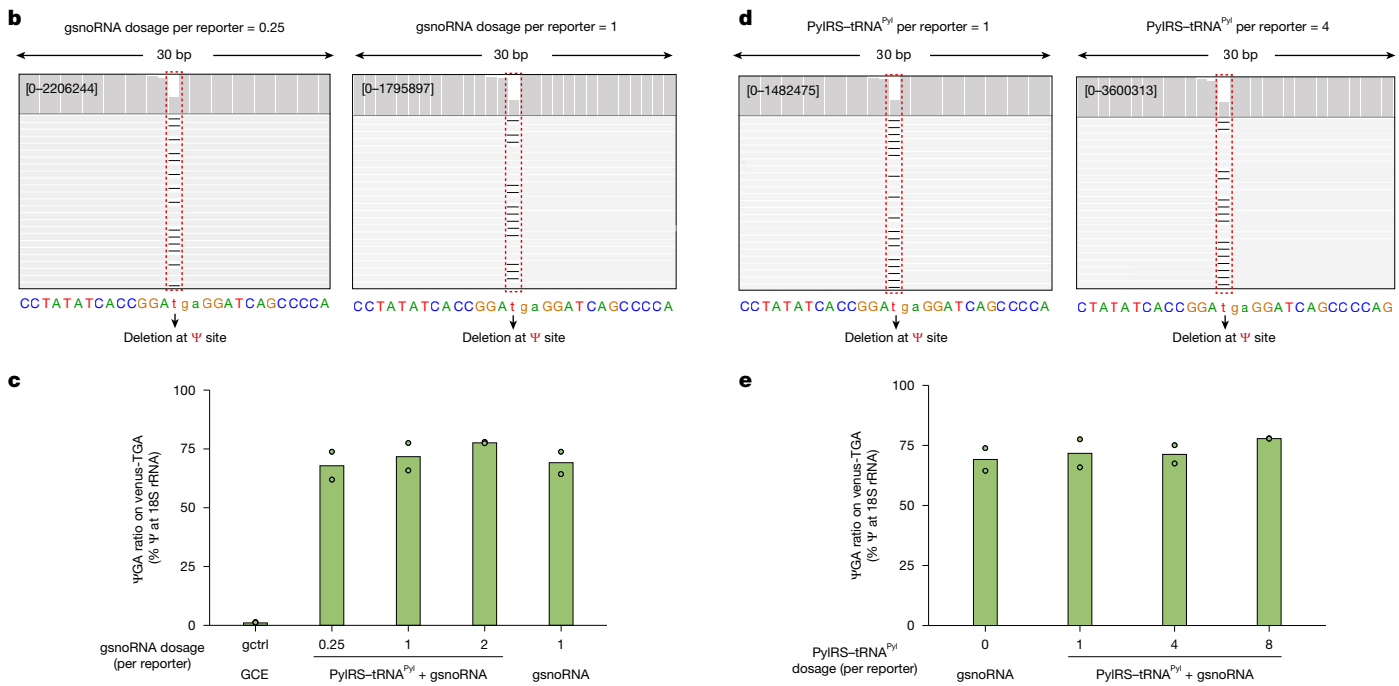


Fig. 1 | See next page for caption.

Fig. 1 | Schematic overview of the RCE strategy and yields of encoded ΨGA codon in the specified mRNA transcript. a, The RCE strategy contains the encoding and decoding processes for a modified RNA codon. For encoding, the RESTART system was used, including a programmable gsnoRNA and DKC1. A gsnoRNA was constructed to target the specified mRNA transcript via base pairing, interacting with the DKC1 and the endogenous proteins to pseudouridylate the U of the intended codon, thus sequence-specifically producing Ψ codon (ΨAA/ΨAG/ΨGA) from stop codons (UAA/UAG/UGA). For decoding, (Ψ codon)-tRNA^{Pyl} decoder tRNAs were generated from tRNA^{Pyl} mutants with a single-nucleotide mutation in the red region for the installed Ψ codons, whereas the endogenous stop codons remained undisturbed. Thus, the RCE system could discriminate the Ψ-modified RNA codons (Ψ codons) from usual stop codons. The hexagon represents a ncAA, the key represents the decoder tRNA, the lock represents the Ψ codon and the yellow star represents the mutation in decoder tRNA. POI, protein of interest. **b**,**c**, Representative IGV views of the installed Ψ at targeted U sites (**b**) and corresponding bar plots of ΨGA codon yields (**c**) in target mRNA under a 1:4 or 1:1 stoichiometric ratio of gsnoRNA-TGA to the targeted Venus-TGA reporter. Following selective chemical labelling of Ψ with the PRAISE method, Ψ sites were identified as deletion signals during sequencing, and normalized to the deletion ratio of Ψ sites (Ψ1347 and Ψ1367) in 18S ribosomal RNA (rRNA) for ΨGA codon yields. Data are shown as mean values (n = 2 independent experiments). gctrl, control non-targeting guide RNA. **d**,**e**, Representative IGV views (**d**) and corresponding ΨGA codon yields (**e**) in intended mRNA with different dosages of RS–tRNA relative to the targeted reporter. Data are shown as mean values (n = 2 biologically independent replicates).

Next, we validated that (ΨGA)-tRNA^{Pyl} exhibited a robust preference for the ΨGA codon on the dual-fluorescence *LDLR*, *AGXT* and *CFTR*-TGA reporters (Supplementary Table 1), whose ΨGA codons reside in different sequence contexts adapted from approximately 200-nt segments of the *CFTR*, *AGXT* and *LDLR* transcripts surrounding the annotated premature termination codons (PTC) in Clinvar[33] (Supplementary Fig. 5). With a 45.1–99.1% yield for the ΨGA codons (Supplementary Fig. 6a), we found that (ΨGA)-tRNA^{Pyl} demonstrated a 2.7- to 11.7-fold ΨGA preference and adequate readthrough efficiencies (approximately 19–21%) in the presence of ncAA-CbzK (carbobenzyloxy-L-lysine) (Fig. 2c–f and Supplementary Fig. 6b,c). Besides HEK293T cells, we also confirmed evident ΨGA preference of the RCE(ΨGA) system to four additional cell lines (Supplementary Fig. 7). Finally, we confirmed that (ΨGA)-tRNA^{Pyl} decoded the ΨGA codon specifically with ncAA but not natural amino acids by identifying readthrough product using liquid chromatography–mass spectrometry (LC–MS) and liquid chromatography–tandem mass spectrometry (LC–MS/MS), and the tRNA charging with acid-denaturing gel northern blots (Extended Data Fig. 3 and Supplementary Fig. 8).

We also inspected the structural basis for the observed ΨGA codon preference. We performed all-atom molecular dynamics simulation based on cryo-electron microscopy (cryo-EM) structures of ribosome, mRNA and tRNA (Extended Data Fig. 4a–c). By extracting the tRNA ASL and mRNA for restrained simulations, we investigated the codon–anticodon pairing probability for nine selected tRNAs with and without Ψ codon preference. We found enhanced pairing probabilities for ΨGA compared with UGA for tRNAs with ΨGA preference (Extended Data Fig. 4d). For instance, *Mm*tRNA^{Pyl}(UCA)-37G formed more stable hydrogen bonds with the ΨGA codon than with the UGA codon (Extended Data Fig. 4e–h), indicating that the strengthened pairing with the ΨGA codon may account for the codon preference. Overall, we successfully established a decoder tRNA, (ΨGA)-tRNA^{Pyl}, that preferentially decoded the ΨGA codon, and thus distinguished the ΨGA codon from endogenous allocations of the 64 genetic codons.

Translatome-wide decoding specificity

We next evaluated the decoding specificity of the RCE(ΨGA) system, which consists of the ΨGA codon and (ΨGA)-tRNA^{Pyl}, by evaluating the potential off-target readthrough events on the UGA codon (which accounts for around 52% of stop codons) across the translatome. Ribosome profiling has been widely utilized to assess global off-target activities of suppressor tRNAs[34], as well as to examine the relevant pathologies resulting from upregulated stop codon readthrough, involving protein mislocalization[35], aggregation[36], instability[37] and the cascades of nontrivial magnitude[38,39]. It should also be suitable for the evaluation of the translatome-wide specificity of the RCE system alongside the GCE system. By design, the RCE(ΨGA) system should incorporate a ncAA at the produced ΨGA codon on specified mRNAs, while leaving the endogenous stop codons less disturbed compared to the GCE(UGA) system. Therefore, we performed ribosome profiling experiments for the RCE(ΨGA) system to measure its off-target activities, which directly reflect interferences on translation termination within cells.

To provide a holistic perspective on stop codon readthrough ratios in the translatome, we calculated the ribosome readthrough score (RRTS)[34,40] values for all transcripts with identified ribosome footprints on the 3′ untranslated region (UTR). As expected, the RCE(ΨGA) and GCE(UGA) systems did not induce off-target readthrough on UAA and UAG codons (Supplementary Fig. 9). For the UGA codon of the intended transcript, RCE(ΨGA) and GCE(UGA) systems exhibited similar on-target RRTS values of 0.366 and 0.375 using the intended UGA as the stop codon (Supplementary Fig. 10). For the UGA codons of globally non-targeting transcripts, the median RRTS value in GCE(UGA) cells, RCE(ΨGA) cells and control cells was found to be 0.057, 0.014 and 0, respectively (Fig. 3a). These findings suggest that RCE exhibited a median fourfold reduction in global off-target readthrough ratios and a similar on-target readthrough efficiency to GCE(UGA) cells, demonstrating a median fourfold ncAA incorporation specificity for the intended transcript in the translatome.

To further investigate the off-target transcripts, we used the RRTS value of a transcript in control cells as background, and designated the transcript as off-target if the transcript possessed obvious RRTS fold changes in RCE(ΨGA) or GCE(UGA) cells. Regardless of the RRTS fold-change cut-off that we used, we consistently identified lower off-target transcripts for RCE(ΨGA) cells than for GCE(UGA) cells (Fig. 3b). For instance, using a strict cut-off of a tenfold change, we observed that the number of off-target transcripts significantly decreased from 63 in GCE(UGA) to 34 in RCE(ΨGA) (Fig. 3b). We also evaluated the gene ontology (GO) terms of off-target transcripts, and determined that GCE(UGA) led to 16 enriched pathways. Notably, the top enriched term appears to be related to ribosome biogenesis. Off-target transcripts in RCE(ΨGA) were enriched across only four pathways (Fig. 3c), demonstrating substantially reduced enrichment in ribosome biogenesis and no enrichment in DNA damage and repair.

Given that natural stop codon readthrough events are associated with diverse pathologies, we next examined the RRTS of reported transcripts with naturally occurring readthrough events at the UGA codon. After excluding transcripts with low footprint coverage, we identified two representative transcripts, *MDH1* (ref. 28) and *THG1L* (ref. 41), which exhibited distinct readthrough in control cells. In RCE(ΨGA) cells, we identified numbers of readthrough events that resembled those in control cells, suggesting a minimal impact of the RCE system on these transcripts. In GCE(UGA) cells, the readthrough activities were significantly enhanced. For *MDH1*, the RRTS in control cells, RCE(ΨGA) cells and GCE(UGA) cells was 0.021, 0.033 and 0.169, respectively. Thus the off-target readthrough ratio on *MDH1* of RCE(ΨGA) exhibited a 12-fold reduction compared with the GCE(UGA) system, with a distinct increase in RRTS of 0.148 induced by the GCE system (Fig. 3d).

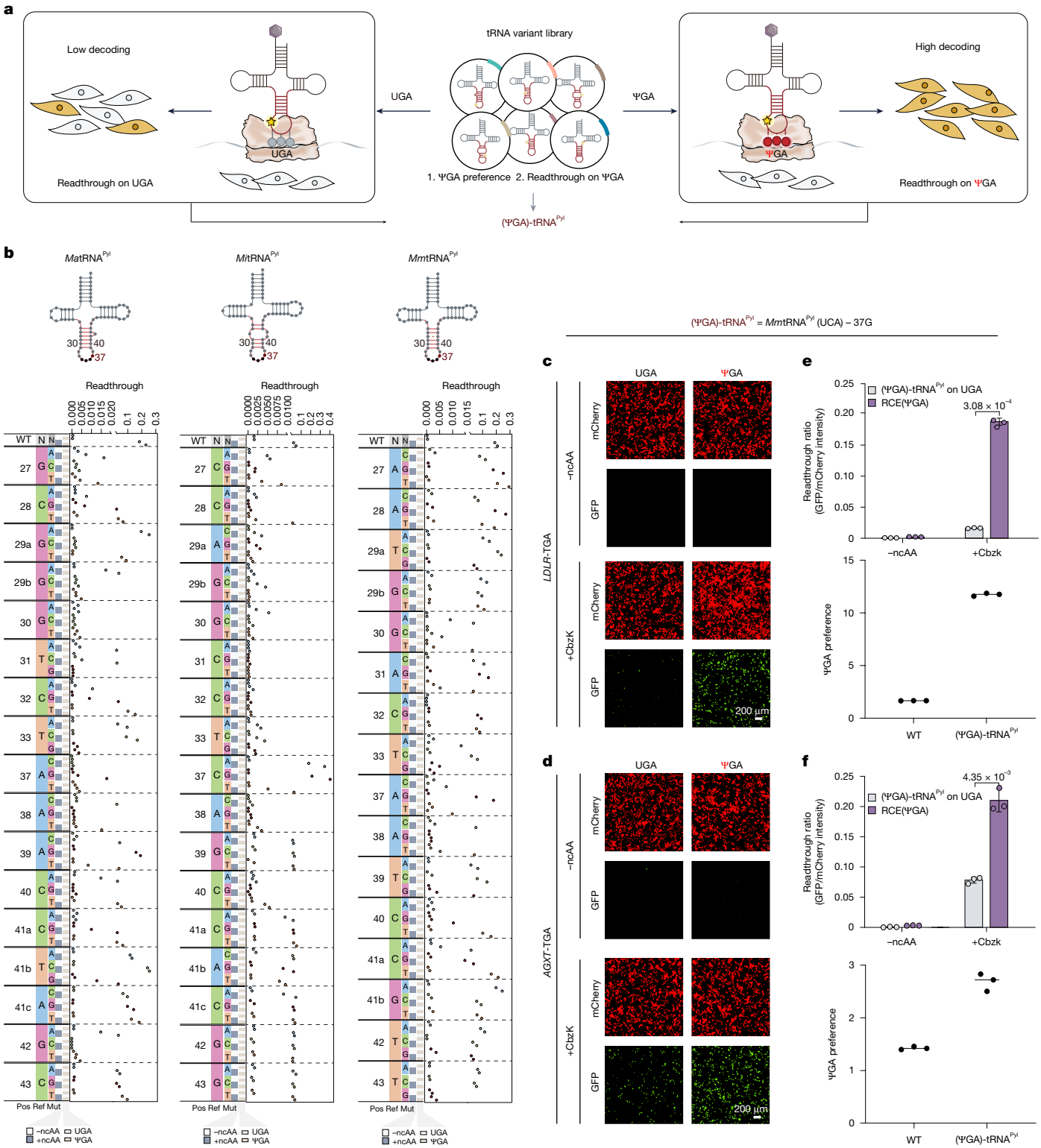


Fig. 2 | Screening and evaluation of the specific and efficient decoder tRNA for the ΨGA codon over the UGA codon. a, Scheme demonstrating the screening strategy of decoder tRNA for the ΨGA codon. Each tRNA variant was assessed with readthrough efficiencies in both UGA and ΨGA conditions. **b**, Top, secondary structures of *Mm*tRNA^{Pyl}(UCA), *Mat*RNA^{Pyl}(UCA) and *Mit*RNA^{Pyl}(UCA), with mutated regions highlighted in red. Bottom, dot plot showing readthrough ratios on the UGA and ΨGA codons of tRNA variants bearing single-nucleotide mutations. Wild-type (WT) nucleotides, mutated nucleotides and position are shown in the Ref, Mut and Pos columns, respectively. Experiments were performed with or without ncAA. *y* axes showing readthrough ratios are truncated to improve visualization. **c**,**d**, Representative fluorescence images showing readthrough of (ΨGA)-tRNA^{Pyl} on UGA and ΨGA codons in the *LDLR*-TGA (**c**) and *AGXT*-TGA (**d**) reporters in the absence and presence of CbzK. Experiments were repeated independently three times. **e**,**f**, Readthrough ratios of (ΨGA)-tRNA^{Pyl} on UGA and ΨGA codons in the *LDLR*-TGA (**e**) and *AGXT*-TGA (**f**) reporters in the absence and presence of CbzK. Data are mean ± s.d. ($n = 3$ biologically independent replicates). Readthrough ratios were normalized using read reporters, which do not contain stop codons in the coding sequences. *P* values were calculated by a two-sided Student's *t*-test.

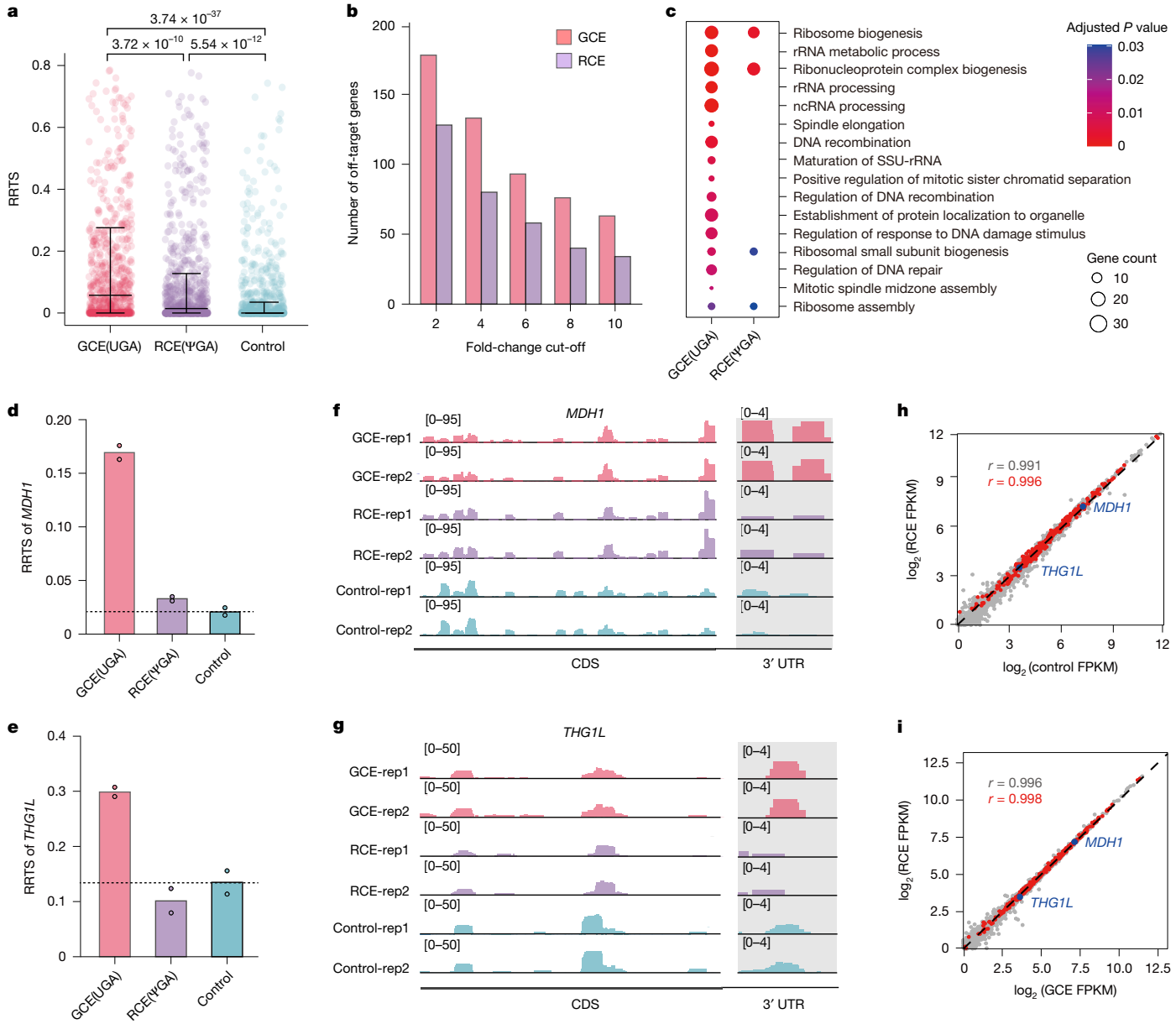


Fig. 3 | The RCE(ΨGA) system exhibited high translatome-wide decoding specificity without transcriptome-wide disturbance. a, Jitter plot of RRTS values for transcripts containing canonical UGA stop codons with 3′ UTR footprints in GCE(UGA), RCE(ΨGA) and control cells via ribosome profiling experiments for (n = 909 transcripts) from n = 2 biologically independent replicates. The centre horizontal line indicates the median RRTS and error bars represent 50% confidence intervals. P values for RRTS values between groups were computed using a two-sided Mann–Whitney U test. **b**, Bar plot depicting the number of potential off-target transcripts detected in GCE(UGA) and RCE(ΨGA) cells at indicated RRTS fold-change thresholds. **c**, Comparison of enriched GO terms of off-target transcripts in GCE(UGA) and RCE(ΨGA) cells. P value was calculated by hypergeometric test and subsequently corrected by the Benjamini–Hochberg method. SSU-rRNA, small subunit rRNA. **d**,**e**, RRTS values of *MDH1* (**d**) and *THG1L* (**e**) in GCE(UGA), RCE(ΨGA) and control cells. Data are shown as mean values (n = 2 biologically independent replicates). **f**,**g**, Zoomed-in IGV views of ribosome footprints on representative *MDH1* (**f**) and *THG1L* (**g**) transcripts in which natural readthrough efficiencies on the canonical stop codon, UGA, were upregulated (n = 2 biologically independent replicates). **h**,**i**, RNA sequencing, illustrating consistent RNA transcription levels between RCE(ΨGA) and control cells (**h**), and between RCE(ΨGA) and GCE(UGA) cells (**i**), including off-target transcripts (red dots), *MDH1* and *THG1L* (blue dots), and other transcripts (grey dots). Data are shown as mean values (n = 2 biologically independent replicates).

For *THG1L*, the RCE(ΨGA) cells possessed no clear off-target readthrough on *THG1L*, whereas the GCE(UGA) system induced an RRTS increase of 0.153 (Fig. 3e). These lower off-target stop codon readthrough ratios in RCE(ΨGA) cells were directly supported by the counts per million mapped reads (CPM) viewed using the Integrated Genome Viewer (IGV) (Fig. 3f,g).

In addition to ribosome profiling, we also conducted RNA-seq experiments to assess any potential effects of the RCE approach on gene expression. We found no interference in the transcriptome of RCE(ΨGA) or GCE(UGA) cells, supported by robust correlations among RCE(ΨGA), GCE(UGA) and control cells (Fig. 3h,i). This indicated that the interferences of codon-expansion methods probably occur in the translatome but not the transcriptome, consistent with previous reports[42]. Collectively, these results demonstrated that the RCE(ΨGA) system exhibits higher specificity during codon expansion than the GCE(UGA) approach.

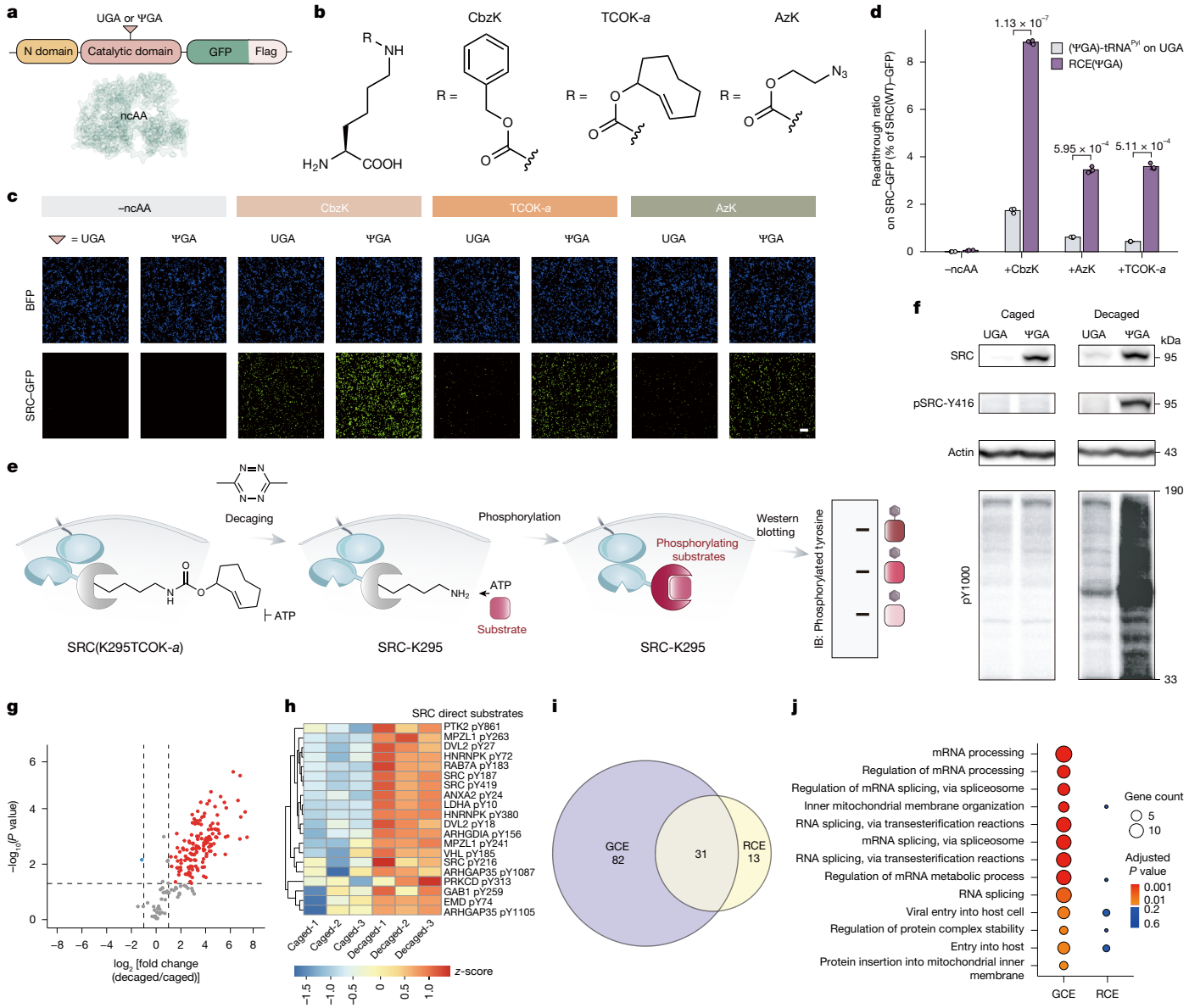


Fig. 4 | RCE(ΨGA) enables site-specific ncAA incorporation and precise modulation of protein activity in living cells. a, Scheme of the SRC(K295*/Y527F)–GFP (SRC–GFP) construct for ncAA incorporation. **b**, Chemical structures of CbzK, TCOK-*a*, and AzK. **c**,**d**, Representative fluorescence images (**c**) of cells and readthrough ratios (**d**) of (ΨGA)-tRNA^{Pyl} on the UGA and ΨGA codons of SRC(K295*/Y527F)–GFP reporter in the absence and presence of different ncAAs. Scale bar, 200 μm. **d**, Readthrough ratios of (ΨGA)-tRNA^{Pyl} on the UGA and ΨGA codons of SRC(K295*/Y527F) in the absence of ncAA and the presence of CbzK, TCOK-*a* or AzK. Data are mean ± s.d. ($n = 3$ biologically independent replicates). *P* values were calculated by a two-sided Student's *t*-test. **e**, Illustration of Me_2Tz-mediated cleavage of the *trans*-cyclooctene (TCO) moiety and the resulting SRC kinase activation. **f**, Western blotting of (ΨGA)-tRNA^{Pyl}-mediated readthrough products from suppression of the UGA or ΨGA codons, as well as the phosphorylated substrates using pY1000 or anti-pSRC Y416 antibodies. Each experiment was repeated twice independently with similar results. **g**, Volcano plot showing the abundance changes of the identified phosphotyrosine (pTyr) sites following SRC(K295TCOK-*a*) decaging. Red dots represent significantly upregulated (more than twofold) pTyr sites. *P* values were calculated by two-sided Student's *t*-tests ($n = 3$ biologically independent replicates). **h**, Heat maps showing the SRC decaging-triggered increase of phosphorylations on proteins that have been previously reported as direct SRC substrates, across the three biologically independent replicates of caged and decaged samples. The colour bar represents the phosphorylation level (*z*-score). **i**, Venn diagram showing a high overlap of ncAA-incorporated endogenous proteins induced by the GCE and RCE systems. **j**, GO enrichment analysis of the off-target misincorporated endogenous proteins, showing that the standard GCE system is much more enriched in misincorporated biological processes than the RCE system. *P* values were calculated by hypergeometric tests and subsequently corrected by the Benjamini–Hochberg method ($n = 3$ biologically independent replicates).

Site-specific ncAA encoding via RCE

After examining the incorporation specificity in the translatome, we demonstrated the applicability of the RCE strategy for incorporating ncAAs with diverse functionalities, which would facilitate the probing and modulation of protein activity, location and interactions within living cells. We selected SRC kinase, a well-known non-receptor protein tyrosine kinase that participates in signalling pathways that control a wide range of biological activities. As lysine 295 in its catalytical pocket is crucial for ATP docking and phosphoryl transfer, we produced a SRC(K295*/Y527F)–GFP construct[43,44] (Fig. 4a), in which a ncAA was incorporated at residue 295 in the catalytic pocket, and the Y527F mutation abolishes autoinhibition by Y527 phosphorylation, thereby mimicking the constitutive phosphorylation activity of SRC.

To validate the incorporation capability of different ncAAs, we selected three representative ncAAs, including an aromatic lysine derivative, CbzK, a chemically caged lysine analogue, TCOK-*a* (axial isomer of *trans*-cyclooctene-caged lysine amino acid), and the bioorthogonally clickable AzK (Fig. 4b). With a 60.9% yield for the ΨGA codon (Supplementary Fig. 11a), we determined that (ΨGA)-tRNA^{Pyl} had a high ΨGA preference (from 5.1- to 8.3-fold) in the presence of CbzK, AzK and TCOK-*a* (Fig. 4c,d). For the ncAA incorporation efficiency, we identified similar readthrough efficiencies of RCE(ΨGA) and GCE(UGA) on the SRC(K295*/Y527F)–GFP construct in the presence of CbzK, AzK or TCOK-*a* (Supplementary Fig. 12a,b). Therefore, the RCE(ΨGA) system incorporated ncAAs into proteins as efficiently as the standard GCE system.

These site-specifically incorporated ncAAs allowed specific protein labelling via the bioorthogonal ligation reaction[45], as well as in situ protein activation via the bioorthogonal cleavage reaction inside cells[46]. In particular, chemically caged ncAAs that can undergo bioorthogonal cleavable reactions have been extensively applied for gain-of-function study of proteins under native cellular context[46]. Among them, TCOK-*a* has been increasingly used as it can undergo rapid and biocompatible inverse electron-demand Diels–Alder (invDA) reaction in the presence of 3,6-dimethyl-1,2,4,5-tetrazine (Me_2Tz)[47] (Fig. 4e). To this end, we utilized this reaction pair and the RCE(ΨGA) method for bioorthogonal SRC kinase activation in living cells. We first demonstrated that RCE(ΨGA) achieved efficient expression of the caged oncogenic SRC mutant (SRC(K295TCOK-*a*/Y527F)–GFP), similarly to the previous GCE(UGA) strategy[44] (Fig. 4f and Supplementary Fig. 11b–d). Of note, incorporation of TCOK-*a* at SRC Lys295 temporally blocked SRC activity, which could be rescued by adding Me_2Tz to trigger the regeneration of the native SRC protein (Fig. 4f and Supplementary Fig. 12c–e).

We then performed functional investigations in a mass spectrometry-based proteomics study. Although investigating SRC substrates remains difficult owing to the dynamic and low-abundance tyrosine phosphorylation, we successfully profiled the phosphoproteome changes upon SRC decaging. Using an affinity purification–mass spectrometry strategy[48], we identified significantly upregulated phosphotyrosyl (pTyr) sites (Fig. 4g and Supplementary Table 4), several of which have been previously shown as direct substrates of SRC (Fig. 4h). Furthermore, many of the identified pTyr sites contain the characteristic SRC substrate motif (Fig. 4i and Supplementary Fig. 12f,g). We also detected phosphorylation at Y416, the auto-phosphorylation site[44], upon SRC activation (Fig. 4f and Supplementary Fig. 12e). These results collectively demonstrated that our gain-of-function SRC decaging strategy effectively elicits cellular responses that are directly attributable to the enzymatic function of SRC. Together, we demonstrated that the RCE(ΨGA) method enabled orthogonal control and functional study of enzyme activity in a gain-of-function manner in living cells.

Next, we investigated the specificity of bioorthogonal protein labelling of the RCE system. We used a GCE(UGA) system with a widely used *Mm*PylRS–tRNA pair[49–51], which we referred to here as TetRS–tRNA^{Tet}, to incorporate TetBu (3-(6-butyl-1,2,4,5-tetrazin-3-yl)-L-phenylalanine), a commercially available ncAA (Extended Data Fig. 5a). For the RCE(ΨGA) system, we adopted tRNA^{Tet}-37G with ΨGA preference (Extended Data Fig. 5b). Although the on-target constructs were expressed at an endogenous concentration in the GCE and RCE system, the off-target signal from the RCE system was significantly lower than the GCE system, demonstrating the ncAA incorporation specificity (Extended Data Fig. 5b–d).

We further illustrated the negligible disturbance of the RCE system on endogenous transcripts by measuring potential off-target TetBu incorporations without transfecting any exogenous reporters (Extended Data Fig. 5e,f). Finally, we detected the off-target TetBu-incorporated proteins via mass spectrometry, and identified 43 and 113 off-target TetBu-incorporated proteins for RCE and GCE, respectively (Extended Data Fig. 6). Of note, the off-target incorporations induced by RCE are nearly a subset of that induced by GCE, further demonstrating the superior specificity of the RCE system (Fig. 4i,j and Supplementary Table 5). Collectively, our results demonstrated that the RCE(ΨGA) method could introduce functional ncAAs as efficiently as the GCE(UGA) strategy while maintaining high specificity on the ΨGA, allowing applications ranging from specific protein labelling to functional modulations in living cells.

Expanding mRNA codons beyond ΨGA

Following the establishment of the RCE(ΨGA) system, we next extended the RCE rationale to create additional Ψ codons for ncAA incorporation by adapting the similar encoding and decoding processes. To develop the RCE(ΨAA) system, we developed two representative reporters, *LDLR*-TAA and *AGXT*-TAA, adapted from the *LDLR*-TGA and *AGXT*-TGA reporters (Supplementary Table 1). When producing the ΨAA codon, we observed high ΨAA yields of ~100% on *AGXT*-TAA reporters with the corresponding gsnoRNA-*AGXT* (Supplementary Fig. 13a). Via a screening process (Supplementary Fig. 14) which is similar to that for ΨGA, we identified tRNA variants showing preference for ΨAA, and selected *Mit*RNA^{Pyl}(UUA)-37G with 2.1-fold ΨAA preference and GCE-comparable efficiency as (ΨAA)-tRNA^{Pyl} (Fig. 5a–e and Supplementary Fig. 13b,c). For the RCE(ΨAG) system, we screened the ΨAG-preferring tRNA variants without obtaining satisfying variants with high ΨAG preference and readthrough. Given that mutations at position 37 improved ΨAG preference and mutations at position 31 exhibited a higher readthrough efficiency (Supplementary Fig. 15), we performed a second round of screening for tRNAs with double mutations at positions 31 and 37. To our delight, the *Mm*tRNA^{Pyl}(CUA)-31G37G exhibited a 2.1-fold ΨAG preference and ~15% ΨAG readthrough efficiency, and thus was chosen as (ΨAG)-tRNA^{Pyl} (Supplementary Table 6 and Supplementary Fig. 16).

We then examined the preference of (ΨAG)-tRNA^{Pyl} on the produced ΨAG codon in *p53* transcripts. We produced a *p53*^{K305*}–EGFP construct, encoding a p53 variant in which the AAG codon for lysine 305 that is crucial for nuclear localization[52–54] was replaced by a UAG codon (Fig. 5f), and then used our RCE(ΨAG) system for TCOK-*a* incorporation to control p53 protein localization. With a high ΨAG yield of 74.2% (Supplementary Fig. 17), (ΨAG)-tRNA^{Pyl} exhibited a 12.3-fold ΨAG preference on the *p53*^{K305*}–EGFP construct (Fig. 5f). The RCE(ΨAG) system demonstrated a yield of 10.7% (Fig. 5f), and the produced p53(K305TCOK-*a*)–EGFP protein was specifically located in the cytoplasm (Fig. 5g). After Me_2Tz-mediated TCOK-*a* decaging, the decaged protein p53-K305–EGFP regained its nuclear localization (Fig. 5g and Supplementary Fig. 18). Notably, the three specific Ψ codon:(Ψ codon)-tRNA^{Pyl} pairs are mutually orthogonal, as the anticodon of (Ψ codon)-tRNA^{Pyl} specifically pairs with its corresponding Ψ codon (Fig. 5h and Supplementary Fig. 19).

RCE is compatible with GCE

Finally, we demonstrated that the RCE strategy could cooperate compatibly with other codon-expansion methods, and in particular the GCE strategy. We integrated the RCE(ΨGA) and the GCE(UAG) strategy to incorporate two different ncAAs (Fig. 5i). As the catalytic activity of SRC kinase relies on its lysine 295 and can be inhibited by phosphorylation at Y527 (ref. 43), its activity could be caged through the incorporation of TCOK-*a* at residue 295 and the ncAA BocK at residue 527.

To illustrate dual ncAA incorporation in SRC, we used the *Ma*RS–*Ma*$^{\Delta N}$tRNA^{Pyl}(CUA)(19) pair[29] as the GCE(UAG) system to introduce BocK. For the RCE(ΨGA) system to install TCOK-*a*, we utilized the *Mm*PylRS-306A384F–*Mm*tRNA^{Pyl}(UCA)-6T37G pair, in which a C at position 6 of *Mm*tRNA^{Pyl}(UCA)-37G was mutated to a T to ensure orthogonality with *Ma*PylRS[29] according to the acceptor arm on *Methanosarcina spelaei* (*Spe*)tRNA^{Pyl}. The *Mm*tRNA^{Pyl}(UCA)-6T37G, still exhibited strong ΨGA preference regardless of the presence of the GCE(UAG) system

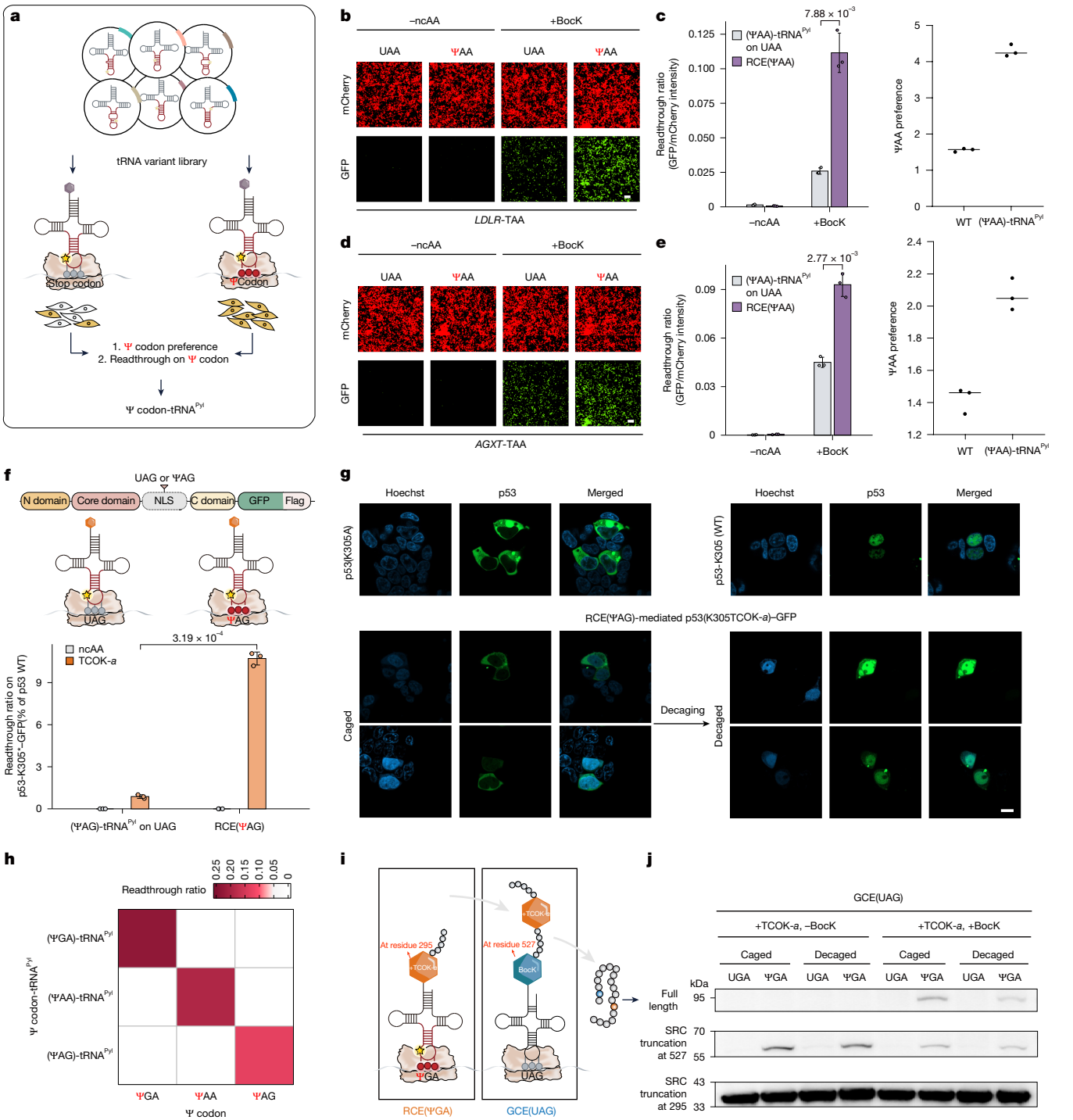


Fig. 5 | The RCE strategy is expandable to other modified RNA codons and is compatible with the GCE strategy. a, Schematic of the screening strategy for (ΨAA)-tRNA^{Pyl}. **b**,**c**, Representative fluorescence images of cells (**b**) and bar plots and dot plots (**c**) showing the readthrough performances of (ΨAA)-tRNA^{Pyl} on UAA and ΨAA codons of the *LDLR*-TAA and *AGXT*-TAA reporters. Scale bar, 200 μm. Data are mean ± s.d. ($n = 3$ biologically independent replicates). **d**,**e**, Representative fluorescence images of cells (**d**) and bar plots and dot plots (**e**) showing the readthrough performances of (ΨAA)-tRNA^{Pyl} on *LDLR*-TAA and *AGXT*-TAA reporters. Scale bar, 200 μm. Data are mean ± s.d. ($n = 3$ biologically independent replicates). *P* values were calculated by a two-sided Student's *t*-test. **f**, Readthrough ratios of (ΨAG)-tRNA^{Pyl} on the UAG and ΨAG codons of the *p53*^{K305*}–GFP construct in the absence and presence of TCOK-*a*. Data are mean ± s.d. ($n = 3$ biologically independent replicates). *P* values were calculated by a two-sided Student's *t*-test. NLS, nuclear localization signal. **g**, Representative fluorescence images of cells containing RCE(ΨAG)-produced p53(K305TCOK-*a*)–GFP protein before and after decaging. Cells expressing p53-K305A–GFP and p53-K305–GFP were used as cytoplasm- and nucleus-localized protein controls, respectively. The experiment was repeated twice independently with similar results. Scale bar, 10 μm. **h**, Readthrough ratio heatmap of three Ψ codons independently decoded by the corresponding (Ψ codon)-tRNA^{Pyl} decoder tRNAs without cross-activity. Data are shown as mean values ($n = 3$ biologically independent replicates). **i**, Schematic illustrating dual ncAA incorporation by coordination of the mutually orthogonal RCE(ΨGA) and GCE(UAG) systems. **j**, Orthogonal TCOK-*a* and BocK dual incorporation (producing SRC(K295TCOK-*a*/Y527BocK) protein) by coordination of RCE(ΨGA) and GCE(UAG) systems. The dual-ncAA-incorporated full-length SRC protein was verified by western blotting, which was repeated twice independently with similar results.

(Fig. 5j), and obtained orthogonality to *Ma*PylRS. This orthogonality was also demonstrated via mass spectrometry[44,55] (Supplementary Fig. 20).

This adaption implied that based on properly arranged orthogonal RS–tRNA pairs[29], we could integrate and program various codon-expansion methods as orthogonal modules for encoding different ncAAs. Thus, we have demonstrated the adaptability of the (Ψ codon)-tRNA^{Pyl} in the RCE approach, enabling RCE to serve as an independent codon-expansion module that work compatibly with the GCE strategy.

Discussion

In summary, we established a general applicable RCE strategy and obtained three triply orthogonal pairs, ΨGA:(ΨGA)-tRNA^{Pyl}, ΨAA:(ΨAA)-tRNA^{Pyl} and ΨAG:(ΨAG)-tRNA^{Pyl}, for site-specific ncAA incorporation into proteins in mammalian cells. Using ribosome profiling and proteomics analysis, we demonstrated the translatome-wide decoding specificity of the RCE strategy, which significantly reduced off-target stop codon readthroughs compared with the standard GCE method. The high specificity of the RCE strategy was verified through multiple approaches, indicating that RCE-based protein decaging offers a general strategy for activation of enzymes of interest. In the encoding component, we identified high Ψ codon yields on these specified mRNA transcripts with specified gsnoRNAs[4]. In the decoding component, the (Ψ codon)-tRNA^{Pyl} decode Ψ codons with robust Ψ codon preferences across various transcripts, consistent with the globally reduced off-target readthrough events. The specificities in encoding and decoding processes contributed to the overall ncAA-incorporating specificity of our RCE strategy, which could be further advanced by engineering the relevant mRNAs, small nucleolar RNAs (snoRNAs) and decoder tRNAs. In addition, although we focused on the stop codons in this study, our RCE strategy could in principle be extended to sense codons, owing to its programmability and specificity during the encoding and decoding processes. Indeed, a new GCE strategy leveraging rare codons has been reported recently[51].

In addition to pseudouridine, the RCE approach may allow for the utilization of various post-transcriptionally modified RNA-expanded codons for translation. More than 150 types of chemical modification have been identified in cellular RNAs so far, most of which can influence the stability, structure and interactions of RNA to a certain extent, including N^6-methyladenosine, N^1-methylpseudouridine, 5-methylcytosine and 2′-*O*-methylation[22,23]. We anticipate that we could generate additional blank RNA codons by adapting the corresponding encoding and decoding processes. The encoding of these varied RNA modifications can be attained with molecular precision—for instance, by using Cas13-directed methyltransferase[56] or CRISPR–Cas9 based enzymes[57]. For decoding, the engineering and characterization of (Ψ codon)-tRNA^{Pyl} decoder tRNAs could enable the development of specific decoder tRNAs for other modified RNA codons. Collectively, the RCE approach enables a subtantial expansion of the four-letter (A, U, C and G) transcription 'alphabet'.

Furthermore, the RCE approach can act as an adaptable strategy that is compatible with the current approaches that exploit translation processes, such as RNA base editors[20] and orthogonally translating organelles[18,19]. RNA base editors can deaminate the C to U, converting sense codons into stop codons on endogenous mRNA transcripts[20]. In the same transcripts, the gsnoRNA for encoding in the RCE system could pseudouridylate the produced U of the targeted stop codon, followed by decoding in the RCE system. Similarly, the RCE approach could operate compatibly with the orthogonally translating organelles to spatially confine UAG codons[18,19]. This is due to the molecular basis of the RCE strategy, in which the Ψ codon and the (Ψ codon)-tRNA^{Pyl}, which are compatible with the RNA motifs and the fused PylRS[18,19], promote phase separation. Collectively, our RCE approach offers a post-transcriptional codon-expansion strategy for ncAA incorporation, in which the modified RNA codons can be assigned independently of the native 64 heritable codons, dynamically enriching the genetic alphabet in mammalian cells.

Online content

Any methods, additional references, Nature Portfolio 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/s41586-025-09165-x.

1. Liu, C. C. & Schultz, P. G. Adding new chemistries to the genetic code. *Annu. Rev. Biochem.* **79**, 413–444 (2010).
2. Lemke, E. A. The exploding genetic code. *ChemBioChem* **15**, 1691–1694 (2014).
3. Chin, J. W. Expanding and reprogramming the genetic code. *Nature* **550**, 53–60 (2017).
4. Song, J. et al. CRISPR-free, programmable RNA pseudouridylation to suppress premature termination codons. *Mol. Cell* **83**, 139–155.e139 (2023).
5. Neumann, H., Wang, K., Davis, L., Garcia-Alai, M. & Chin, J. W. Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature* **464**, 441–444 (2010).
6. Orelle, C. et al. Protein synthesis by ribosomes with tethered subunits. *Nature* **524**, 119–124 (2015).
7. Fried, S. D., Schmied, W. H., Uttamapinant, C. & Chin, J. W. Ribosome subunit stapling for orthogonal translation in *E. coli*. *Angew. Chem. Int. Ed.* **54**, 12791–12794 (2015).
8. Isaacs, F. J. et al. Precise manipulation of chromosomes in vivo enables genome-wide codon replacement. *Science* **333**, 348–353 (2011).
9. Lajoie, M. J. et al. Genomically recoded organisms expand biological functions. *Science* **342**, 357–360 (2013).
10. Ostrov, N. et al. Design, synthesis, and testing toward a 57-codon genome. *Science* **353**, 819–822 (2016).
11. Wang, K. et al. Defining synonymous codon compression schemes by genome recoding. *Nature* **539**, 59–64 (2016).
12. Fredens, J. et al. Total synthesis of *Escherichia coli* with a recoded genome. *Nature* **569**, 514–518 (2019).
13. Nyerges, A. et al. Synthetic genomes unveil the effects of synonymous recoding. Preprint at *bioRxiv* https://doi.org/10.1101/2024.06.16.599206 (2024).
14. Grome, M. W. et al. Engineering a genomically recoded organism with one stop codon. *Nature* **639**, 512–521 (2025).
15. Malyshev, D. A. et al. A semi-synthetic organism with an expanded genetic alphabet. *Nature* **509**, 385–388 (2014).
16. Fischer, E. C. et al. New codons for efficient production of unnatural proteins in a semisynthetic organism. *Nat. Chem. Biol.* **16**, 570–576 (2020).
17. Thompson, D. B. et al. The future of multiplexed eukaryotic genome engineering. *ACS Chem. Biol.* **13**, 313–325 (2018).
18. Reinkemeier, C. D., Girona, G. E. & Lemke, E. A. Designer membraneless organelles enable codon reassignment of selected mRNAs in eukaryotes. *Science* **363**, eaaw2644 (2019).
19. Reinkemeier, C. D. & Lemke, E. A. Dual film-like organelles enable spatial separation of orthogonal eukaryotic translation. *Cell* **184**, 4886–4903.e4821 (2021).
20. Hao, M. et al. Tracking endogenous proteins based on RNA editing-mediated genetic code expansion. *Nat. Chem. Biol.* **20**, 721–731 (2024).
21. Ranjan, N. & Leidel, S. A. The epitranscriptome in translation regulation: mRNA and tRNA modifications as the two sides of the same coin? *FEBS Lett.* **593**, 1483–1493 (2019).
22. Boccaletto, P. et al. MODOMICS: a database of RNA modification pathways. 2021 update. *Nucleic Acids Res.* **50**, D231–D235 (2022).
23. Shi, H., Chai, P., Jia, R. & Fan, X. Novel insight into the regulatory roles of diverse RNA modifications: Re-defining the bridge between transcription and translation. *Mol. Cancer* **19**, 78 (2020).
24. Karijolich, J. & Yu, Y.-T. Converting nonsense codons into sense codons by targeted pseudouridylation. *Nature* **474**, 395–398 (2011).
25. Luo, N. et al. Near-cognate tRNAs increase the efficiency and precision of pseudouridine-mediated readthrough of premature termination codons. *Nat. Biotechnol.* **48**, 114–123 (2025).
26. Charette, M. & Gray, M. W. Pseudouridine in RNA: what, where, how, and why. *IUBMB Life* **49**, 341–351 (2000).
27. Zhang, M. et al. Quantitative profiling of pseudouridylation landscape in the human transcriptome. *Nat. Chem. Biol.* **19**, 1185–1195 (2023).
28. Hofhuis, J. et al. The functional readthrough extension of malate dehydrogenase reveals a modification of the genetic code. *Open Biol.* **6**, 160246 (2016).
29. Dunkelmann, D. L., Willis, J. C. W., Beattie, A. T. & Chin, J. W. Engineered triply orthogonal pyrrolysyl–tRNA synthetase/tRNA pairs enable the genetic encoding of three distinct non-canonical amino acids. *Nat. Chem.* **12**, 535–544 (2020).
30. Seelam Prabhakar, P., Takyi, N. A. & Wetmore, S. D. Posttranscriptional modifications at the 37th position in the anticodon stem-loop of tRNA: structural insights from MD simulations. *RNA* **27**, 202–220 (2021).
31. Nozawa, K. et al. Pyrrolysyl-tRNA synthetase–tRNAPyl structure reveals the molecular basis of orthogonality. *Nature* **457**, 1163–1167 (2009).
32. Suzuki, T. et al. Crystal structures reveal an elusive functional domain of pyrrolysyl-tRNA synthetase. *Nat. Chem. Biol.* **13**, 1261–1266 (2017).
33. Landrum, M. J. et al. ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res.* **42**, D980–D985 (2014).
34. Wang, J. et al. AAV-delivered suppressor tRNA overcomes a nonsense mutation in mice. *Nature* **604**, 343–348 (2022).

35. Hollingsworth, T. J. & Gross, A. K. The severe autosomal dominant retinitis pigmentosa rhodopsin mutant Ter349Glu mislocalizes and induces rapid rod cell death. *J. Biol. Chem.* **288**, 29047–29055 (2013).
36. Vidal, R. et al. A stop-codon mutation in the BRI gene associated with familial British dementia. *Nature* **399**, 776–781 (1999).
37. Shibata, N. et al. Degradation of stop codon read-through mutant proteins via the ubiquitin–proteasome system causes hereditary disorders. *J. Biol. Chem.* **290**, 28428–28437 (2015).
38. Namy, O., Duchateau-Nguyen, G. & Rousset, J.-P. Translational readthrough of the PDE2 stop codon modulates cAMP levels in *Saccharomyces cerevisiae*. *Mol. Microbiol.* **43**, 641–652 (2002).
39. Capone, J. P., Sharp, P. A. & RajBhandary, U. L. Amber, ochre and opal suppressor tRNA genes derived from a human serine tRNA gene. *EMBO J.* **4**, 213–221 (1985).
40. Wangen, J. R. & Green, R. Stop codon context influences genome-wide stimulation of termination codon readthrough by aminoglycosides. *eLife* **9**, e52611 (2020).
41. Schueren, F. et al. Peroxisomal lactate dehydrogenase is generated by translational readthrough in mammals. *eLife* **3**, e03640 (2014).
42. Shi, N. et al. Restoration of dystrophin expression in mice by suppressing a nonsense mutation through the incorporation of unnatural amino acids. *Nat. Biomed. Eng.* **6**, 195–206 (2022).
43. Yeatman, T. J. A renaissance for SRC. *Nat. Rev. Cancer* **4**, 470–480 (2004).
44. Zhang, G. et al. Bioorthogonal chemical activation of kinases in living systems. *ACS Cent. Sci.* **2**, 325–331 (2016).
45. Sletten, E. M. & Bertozzi, C. R. Bioorthogonal chemistry: fishing for selectivity in a sea of functionality. *Angew. Chem. Int. Ed.* **48**, 6974–6998 (2009).
46. Li, J. & Chen, P. R. Development and application of bond cleavage reactions in bioorthogonal chemistry. *Nat. Chem. Biol.* **12**, 129–137 (2016).
47. Li, J., Jia, S. & Chen, P. R. Diels–Alder reaction-triggered bioorthogonal protein decaging in living cells. *Nat. Chem. Biol.* **10**, 1003–1005 (2014).
48. Bian, Y. et al. Ultra-deep tyrosine phosphoproteomics enabled by a phosphotyrosine superbinder. *Nat. Chem. Biol.* **12**, 959–966 (2016).
49. Jang, H. S., Jana, S., Blizzard, R. J., Meeuwsen, J. C. & Mehl, R. A. Access to faster eukaryotic cell labeling with encoded tetrazine amino acids. *J. Am. Chem. Soc.* **142**, 7245–7249 (2020).
50. Bryson, D. I. et al. Continuous directed evolution of aminoacyl-tRNA synthetases. *Nat. Chem. Biol.* **13**, 1253–1260 (2017).
51. Ding, W. et al. Rare codon recoding for efficient noncanonical amino acid incorporation in mammalian cells. *Science* **384**, 1134–1142 (2024).
52. Liang, S.-H. & Clarke, M. F. Regulation of p53 localization. *Eur. J. Biochem.* **268**, 2779–2783 (2001).
53. O'Keefe, K., Li, H. & Zhang, Y. Nucleocytoplasmic shuttling of p53 is essential for MDM2-mediated cytoplasmic degradation but not ubiquitination. *Mol. Cell. Biol.* **23**, 6396–6405 (2003).
54. Gautier, A. et al. Genetically encoded photocontrol of protein localization in mammalian cells. *J. Am. Chem. Soc.* **132**, 4086–4088 (2010).
55. Guangcan, S. et al. How to use open-pFind in deep proteomics data analysis?—A protocol for rigorous identification and quantitation of peptides and proteins from mass spectrometry data. *Biophys. Rep.* **7**, 207–226 (2021).
56. Wilson, C., Chen, P. J., Miao, Z. & Liu, D. R. Programmable m^6A modification of cellular RNAs with a Cas13-directed methyltransferase. *Nat. Biotechnol.* **38**, 1431–1440 (2020).
57. Liu, X.-M., Zhou, J., Mao, Y., Ji, Q. & Qian, S.-B. Programmable RNA N^6-methyladenosine editing by CRISPR–Cas9 conjugates. *Nat. Chem. Biol.* **15**, 865–871 (2019).

Methods

Cell culture

HEK293T, U-2 OS, COS-7 and HeLa cells were cultured in DMEM medium (Corning, 10-013-CVR) and CHO-K1 cells were cultured in DMEM/F12 medium (Gibco). All cells were cultured with medium containing 10% FBS and 1% penicillin/streptomycin (both from Gibco, v/v) at 37 °C with 5% CO_2. To passage cells, they were initially rinsed with PBS (Corning), and then treated with 0.25% Trypsin (Gibco) before incubation (37 °C, 1 min). Following this, the trypsin was neutralized by adding FBS-containing medium. The cells were subsequently collected by centrifugation (500*g*, 5 min), counted, and divided for various experimental uses. All cells were confirmed to be free from mycoplasma contamination using a mycoplasma detection kit (TransGene Biotech, FM311-01) prior to use.

Plasmid construction

Dual-fluorescent reporters were generated based on previously described dual-colour reporters using site-directed mutagenesis with Fast*Pfu* polymerase (Transgene) following the manufacturer's guidelines. For the construction of hU6-driven gsnoRNA expression constructs, gsnoRNA fragments were amplified using an overlapping PCR strategy involving one sense strand and three antisense strands, and subsequently integrated into the pLenti-sgRNA-lib 2.0 backbone (Addgene, #89638) using Golden Gate cloning. The secondary structures of gsnoRNAs were predicted using RNAfold[58,59] (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi) with default settings. The sequences of gsnoRNAs and synthetases are provided in Supplementary Table 1. Synthetases and tRNAs were custom synthesized by Beijing Qsingke Biotechnology Company. *p53* and *SRC* genes were also synthesized by Beijing Qsingke Biotechnology Company and inserted into the pcDNA3.1-GFP vector using Gibson assembly with NEBuilder HiFi DNA Assembly Master Mix (NEB). RS3#-2×(ΨAG)-tRNA^{Pyl} and RS3#-2×(ΨGA)-tRNA^{Pyl} constructs were generated by recombining PCR-amplified RS3# and Decoder sequences using NEBuilder HiFi DNA Assembly Master Mix (NEB). All clones were verified by Sanger sequencing. Plasmids for transfection were purified using the EndoFree Mini Plasmid Kit II (Tiangen Biotech).

Screening assay for decoder tRNA based on high-content imaging

HEK293T cells were rinsed with PBS and treated with 0.25% Trypsin at 37 °C for 1 min. Trypsin activity was halted by adding FBS-containing medium. After centrifugation at 500*g* for 5 min, the cells were enumerated and diluted to a concentration of 4×10^5 cells per ml. Subsequently, the cells were seeded into individual wells of 96-well flat-bottom plates (LABSELECT, 11514) that had been pre-coated overnight with a 20 ng μl^{-1} poly-D-lysine (Beyotime, ST508) solution and rinsed with water. Following 16–24 h of incubation, each tRNA mutant was separately assessed under both the UGA and ΨGA conditions.

For each tRNA mutant, a ΨGA mixture was prepared by combining 62.5 ng of the Screen-TGA reporter, 15.6 ng of gsno-TGA, 7.8 ng of DKC1-*iso3*, 62.5 ng of synthetases, and 122.5 ng of the tRNA mutant with 0.5 μl of Lipofectamine LTX reagent (Invitrogen) and 0.2 μl of PLUS reagent (Invitrogen), following the manufacturer's instructions. Similarly, for the UGA mixture, 62.5 ng of the Screen-TGA reporter, 15.6 ng of gctrl, 7.8 ng of DKC1-vector, 62.5 ng of synthetases, and 122.5 ng of the tRNA mutant were mixed with 0.5 μl of Lipofectamine LTX reagent (Invitrogen) and 0.2 μl PLUS reagent (Invitrogen) in the same manner.

The UGA and ΨGA mixtures were added to separate wells containing HEK293T cells cultured in a medium supplemented with 200 μM ncAA (CbzK or BocK, corresponding to the synthetases). After 24 h, the transfected cells were provided with fresh culture medium. At 48–72 h post-transfection, cell imaging was conducted using the ImageXpress Micro 4 high-content imaging system (Molecular Devices). Four images from different regions of a single well on the 96-well plates were captured using a 10× microscope objective, and these images were subsequently analysed automatically using MetaXpress software (MX version 6.2.3.733; CME version 6.2.3.991). The intensities of mCherry or EGFP were determined as the intensity per cell multiplied by the number of positive cells in the fluorescence images. The readthrough ratio was calculated as the GFP intensity divided by the mCherry intensity, and then normalized to the DF-Read reporter without any truncations.

Readthrough product enrichments

HEK293T cells were seeded into 6-well plates (Corning, 3516) and incubated for 16–24 h. For general readthrough reporters with adequate gsnoRNAs, 1 μg of expression reporter, 250 ng of gsnoRNA, 125 ng of DKC1-*iso3*, and 1 μg of RS-2×Decoder plasmids were transfected using PEI (Yeasen, 40816ES02) following to the manufacturer's directions in the presence of ncAAs. The culture medium was changed after 24 h of transfection. At 96 h post-transfection, the cells were rinsed with PBS and collected with PBS. Following centrifugation, the cells were resuspended with a TBS-containing protease inhibitor (Roche) and 1 mM DTT. After ultrasonic lysis, the cell lysate was centrifuged at 21,000 rpm at 4 °C. The supernatant was obtained and incubated for 2 h alongside anti-Flag M2 magnetic beads (Sigma Aldrich) for flag-tagged readthrough product purification following the manufacturer's directions. The beads were rinsed with TBST three times, TBS three times, and finally with H_2O. Samples with a concentration of 2 mg ml^{-1} of 3×Flag peptide in TBS were used to elute the enriched readthrough products at 37 °C for 2 h.

Amino acid characterization at specified locations by mass spectrometry

Following the enrichment of readthrough products as described above, the elution was supplemented with 5× SDS loading buffer (Beyotime), and subjected to boiling at 95 °C for 30 min. Subsequently, centrifugation at 21,000 rpm for 10 min facilitated the separation of the supernatant, which was then loaded onto a 4%–15% SDS–PAGE gel (Beyotime) along with Blue Plus V Protein Marker (DM141-01, Transgene) markers. After staining with Coomassie blue, the correct protein bands were excised and destained using a solution containing 25 mM ammonium bicarbonate in 50% acetonitrile. Following reduction with dithiothreitol and alkylation with iodoacetamide, the proteins were subjected to overnight digestion with porcine trypsin (Sequencing grade modified; Pierce) at 37 °C. The resulting tryptic peptides were extracted from the gel pieces using acetonitrile supplemented with 0.1% formic acid. Subsequently, the samples were dried using a vacuum centrifuge concentrator at 30 °C and reconstituted in 10 μl of 0.1% formic acid/H_2O. LC–MS/MS analysis was conducted using an Orbitrap Fusion Lumos mass spectrometer (Thermo Scientific). Data analysis was performed using Proteome Discoverer (Thermo) against the UniProt database (https://www.uniprot.org) using human proteome (UP000005640), along with a custom database (Supplementary Table 1) containing reporter sequences featuring 20 different natural amino acids at positions corresponding to the premature termination codon location. Variable modifications of CbzK, TCOK-*a* and BocK were considered during the search process according to their molecular weight. Abundances of peptides containing different amino acids were used to calculate the incorporation percentage of different amino acids.

Western blotting

HEK293T cells were seeded into 24-well plates (Corning, 3526) and incubated for 16–24 h. For the ΨGA condition of HA–SRC(K295*/Y527F)–GFP, 500 ng of HA-*SRC*$^{K295*/Y527F}$–GFP, 125 ng of gsnoRNA-SRC, 62.5 ng of DKC1-*iso3*, and 500 ng of RS3#-2×(ΨGA)-tRNA^{Pyl} plasmids were transfected using Lipofectamine LTX reagent (Invitrogen) supplemented with PLUS reagent (Invitrogen) following the manufacturer's directions. For the UGA condition of HA–SRC–(K295*/Y527F)–GFP,

500 ng of HA-*SRC*$^{K295*/Y527F}$–GFP, 125 ng of gctrl, 62.5 ng of DKC1-*iso3*, and 500 ng of RS3#-2×(ΨGA)-tRNA^{Pyl} plasmids were transfected using Lipofectamine LTX reagent (Invitrogen) supplemented with PLUS reagent (Invitrogen). The UGA and ΨGA conditions encompassed two copies, one was supplied to cells incubated with 100 μM ncAA, while the other was provided to cells under standard cell culture conditions as a control. For evaluating ncAA incorporation, the TetRS-2×(ΨGA)-tRNA^{Tet} plasmid was used for RCE, the TetRS-2×tRNA^{Tet} (wild-type tRNA^{Tet}) plasmid was used for GCE, others remain unchanged in transfection.

For the ΨGA condition of HA–SRC(K295*/Y527*)–GFP, 500 ng of HA-*SRC*$^{K295*/Y527*}$–GFP, 125 ng of gsno-SRC, 62.5 ng of DKC1-*iso3*, 500 ng of RS3#-2×(ΨGA)-tRNA^{Pyl} (based on *Spe*tRNA^{Pyl}), 500 ng of *Ma*PylRS, and 980 ng of *Ma*$^{\Delta N}$tRNA^{Pyl}(CUA)(19) plasmids were transfected using Lipofectamine LTX PLUS reagent (Invitrogen). For the UGA condition of HA–SRC(K295*/Y527*)–GFP, 500 ng of HA-*SRC*$^{K295*/Y527*}$–GFP, 125 ng of gctrl, 62.5 ng of DKC1-vector, 500 ng of RS3#-2×(ΨGA)-tRNA^{Pyl} (based on *Spe*tRNA^{Pyl}), 500 ng of *Ma*PylRS, and 980 ng of *Ma*$^{\Delta N}$tRNA^{Pyl}(CUA)(19) plasmids were transfected using Lipofectamine LTX PLUS reagent. The UGA and ΨGA conditions both encompassed four copies, the first was supplied to cells under common cell culture conditions, the second was supplied to cells incubated with 100 μM TCOK-*a*, the third was supplied to cells incubated with 100 μM BocK, while the fourth was supplied to cells incubated with 100 μM TCOK-*a* and 100 μM BocK.

At 24–48 h post-transfection, cells were rinsed with cold PBS and centrifuged at 500*g* at 4 °C for 5 min. The cell pellets were mixed with 1× SDS loading buffer, boiled at 95 °C for 30 min, and centrifuged at 21,000 rpm for 10 min. Proteins in supernatants were separated via electrophoresis on 4%–15% SDS–PAGE gel and transferred to PVDF membranes (Millipore). These membranes were blocked using 5% BSA and incubated with primary antibodies for 2 h at room temperature, rinsed three times with TBST, and then incubated with horseradish peroxidase (HRP)-conjugated antibodies for 1 h at room temperature. The protein bands were visualized using a ChemiDoc MP Imaging System (Bio-Rad). The intensity data were collected by Fiji (64-bit for Windows; https://imagej.net/software/fiji/downloads, released 30 May 2017).

For immunoblots shown in figures, the corresponding antibody and their dilution are as follows: mouse anti-Flag (Sigma Aldrich, F1804, 1:2,000 or 1:750 dilution); β-actin mouse monoclonal antibody (CWBIO, CW0096M, 1:5,000 dilution); phosphotyrosine (P-Tyr-1000) MultiMab rabbit monoclonal antibody mix (herein referred as pY1000) (CST, 8954, 1:2,000 or 1:750 dilution); phospho-SRC family (Tyr416) (E6G4R) rabbit monoclonal antibody 59548 (herein referred as pSRC Y416) (CST, 1:1,000 dilution); eIF2α antibody 9722 (CST, 1:1,000 dilution); phospho-eIF2α (Ser51) (D9G8) XP rabbit monoclonal antibody 3398 (CST, 1:1,000 dilution); ATF-4 (D4B8) rabbit monoclonal antibody 11815 (CST, 1:1,000 dilution); biotin antibody (33):sc-101339 (Santa Cruz, 1:2,000 dilution); p53 antibody 9282 (CST, 1:1,000 dilution); phospho-p53 (Ser15) antibody 9284 (CST, 1:1,000 dilution); goat anti-mouse IgG, HRP conjugated (CWBIO, CW0102, 1:5,000 dilution); goat anti-rabbit IgG, HRP conjugated (CWBIO, CW0103, 1:5,000 dilution).

Confocal imaging of subcellular localization

HEK293T cells cultured in wells were diluted to 4×10^5 cells per ml as outlined above and seeded on 20-mm glass-bottom dishes following coating with 20 ng μl^{-1} poly-D-lysine (Beyotime) solution overnight and pre-rinsed with water. After 16–24 h of incubation, in the presence of TCOK-*a*, 500 ng of *p53*^{K305*}–EGFP reporter, 125 ng of gsno-P53, 62.5 ng of DKC1-*iso3*, and 500 ng of RS3#-2×(ΨAG)-tRNA^{Pyl} were transfected into one specific dish according to numbering, using Lipofectamine LTX reagent (Invitrogen) supplemented with PLUS reagent (Invitrogen) following the manufacturer's directions. Each transfection condition consisted of two dishes providing TCOK-*a*; one was incubated with 200 μM Me_2Tz in medium to decage TCOK-*a*, while the other was recycled with the standard medium. After 2 h, cells were incubated with 10 μg ml^{-1} Hoechst (Beyotime) in DMEM for 10 min and rinsed with DMEM three times. A Spin SR (Evident, Olympus) was used to capture images, measure colocalization and fluorescence intensity using 63× magnification.

mRNA sequencing

Total RNAs were extracted from the same cell lysates utilized in the ribosome profiling step, and used for RNA-seq by Guangzhou Epibiotek. Library concentrations were determined using a Qubit 2.0 fluorometer with the Qubit dsDNA HS Assay kit (Invitrogen), while size distribution was assessed using the Agilent 4150 TapeStation System with High Sensitivity D1000 ScreenTape. Finally, sequencing of the libraries was performed on the Illumina Hiseq X-ten platform, generating 2 ×150 bp paired-end raw reads.

Analysis of RNA-seq data

The raw reads were quality-controlled using Trim galore[60] (version 0.6.6) to exclude adaptor sequences and low-quality reads. Clean reads (length > 30 bp) were mapped to the human reference genome (GRCh38) with HISAT2 (ref. 61) (version 4.8.5). Expression abundance values were computed using featureCount[62] (version 2.0.4), and genes with a reads per kilobase per million mapped reads (RPKM) value below 1 were removed for subsequent analysis.

Quantification of Ψ via PRAISE

HEK293T cells were plated in 24-well plates and allowed to incubate for 16–24 h. Each well received transfection with 500 ng of reporter, along with either 125 ng or 500 ng of Gsno (RCE samples) or Gctrl (GCE samples), 62.5 ng of DKC1-*iso3*, and 500 μg of RS3#-2×Decoder plasmids using Lipofectamine LTX reagent (Invitrogen) supplemented with PLUS reagent (Invitrogen), following the manufacturer's directions. Total RNA extraction was performed using TRIzol (Life Technologies), followed by isopropanol precipitation, per the manufacturer's protocol (Invitrogen). A portion of 200 ng of total RNA from each sample was subjected to DNase I treatment at 37 °C for 30 min. Subsequently, RNA was fragmented to approximately 150 nt using magnesium RNA fragmentation buffer (New England Biolabs) for 4 min at 94 °C, followed by cooling on ice. The reaction was quenched using RNA fragmentation stop solution and then purified by ethanol precipitation. Following ethanol precipitation, RNA was resuspended in 5 μl of nuclease-free water. A fresh bisulfite solution was freshly prepared by dissolving 4.05 g of sodium bisulfite (Sigma Aldrich) in 5.5 ml of RNase-free water, adjusting the pH to 5.1 with 10 M sodium hydroxide, and adjusting the volume to 10 ml with water. Additionally, a 100 mM hydroquinone solution was freshly prepared by adding 11.01 mg of hydroquinone (Sigma Aldrich) to 1 ml of RNase-free water. A sample of 5 μl of RNA fragments was dissolved in 50 μl of bisulfite/sulfite solution, consisting of a 100:1 mixture of bisulfite/sulfite solution and 100 mM hydroquinone, and heated to 70 °C for 5 h. The reaction mixture was then desalted by moving it twice through Micro Bio-Spin 6 chromatography columns (Bio-Rad). The desalted RNA was transferred to a new 1.5-ml nuclease-free tube and adjusted to 100 μl with RNase-free water, followed by incubation with an equal volume of 1 M Tris-HCl (pH 9.0) at 75 °C for 30 min. The reaction was promptly halted by chilling on ice and ethanol precipitation.

For targeted amplicon sequencing, the bisulfite/sulfite-treated RNA was subsequently reverse transcribed into cDNA using random hexamers (Thermo Fisher) with Maxima H minus Reverse Transcriptase (Thermo Fisher). For transcriptomic sequencing, mRNA or total RNA for each sample was subjected to library construction using SMARTer Stranded Total RNA-Seq Kit v3 Pico Input Mammalian (Takara Bio, 634485) according to the manufacturer's protocol, with a substitution of reverse transcriptase Maxima H minus Reverse Transcriptase (Thermo Fisher).

Analysis of PRAISE data

The identification of Ψ sites and the assessment of modification levels were conducted according to the established PRAISE pipeline (https://github.com/Zhe-jiang/PRAISE). Initially, only read 1 and 2 underwent further processing. Raw reads underwent adaptor removal and quality control utilizing cutadapt (version 2.10)[60]. PCR redundancy was then eliminated using Seqkit (version 0.14.0)[63] according to the unique molecular identifiers (UMIs) of 4 bp and 6 bp located at the 5′ and 3′ ends of reads R1, respectively. UMIs were sequentially removed by umi_tools (version 1.0.0)[64]. The obtained cleaned reads were aligned to the reference sequences using the PRAISE tool (version 4.8.5) with default parameters[27]. Based on the realigned bam file, we first calculated the deletion rate for Ψ sites, and then normalized them to the averaged deletion ratio of 18S Ψ1347 and Ψ1367 so as to obtain the actual Ψ level. To identify the transcriptome-wide off-target Ψ editing of the RCE system, we used following criteria: (1) the deletion ratio in control samples should be less than 5%; (2) the difference in deletion ratio between RCE samples and the control samples should be greater than 10%; (3) the deletion ratio in RCE samples should be statistically significantly great than that in control samples with a false discovery rate (FDR) value < 0.05. Primers for targeted deep sequencing were listed in Supplementary Table 7.

Ribosome profiling

HEK293T cells were seeded into 10-cm dishes and incubated for 16–24 h as described above. For RCE samples, 40 μg of *SRC*$^{K295*/Y527F}$–GFP reporter, 10 μg of gsno-SRC, 5 μg of DKC1-*iso3*, and 40 μg of RS3#-2×(ΨGA)-tRNA^{Pyl} plasmids were transfected into four dishes using Lipofectamine LTX reagent (Invitrogen) supplemented with PLUS reagent (Invitrogen) alongside 200 μM CbzK. For GCE samples, 40 μg of *SRC*$^{K295*/Y527F}$–GFP reporter, 10 μg of gctrl, 5 μg of DKC1-vector, and 40 μg of RS3#-2×*Mm*tRNA^{Pyl} (wild-type *Mm*tRNA^{Pyl}) plasmids were transfected into four dishes using Lipofectamine LTX reagent (Invitrogen) supplemented with PLUS reagent (Invitrogen) alongside 200 μM CbzK. The culture medium was changed after 12 h of transfection. At 48 h post-transfection, the cells were cycled into fresh DMEM medium containing 10% FBS and 0.1 mg ml^{-1} cycloheximide (CHX), and incubated at 37 °C with 5% CO_2 for 10 min. The cells were rinsed with pre-chilled PBS containing 0.1 mg ml^{-1} CHX 3 times and collected with the pre-chilled PBS containing 0.1 mg ml^{-1} CHX.

The primary procedures followed the ribosome profiling protocol[65] and were conducted by Guangzhou Epibiotek, incorporating the use of RNase I to enhance base resolution during treatment of the cell lysates. RNA fragments between 26 and 32 nucleotides in length were isolated via urea–PAGE and utilized for library construction.

Analysis of Ribo-seq data

Raw sequencing reads had their adaptors trimmed and low-quality bases were trimmed using Trim_galore[60] along with standard settings. Reads mapping to rRNA and tRNA were removed, and the remaining reads were mapped to the human genome (hg38, GENCODE v44). The mapping was performed using bowtie[66] (version 1.3.0) with the following parameters to allow one mismatch and acquire uniquely mapped reads: '-m 1 -v 1 --best –strata'. In the reference annotation file, only the annotated transcript containing the longest coding sequence (CDS) for each gene was utilized for further evaluation. RiboWaltz[67] (version 1.2.0) was used to characterize the P-sites for each ribosome-protected fragment.

To identify transcripts that have undergone stop codon readthrough, we used RRTS[40]. RRTS represents a ratio of the ribosome density between the natural stop codon and the next in-frame stop codon to the ribosome density in the CDS. To increase the read depth of the initial samples, biological replicates were pooled during the comparison of RRTS between samples. A Mann–Whitney *U* test was utilized to assess significant differences in RRTS values across diverse samples. The genes that experience a twofold increase in RRTS after the treatment of GCE or RCE tools are defined as potential off-targets to conduct GO enrichment analysis. To alleviate background noise, we additionally require that the RRTS of off-target genes must exceed 0.02. GO enrichment analysis was performed using R package clusterProfiler[68] (v4.8.3). We listed the transcripts with off-target UGA codon readthrough events by ribosome profiling in Supplementary Table 6.

Quantitative tyrosine phosphoproteomic profiling

Cells were collected in lysis buffer, which contained 50 mM Tris-HCl, pH 8.5, 7 M urea, 1% Triton X-100, 7 M urea, 1 mM PMSF, protease inhibitors mixture (RHAWN) and phosphatase inhibitor mixtures (HY-K0022). The protein concentration was measured by the BCA protein assay kit (Pierce). Cellular proteins were precipitated by the chloroform-methanol method, and were solubilized in 50 mM Tris-HCl buffer (with 8 M urea, pH 8.5) with sonication. The soluble proteins were reduced with DTT, alkylated with iodoacetamide and digested with trypsin (Sigma Aldrich). The digested proteins were acidified by 10%TFA to pH 2-3, purified by the CAE-Ti-IMAC (J&K), and further enriched by the pTyr superbinder[48,69]. The obtained peptides were desalted and identified with LC–MS/MS (Orbitrap Fusion Lumos Tribid LC–MS).

The phosphoproteomic data were processed following reported protocol[70], using Maxquant (version 2.6.5) with the *Homo sapiens* (organism_id:9606) proteome databases (as UP000005640) downloaded from Uniprot (https://www.uniprot.org). The phosphotyrosine (pTyr) proteomic data were processed as previous study. We first removed the pTyr sites from reverse sequences and potential contaminants or showing localization probability less than 0.75. The intensities were $\log_2$-transformed, and the Pearson correlation values were calculated. Data imputation was performed using Perseus software (version 4.8.5) with parameters: downshift = 1.8 and width = 0.3. pTyr sites with $\log_2$ fold change greater than one and analysis of variance adjusted P value were kept as significantly changed pTyr sites.

Proteomic profiling the mis-incorporations

HEK293T cells were seeded into 10-cm dishes and incubated for 16–24 h as described above. For RCE samples, 40 μg of *SRC*$^{K295*/Y527F}$–GFP reporter, 10 μg of gsno-SRC, 5 μg of DKC1-*iso3*, and 40 μg of TetRS-2×(ΨGA)-tRNA^{Tet} plasmids were transfected into four dishes using Lipofectamine LTX reagent (Invitrogen) supplemented with PLUS reagent (Invitrogen) alongside 200 μM CbzK. For GCE samples, 40 μg of *SRC*$^{K295*/Y527F}$–GFP reporter, 10 μg of gctrl, 5 μg of DKC1-vector, and 40 μg of TetRS-2×tRNA^{Tet} (wild-type tRNA^{Tet}) plasmids were transfected into 4 dishes using Lipofectamine LTX reagent (Invitrogen) supplemented with PLUS reagent (Invitrogen) alongside 200 μM CbzK. At 48 h post-transfection, the cell were collected with modified RIPA buffer following previous protocol[51]. The proteins in cell lysate were denatured with PBS buffer (containing 8 M urea), reacted with TCO-biotin probe (Confluore) for 2 h, precipitated with chloroform-methanol method. The proteins were redissolved in 2% SDS with sonication, diluted with PBS to final SDS concentration at 0.2%, and incubated with Streptavidin agarose resin beads (Thermo, 20353). The peptides on beads were reduced with DTT, alkylated with iodoacetamide, digested with trypsin, desalted and identified under data-independent acquisition mode via LC–MS/MS (Bruker, TIMS-TOF Pro2).

Data were analysed with DIA-NN software (version 1.9.2) under default parameters[71], with the *H. sapiens* (organism_id:9606) proteome database (as UP000005640) downloaded from Uniprot (https://www.uniprot.org)[70]. Proteins were filtered to include only those identified in both replicates of at least one condition. The filtered proteins were then used as input for DEP Bioconductor package (version 1.14.0)[72] to conduct differential enrichment analysis. Missing values were imputed using the MinProb method with $q = 0.01$.

Proteins were considered differentially enriched if they exhibited $\log_2$ fold changes greater than 1 relative to the control samples, with an adjusted P value < 0.05.

Detecting tRNA aminoacylation

Total RNA was extracted using TRIzol reagent (Invitrogen), chloroform (TGREAG), 3 M sodium acetate (pH 4.5) and ethanol (TGREAG). The RNA samples were dissolved in 0.1 M sodium acetate (pH 4.5), mixed with 2× RNA loading buffer (0.1 M sodium acetate pH 4.5, 90% glycerol, 0.03% xylene cyanide, and 0.03% bromophenol blue), and fractionated by electrophoresis at a constant 18 W on a 10% urea–PAGE gel (0.1 M sodium acetate, pH 4.5) at 4 °C. Each blots contain the control tRNA samples from cells which were transfected with wild-type $tRNA^{Pyl}$ or decoder tRNA but lack ncAA. We use the dye indicator, Xylenecyanol FF, which is reported to indicate 55-nt RNA in the 10% urea–PAGE gel[73]. The RNA was then transferred to nylon membrane (MILLIPORE) which was then crosslinked using UVLink 1000 Crosslinker (Jena). The tRNA was detected by Digoxigenin (DIG)-labelled DNA probes (Sangon), and visualized with Anti-Digoxigenin-AP (Roche). DIG-labelled probes were as follows: *Mm*tRNA, CGGAAACCCCGGGAATCTAACCCGGCTGAACGGA; *Ma*tRNA, CGAGAGACCGGGGCGTCGAACCCCGCTGGCTAGG; *Mi*tRNA, CGAAGTGCCCGGGAGTTGAACCCGGCTGCCGTGG. Each probe was modified with DIG at both 5′ and 3′ ends.

Molecular dynamics simulation

The mRNA-tRNA–ribosome complex was built from crystal structures, with the mRNA and ribosome (Protein Data Bank (PDB) ID: 4JYA) and tRNA (PDB ID: 5UD5) from crystal structures. Molecular dynamics workflows were implemented using AmberTools (version 23) for force field parameter assignment and system setup[74], OpenMM (version 8.1.2) for simulation execution[75], and MDTraj (version 1.9.8) for trajectory analysis[76]. The simulation system utilized Amber99sb*-ildn[77] for proteins, OL3 (ref. 78) for nucleic acids, GAFF2 (ref. 79) with AM1-BCC charges for the cofactor, and Amber parameters with RESP charges[80] for Ψ (derived from HF/6-31 G* calculations). The structure was then solvated using TIP3P waters[80] in a 12 Å buffer supplemented with 0.2 M $MgCl_2$, followed by system preparation through energy minimization and sequential equilibration stages.

Initial equilibration comprised 40 ns constant-temperature, constant-pressure (*NPT*; number of particles *N*, pressure *P*, temperature *T*) ensemble simulation under 100 kJ mol^{-1} $Å^{-1}$ restraints applied to the ribosome–mRNA-tRNA complex. Subsequent relaxation was achieved through three consecutive 20 ns NPT simulations with progressively reduced restraints of 50, 20, 5 kJ/mol/Å. In the production phase, 50 ns NPT simulation was performed for tRNA while maintaining 5 kJ/mol/Å restraints on the ribosome–mRNA backbone. The 5 kJ/mol/Å restraints on the ribosome–mRNA backbone were not lifted because the mRNA provided in the crystal structure contains only 5 bases. Additional restraints were necessary to maintain consistency with the crystal structure.

The tRNA-mRNA binding structure were extracted form mRNA-tRNA–ribosome complex, and performed restrained simulation based on the codon–anticodon binding geometries were aligned to the corresponding region in 4JYA. Solvation and equilibration followed established protocols, with production simulations extended to 200 ns.

Reporting summary

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

Data availability

All next-generation sequencing data generated for this study have been deposited in the NCBI Sequence Read Archive (SRA) under accession code PRJNA1090628. The human reference genome GRCh38 (December 2013) was downloaded from the following link: https://hgdownload.soe.ucsc.edu. The *H. sapiens* proteome database (UP000005640) was downloaded from https://www.uniprot.org. The custom database containing reporter sequences featuring 20 different natural amino acids at positions corresponding to the premature termination codon location is provided in Supplementary Table 1. The tRNA reference sequences were derived from GtRNAdb[81] (http://gtrnadb.ucsc.edu/; accessed September 2024). ISR-related genes were collected from the GeneCards database (https://www.genecards.org/; accessed September 2024). Source data are provided with this paper.

Code availability

Custom codes are available on GitHub (https://github.com/yanxueqing621/RCE_project). These scripts include the pipelines for off-target Ψ sites identification based on PRAISE sequencing data, evaluation of potential off-target readthrough events based on Ribo-seq data, and RNA-seq data analysis.

58. Lorenz, R. et al. ViennaRNA Package 2.0. *Algorithms Mol. Biol.* **6**, 26 (2011).
59. Gruber, A. R., Lorenz, R., Bernhart, S. H., Neuböck, R. & Hofacker, I. L. The Vienna RNA Websuite. *Nucleic Acids Res.* **36**, W70–W74 (2008).
60. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10–12 (2011).
61. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**, 907–915 (2019).
62. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2013).
63. Shen, W., Le, S., Li, Y. & Hu, F. SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLoS ONE* **11**, e0163962 (2016).
64. Smith, T., Heger, A. & Sudbery, I. UMI-tools: modeling sequencing errors in unique molecular identifiers to improve quantification accuracy. *Genome Res.* **27**, 491–499 (2017).
65. Ingolia, N. T., Brar, G. A., Rouskin, S., McGeachy, A. M. & Weissman, J. S. The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments. *Nat. Protoc.* **7**, 1534–1550 (2012).
66. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
67. Lauria, F. et al. riboWaltz: optimization of ribosome P-site positioning in ribosome profiling data. *PLoS Comp. Biol.* **14**, e1006169 (2018).
68. Wu, T. et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. *Innovation* **2**, 100141 (2021).
69. Dong, M. et al. Sensitive, robust, and cost-effective approach for tyrosine phosphoproteome analysis. *Anal. Chem.* **89**, 9307–9314 (2017).
70. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* **26**, 1367–1372 (2008).
71. Demichev, V., Messner, C. B., Vernardis, S. I., Lilley, K. S. & Ralser, M. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat. Methods* **17**, 41–44 (2020).
72. Shaw, J. J. & Green, R. Two distinct components of release factor function uncovered by nucleophile partitioning analysis. *Mol. Cell* **28**, 458–467 (2007).
73. Gagnon, K. T. & Maxwell, E. S. Electrophoretic mobility shift assay for characterizing RNA–protein interaction. *Methods Mol. Biol.* **703**, 275–291 (2011).
74. Case, D. A. et al. AmberTools. *J. Chem. Inf. Model.* **63**, 6183–6191 (2023).
75. Eastman, P. et al. OpenMM 8: molecular dynamics simulation with machine learning potentials. *J. Phys. Chem. B* **128**, 109–116 (2024).
76. McGibbon, RobertT. et al. MDTraj: a modern open library for the analysis of molecular dynamics trajectories. *Biophys. J.* **109**, 1528–1532 (2015).
77. Lindorff-Larsen, K. et al. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins Struct. Funct. Bioinformatics* **78**, 1950–1958 (2010).
78. Zgarbová, M. et al. Refinement of the Cornell et al. nucleic acids force field based on reference quantum chemical calculations of glycosidic torsion profiles. *J. Chem. Theory Comput.* **7**, 2886–2902 (2011).
79. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **25**, 1157–1174 (2004).
80. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
81. Chan, P. P. & Lowe, T. M. GtRNAdb 2.0: an expanded database of transfer RNA genes identified in complete and draft genomes. *Nucleic Acids Res.* **44**, D184–D189 (2016).

Acknowledgements The authors thank the National Center for Protein Sciences at Peking University in Beijing, China, for assistance with the 4150 TapeStation System (G. Li), mass spectrometry (D. Liu and Q. Zhang), Zeiss LSM 980 confocal microscope and Spin SR confocal microscope (L. Fu); G. Jia, R. Liu, F. Lin, X. Rao, C. Shao, M. Zhang and Y. Ma for discussions and materials; the Center for Quantitative Biology at Peking University for assistance with the ImageXpress Micro 4 high-content imaging system; and X. Li for help. Part of the analysis was performed on the High-Performance Computing Platform of the Center for Life Science

(Peking University). We acknowledge funding from the National Natural Science Foundation of China (22137001 to P.R.C., 22337001 to C.Y., 92353000 to P.R.C. and 22425071 to C.Y.), the Ministry of Science and Technology (2023YFA1506500 and 2021YFA1302600 to P.R.C.; 2023YFC3402200 to C.Y.), Beijing Municipal Science and Technology Commission Project (Z231100002723005 to C.Y.), New Cornerstone Science Foundation (to C.Y.) through the XPLORER PRIZE, as well as New Cornerstone Science Foundation (to P.R.C.) through the New Cornerstone Investigator Program. This work was also supported by the Ministry of Agriculture and Rural Affairs of China.

Author contributions P.R.C., C.Y. and J.L. conceived the research. J.L. guided, designed and performed the experiments. X.Y. analysed the sequencing data. H. Wu analysed and organized the fluorescence data. Z.J. and Y.S. contributed in the screening, immunoprecipitation and proteomic experiments. X.W., R.G. and H. Wen conducted molecular dynamics simulation. Y.R. assisted the experiments in screening. C.L. performed acid-denaturing northern blots. Y.Z. and Z.J. guided the phosphoproteomic profiling. Y. M. conducted the tRNA sequencing and PRAISE sequencing experiments. All authors commented and approved the paper.

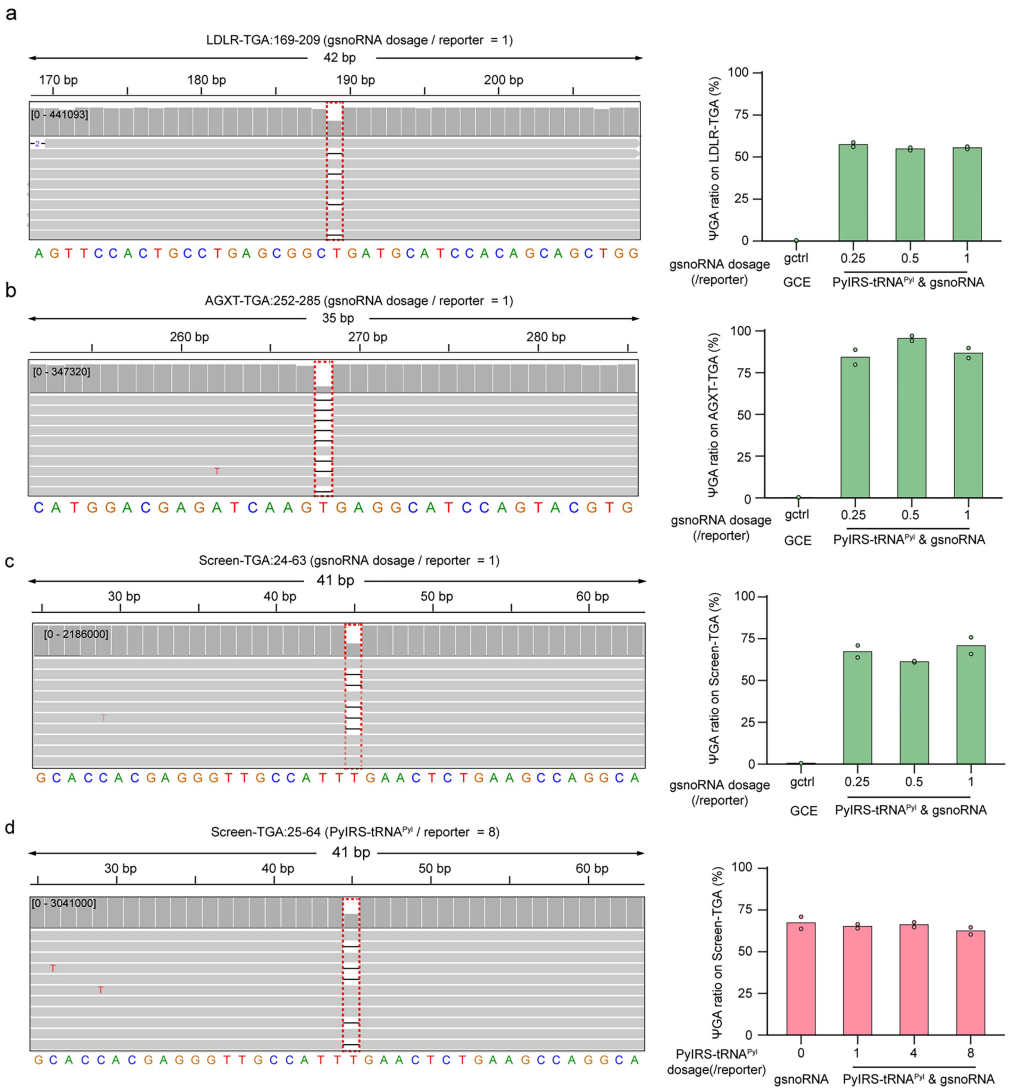
Competing interests C.Y. is an inventor on patents related to the RESTART technology (PCT/CN2022/095172). The other authors declare no competing interests.

Additional information

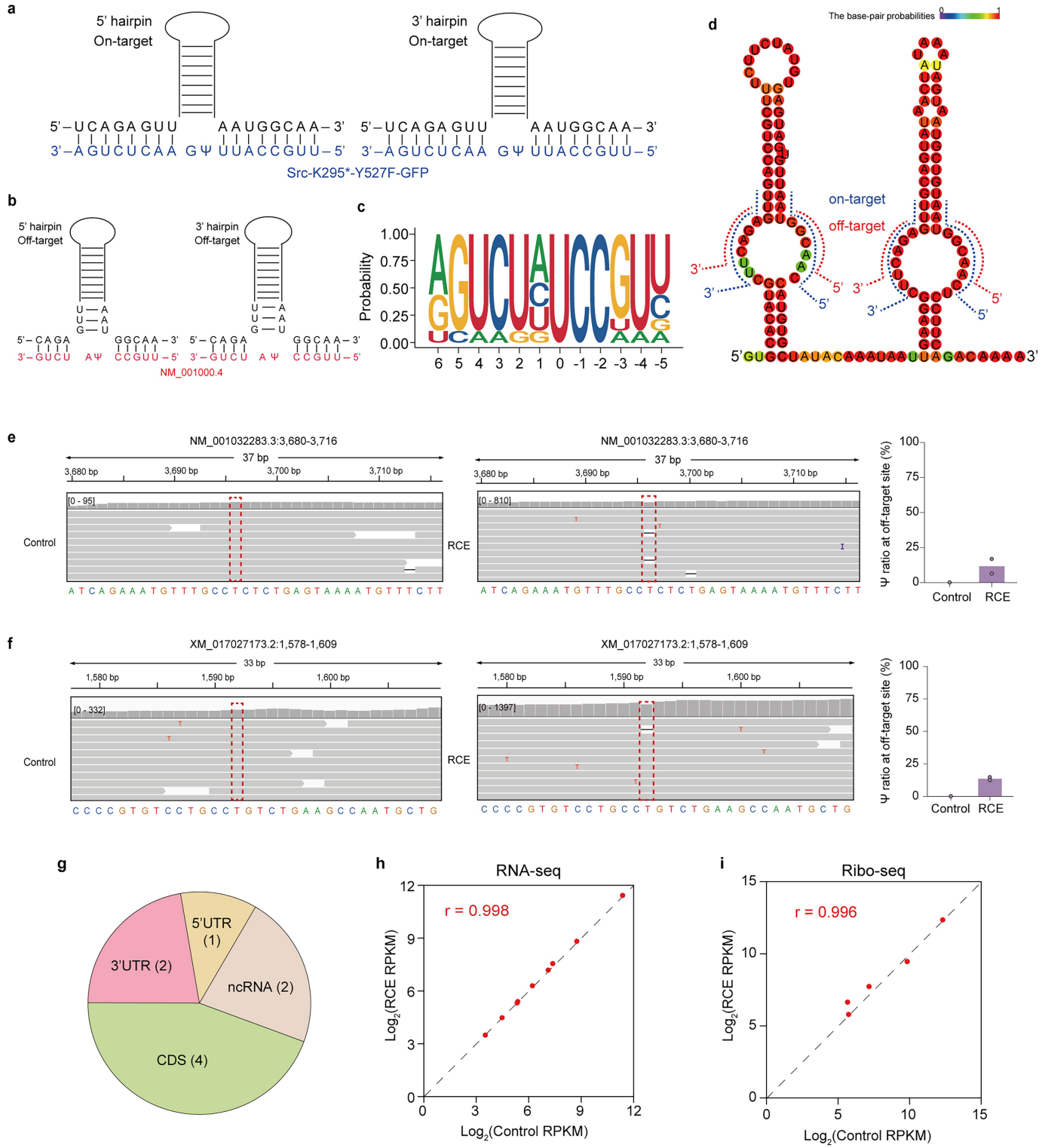
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-025-09165-x.

Correspondence and requests for materials should be addressed to Chengqi Yi or Peng R. Chen.

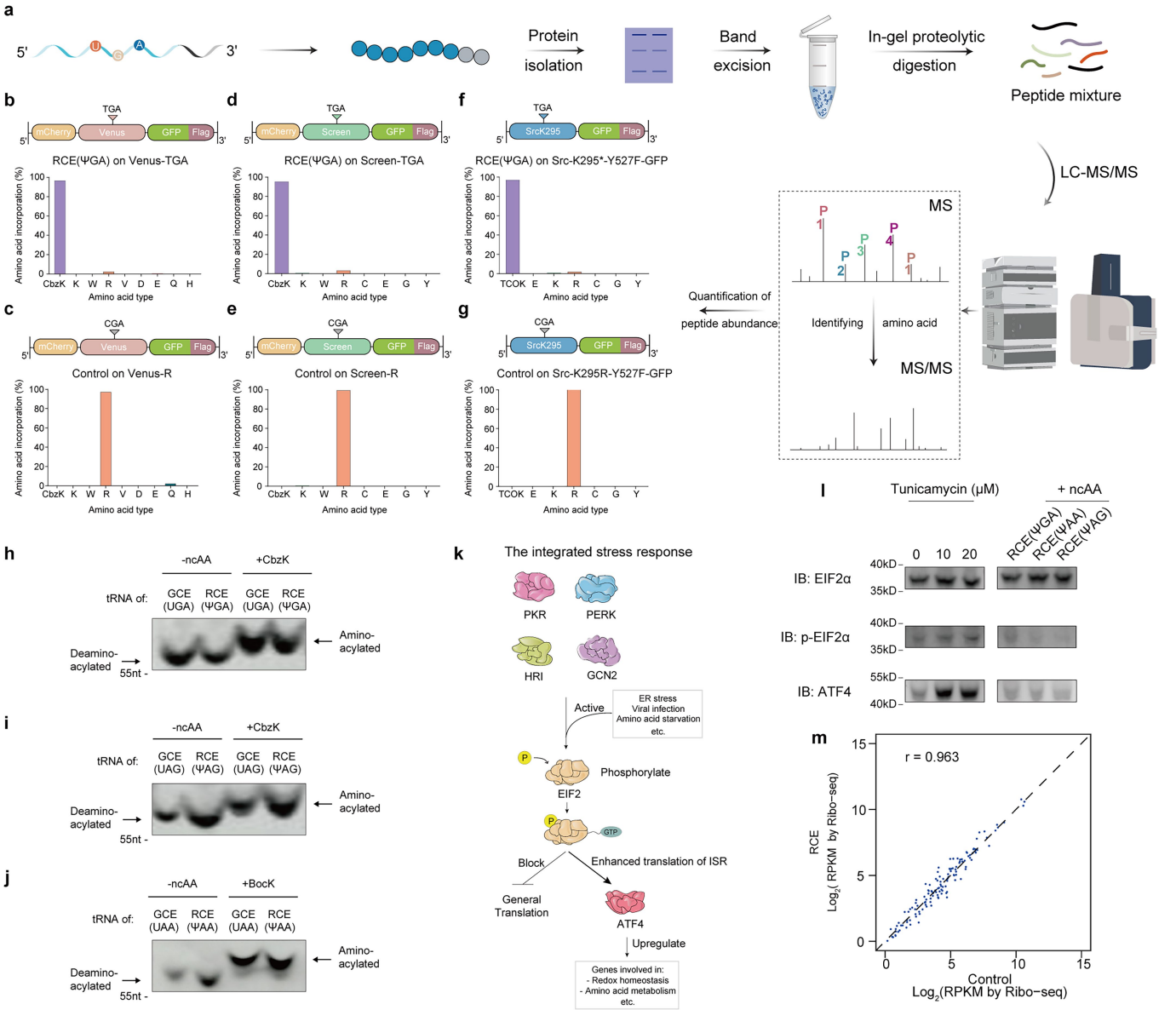
Peer review information *Nature* thanks Kathrin Lang and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.



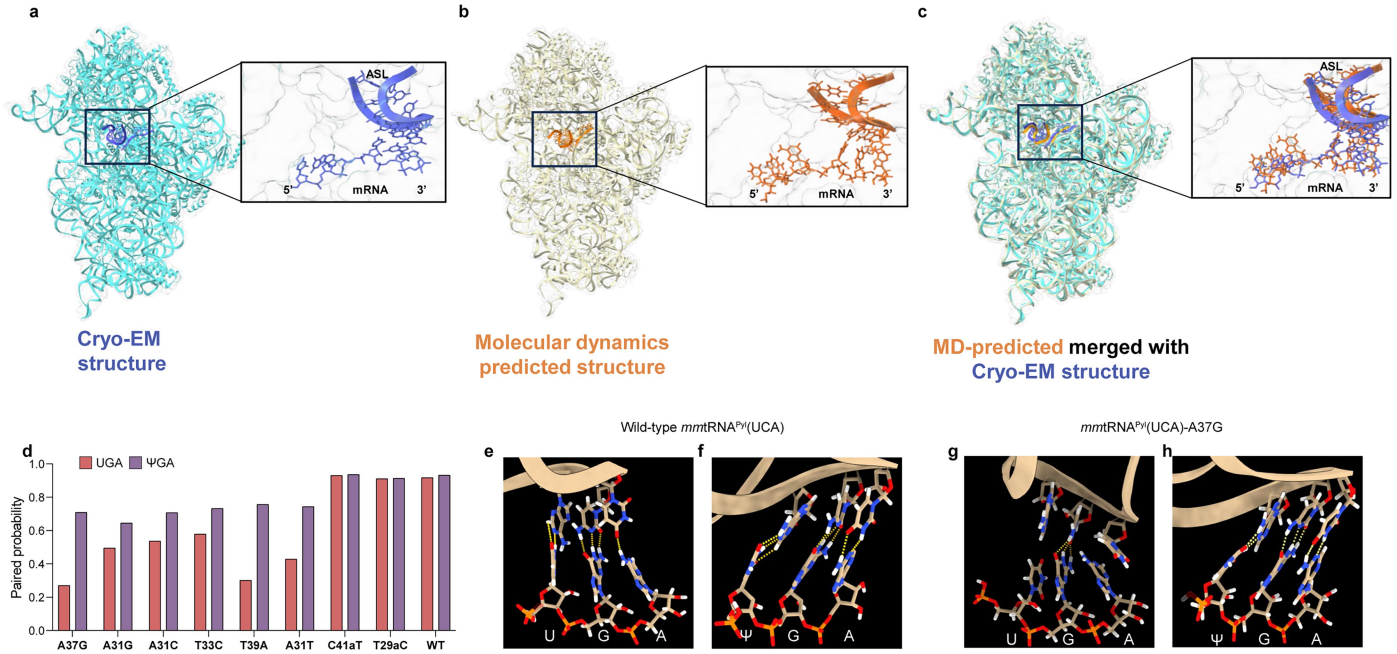
Extended Data Fig. 1 | The yields of ΨGA codons within targeted mRNAs under different dosage of gsnoRNA and PylRS-tRNA^{Pyl}. a-c, The IGV views and ΨGA codon yields within three targeted reporters: LDLR-TGA **(a)**, AGXT-TGA **(b)**, and Screen-TGA **(c)** at 0%, 25%, 50%, and 100% of the corresponding guide snoRNA. **d**, The IGV views and ΨGA codon yields within targeted Screen-TGA reporter under zero, 1:1, 1:4 or 1:8 stoichiometry to PylRS-tRNA^{Pyl}. Data of bar plots are presented as the mean values (n = 2 biologically independent replicates).



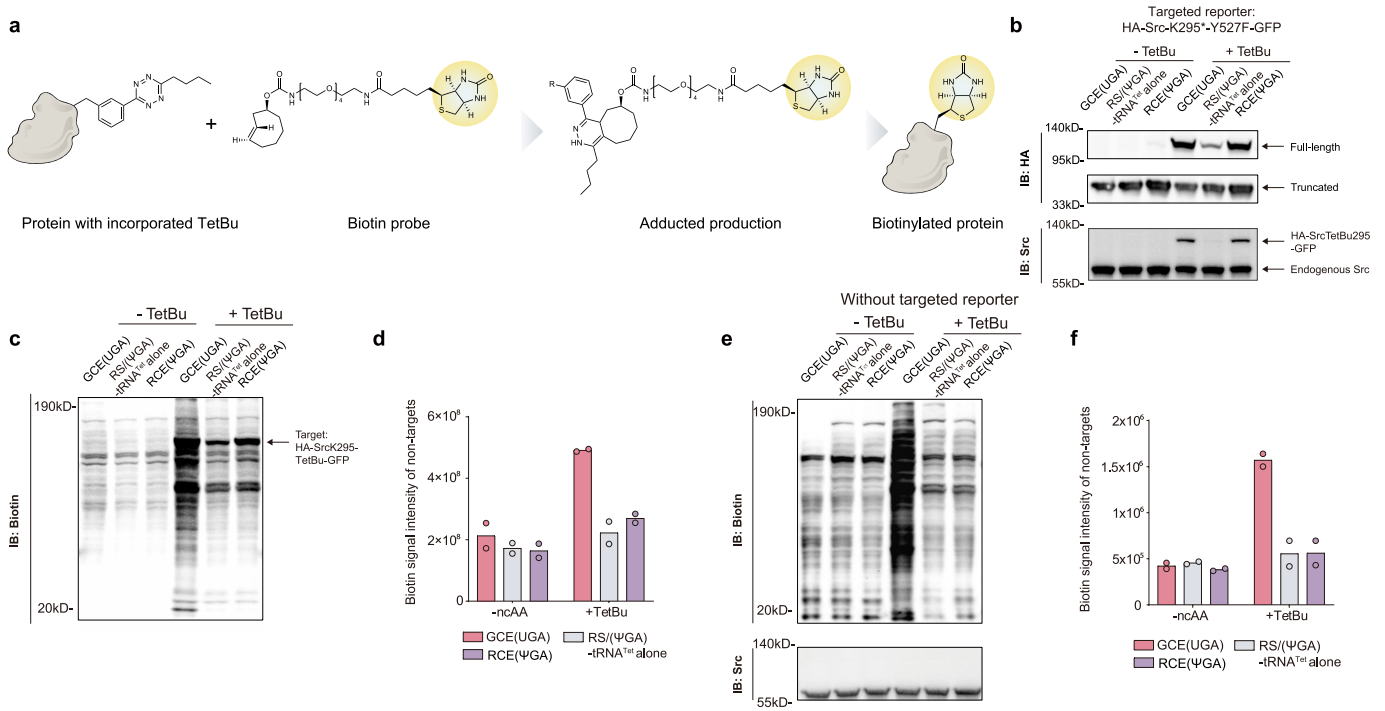
Extended Data Fig. 2 | High encoding specificity of RCE strategy verified by the limited off-target Ψ edits as well as low off-target Ψ editing ratios across the transcriptome. a, Complementary regions between gsnoRNA (black) and target site in SrcK295TGA-GFP reporter (blue). **b**, Complementary regions between gsnoRNA (black) and an off-target site in the endogenous gene (red). **c**, Sequence motif of the RCE-induced off-target Ψ sites. **d**, The on-target and off-target regions within the secondary structure of the gsnoRNA-Src predicted by RNAfold. The base-pair probabilities are indicated with color. Red and blue colors represent high and low probabilities, respectively. **e**,**f** IGV views representing the reads mapped at the off-target sites in 3' UTRs of NM_001032283.3 **(e)** and XM_017027173.2 **(f)**. **g**, Pie chart showing the distribution of RCE-induced off-target Ψ sites in the transcriptome. **h**, Scatterplot showing the expression level of genes containing the off-target Ψ sites by RNA-seq. Data are represented as the mean values (n = 2 biologically independent replicates). **i**, Scatterplot showing the translation level of genes containing the off-target Ψ sites by Ribo-seq. Data are represented as the mean values (n = 2 biologically independent replicates). The correlations between different conditions were measured via Pearson's correlation score by defaulted parameters.



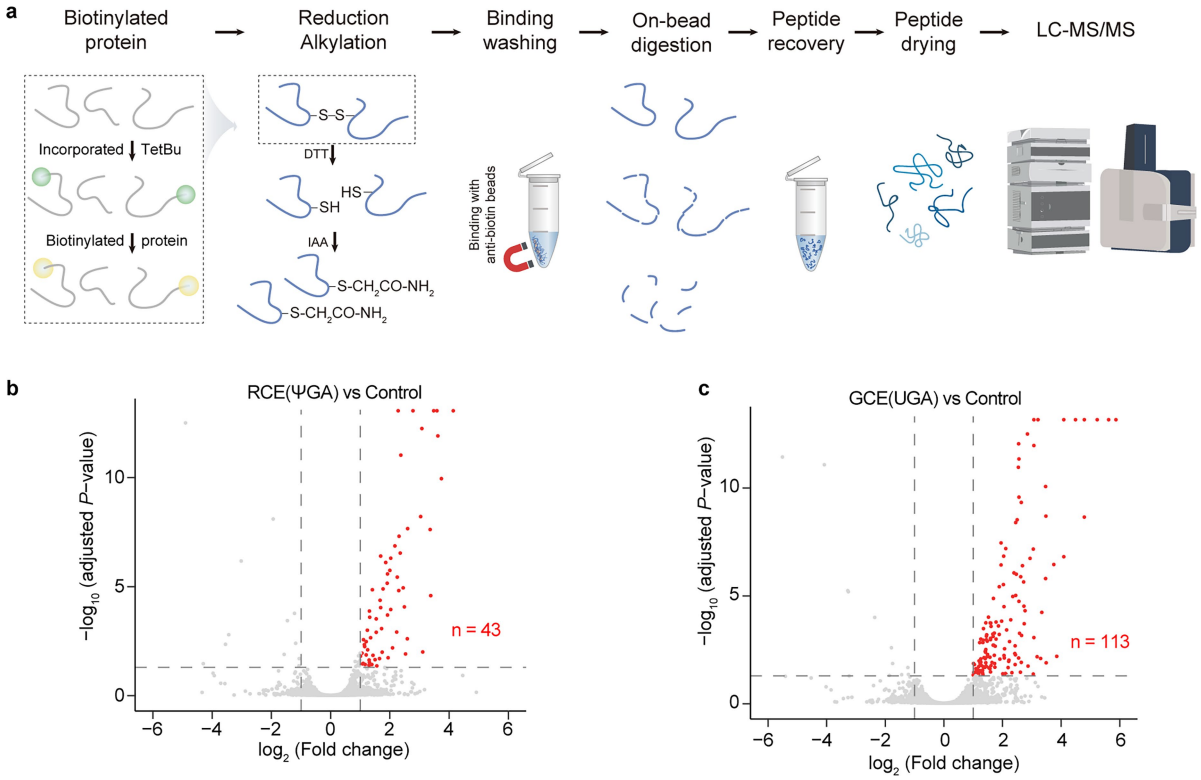
Extended Data Fig. 3 | Evaluating the ncAA incorporation fidelity of the RCE system. a, Schematic illustration of the workflow for identifying peptides with different amino acid incorporation via LC-tandem MS. **b-g**, The RCE system allowed site-specific incorporation of ncAAs with high fidelity on the Venus-TGA **(b)**, Screen-TGA **(d)**, and Src-K295*-Y527F-GFP **(f)** reporters. The corresponding Venus-R **(c)**, Screen-R **(e)**, and Src-K295R-Y527F-GFP **(g)** reporters were used for comparison of natural amino acid incorporation. **h-j**, Acid denaturing Northern Blots illustrating that the decoder tRNAs were highly charged with ncAA in all three RCE systems: RCE (ΨGA) **(h)**, RCE (ΨAG) **(i)**, RCE (ΨAA) **(j)**. Each experiment was independently repeated twice with similar results. **k**, Diagram of the integrated stress response pathway. **l**, Western blotting demonstrating that the integrated stress response was not activated in the RCE systems. Tunicamycin treatment is used as a control of ISR activation. Each experiment was independently repeated twice with similar results. **m**, Scatterplot showing the translation level of ISR-related genes measured by Ribo-seq. Data are represented as the mean values (n = 2 biologically independent replicates).



Extended Data Fig. 4 | MD simulation-based prediction supports the ΨGA preference of the screened *mm*tRNA^{Pyl}(UCA)-A37G variant. a-c, Structural comparison among the ribosome-tRNA-mRNA complex from cryo-EM (PDB ID: 4JYA) **(a)**, and all-atom MD simulation **(b)**, along with the merged complex structures between cryo-EM and MD **(c)**. The results show that the tRNA and mRNA align well between the structures from cryo-EM and MD simulation. **d**, The paired probability of the averaged trajectories demonstrating the binding between tRNA variants and UGA- or ΨGA-containing mRNAs. **e-h**, Centroids of the tRNA-mRNA paring structures of the wild-type *mm*tRNA^{Pyl}(UCA) **(e,f)** and the selected *mm*tRNA^{Pyl}(UCA)-A37G **(g,h)**, illustrating that *mm*tRNA^{Pyl}(UCA)-37G forms more stable hydrogen bonds to ΨGA codon than to UGA codon.



Extended Data Fig. 5 | The RCE system exhibited a high translational specificity when the reporter construct was expressed at an endogenous level. a, Incorporated tetrazine-containing ncAA–TetBu could be labeled with TCO-biotin via bioorthogonal ligation. **b**, Western blotting of N-terminal HA-tagged products demonstrating that the protein product of the on-target reporter was expressed at endogenous concentration. RS/(ΨGA)-tRNA^{Tet} alone was used to show the readthrough of tRNA^{Tet} on the UGA codon of targeted reporter. The experiment was independently repeated twice with similar results. **c,d**, Western blotting (**c**) and quantitative analysis (**d**) showing the biotin signal for the GCE and RCE systems, demonstrating that the RCE system exhibited a low mis-incorporation level close to the background levels of endogenous biotinylated proteins. Data of bar plots are presented as the mean values (n = 2 biologically independent experiments). **e,f**, Western blotting (**e**) and quantitative analysis (**f**) comparing the off-target TetBu incorporation of the GCE(UGA), RS/(ΨGA)-tRNA^{Tet} alone, and RCE(ΨGA) systems, without the targeted reporter, further demonstrating the ncAA incorporation specificity of the RCE system. Data of bar plots are presented as the mean values (n = 2 biologically independent experiments).



Extended Data Fig. 6 | Proteome analysis demonstrated high translational specificity of the RCE(ΨGA) system. a, Schematic illustration of the affinity-based enrichment method for detecting the biotinylated peptides. **b,c**, Volcano plots showing the ncAA-incorporated endogenous proteins induced by the RCE(ΨGA) system **(b)** and the GCE(UGA) system **(c)**. Red dots represent ncAA-incorporated endogenous proteins. *P*-values were calculated by linear models and subsequently adjusted via the Benjamini-Hochberg method. X-axis is fold change of protein abundance in the GCE(UGA)- or RCE(ΨGA)-treated cells compared to the control cells.

nature portfolio

Corresponding author(s): Chengqi Yi and Peng R. Chen

Last updated by author(s): Apr 14, 2024

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	Next-generation sequencing data was collected and demultiplexed by Illumina NovaSeq and BGI MGI platform. Cells were imaged using the ImageXpress Micro 4 high-content imaging system. Subcellular colocalization were imaged using Evident(Olympus) Spin SR microscope. The blots were imaged using ChemiDoc MP Imaging System (Biorad). Proteomic data were collected using LC-MS/MS (Orbitrap Fusion Lumos Tribid LC-MS) and LC-MS/MS (Bruker, TIMS-TOF Pro2).
Data analysis	RNAfold (online version) Trim_galore (version 0.6.6) HISAT2 (version 4.8.5) featureCount (version 2.0.4) bowtie (version 1.3.0) RiboWaltz (version 1.2.0) clusterProfiler (version 4.8.3) cutadapt (version 2.10) Seqkit (version 0.14.0) PRAISE tool (version 4.8.5) Perseus (version 2.1.3.0) DIA-NN (version 1.9.2) umi_tools (version 1.0.0) OpenMM (version 8.1.2)

Ambertools (version 23)
MDtraj (version 1.9.8)
MetaXpress (MX version 6.2.3.733;CME Version 6.2.3.991)
Fiji(64-bit for windows, released at May 30, 2017)
The custom codes are provided to public in the Github(http://github.com/yanxueqing621/RCE_project).

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 next-generation sequencing data generated for this study has been deposited in the NCBI Sequence Read Archive, under accession code PRJNA1090628. The reference genome GRCh38 was download from the following link: https://hgdownload.soe.ucsc.edu/downloads.html. UniProt database was download from the following link: https://www.uniprot.org. The tRNA reference sequences were derived from GtRNAdb(http://gtrnadb.ucsc.edu/). The customed protein databases of proteins with different amino acids were provided in the Supplementary Table 1, whcih should be used together with the human proteome database above in case of quality control.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	Not involving animals or humans
Reporting on race, ethnicity, or other socially relevant groupings	Not involving animals or humans
Population characteristics	Not involving animals or humans
Recruitment	Not involving animals or humans
Ethics oversight	Not involving animals or humans

Note that full information on the approval of the study protocol must also be provided in the manuscript.

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	No statistical methods were used to predetermine sample size. For experiments with analysis, a minimum of 2 technical replicates were performed to confirm reproducibility.For the sequencing experiments, the sample size was determined based on our prior experiences (Song et al., Mol Cell. 2023(PMID: 36521489);Zheng et al, Nat Methods. 2015 (PMID: 26214130)). For the western blotting experiments, we found the two biological replicates are similar; if there is a need to quantify the graylevel, we use three biological replicates for statistic analysis (Weng et al., Nat Chembiol. 2024(PMID:38167916)). For statistic analysis of localization, we have used a sample size of 50 for quantitation, which is above the published sample size used for the same purpose (Li et al., Nat Chem.,2024(39501047); Gautier et al.,J.A.C.S., 20218600).
Data exclusions	No data was excluded.
Replication	To verify the reliability of the results, we performed experiments in different cell lines (HEK293T, CHO-K1, COS-7, and U-2 OS),and evaluated via multiple methods (fluorescence, sequencing, and proteomic profiling). We also performed 2 or more independent experiments and all atempts to replicate data were successful.

Randomization	Because the sample size is small, randomization was not relevant for this study. Covariates were controlled for by running controls in parallel whenever applicable.
Blinding	Blinding was not performed as experimental conditions were evident. All samples were treated identically through standard procedures.

Behavioural & social sciences study design

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

Study description	
Research sample	
Sampling strategy	
Data collection	
Timing	
Data exclusions	
Non-participation	
Randomization	

Ecological, evolutionary & environmental sciences study design

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

Study description	
Research sample	
Sampling strategy	
Data collection	
Timing and spatial scale	
Data exclusions	
Reproducibility	
Randomization	
Blinding	

Did the study involve field work? ☐ Yes ☐ No

Field work, collection and transport

Field conditions	
Location	
Access & import/export	
Disturbance	

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
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

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

Antibodies

Antibodies used

ANTI-FLAG® M2 magnetic beads (Sigma Aldrich, M8823);
mouse anti-Flag (Sigma-Aldrich, F1804, 1:2000 or 1:750 dilution);
Anti β-Actin Mouse Monoclonal (clone name: 3E8) Antibody (CWBIO, CW0096M, 1:5000 dilution);
Phospho-Tyrosine (P-Tyr-1000) MultiMab® Rabbit mAb mix (Cell Signaling Technology, #8954, 1:2000 or 1:750 dilution)
Goat Anti-Mouse IgG, HRP Conjugated (CWBIO, CW0102, 1:5000 dilution);
Goat Anti-Rabbit IgG, HRP Conjugated (CWBIO, CW0103, 1:5000 dilution);
Phospho-Src Family (Tyr416) (E6G4R) Rabbit mAb #59548(CST,1:1000 dilution);
eIF2α Antibody #9722 (CST, 1:1000 dilution);
Phospho-eIF2α (Ser51) (D9G8) XP® Rabbit mAb #3398 (CST, 1:1000 dilution);
ATF-4 (D4B8) Rabbit mAb #11815 (CST, 1:1000 dilution);
Biotin Antibody (33): sc-101339 (Santa Cruz, 1:2000 dilution);
p53 Antibody #9282 (CST, 1:1000 dilution);
Phospho-p53 (Ser15) Antibody #9284 (CST, 1:1000 dilution).

Validation

ANTI-FLAG® M2 magnetic beads: validated by manufacturer by enrichments of two different FLAG®-tagged baits either from a HeLa-S3 nuclear extract (Nuc. Extr.) or from a HeLa-S3 cytoplasmic fraction (Cyto. Extr.).
mouse anti-Flag: validated by manufacturer by dot blot against FLAG-BAP fusion protein;
Anti β-Actin Mouse Monoclonal (clone number: 3E8) Antibody: validated by manufacturer by western blotting against whole cell lysates from Hela cell line;
Phospho-Tyrosine (P-Tyr-1000) MultiMab® Rabbit mAb mix: validated by manufacturer by Western blot analysis of extracts from A-431 cells, untreated (-) or treated with Human Epidermal Growth Factor (hEGF) #8916 (100 ng/ml, 5 min; +)
Goat Anti-Mouse IgG, HRP Conjugated & Goat Anti-Rabbit IgG, HRP Conjugated: Conjugates have been been specifically tested and qualified for Western blot and ELISA assay applications by manufacturer;
Phospho-Src family (Tyr416): validated by western blot analysis of extarcts from Colo 201 cells, untreated (-) or treated with Dasatinib #9052 (200nM, 24hr;+) by the manufacturer;
eIF2a antibody #9722 AND phospho-eIF2a (Ser51)#3398: validated by western blot analysis of extracts from PC12 cells, untreated or thapsigargin-treated by manufacturer;
ATF4 (D4B8): validated by western blot analysis of extracts from 293 and HeLa cells, untreated (-) or tunicamycin-treated;
Biotin Antibody (33): validated by western blot analysis of biotin conjugated goat IgG and normal goat IgG by manufacturer;
p53 antibody #9282 was validated by western blot analysis of cell extracts and simpleChIP Enzymatic Chromatin IP Kits by the manufacturer.
Phospho-p53 (Ser15) Antibody #9284 was validated by western blot analysis of a p53 fusion protein, untreated or phosphorylated by DNA-PK, using Phospho-p53 (Ser15) Antibody (upper) and p53 Antibody #9282 (lower)by the manufacturer.

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)

source: ATCC;
cell lines used: HEK293T cells (ATCC CRL-3216), U-2 OS (HTB-96), Hela (CCL-2), COS-7 (CRL-1651),CHO-K1(CCL-61)

Authentication

No method of cell line authenticated was used.

Mycoplasma contamination

Cells tested negative for mycoplasma as detailed in the Methods.

Commonly misidentified lines
(See ICLAC register)

No commonly misidentified cell lines were used.

Palaeontology and Archaeology

Specimen provenance	
Specimen deposition	
Dating methods	

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Ethics oversight	

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals	
Wild animals	
Reporting on sex	
Field-collected samples	
Ethics oversight	

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	
Study protocol	
Data collection	
Outcomes	

Dual use research of concern

Policy information about dual use research of concern

Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

No	Yes	
☐	☐	Public health
☐	☐	National security
☐	☐	Crops and/or livestock
☐	☐	Ecosystems
☐	☐	Any other significant area

Experiments of concern

Does the work involve any of these experiments of concern:

No	Yes	
☐	☐	Demonstrate how to render a vaccine ineffective
☐	☐	Confer resistance to therapeutically useful antibiotics or antiviral agents
☐	☐	Enhance the virulence of a pathogen or render a nonpathogen virulent
☐	☐	Increase transmissibility of a pathogen
☐	☐	Alter the host range of a pathogen
☐	☐	Enable evasion of diagnostic/detection modalities
☐	☐	Enable the weaponization of a biological agent or toxin
☐	☐	Any other potentially harmful combination of experiments and agents

Plants

Seed stocks

Novel plant genotypes

Authentication

ChIP-seq

Data deposition

☐ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links
May remain private before publication.

Files in database submission

Genome browser session
(e.g. UCSC)

Methodology

Replicates

Sequencing depth

Antibodies

Peak calling parameters

Data quality

Software

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

Instrument

Software

Cell population abundance

Gating strategy

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

Magnetic resonance imaging

Experimental design

Design type

Design specifications

Behavioral performance measures

Acquisition

Imaging type(s)

Field strength

Sequence & imaging parameters

Area of acquisition

Diffusion MRI ☐ Used ☐ Not used

Preprocessing

Preprocessing software

Normalization

Normalization template

Noise and artifact removal

Volume censoring

Statistical modeling & inference

Model type and settings

Effect(s) tested

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference

(See Eklund et al. 2016)

Correction

Models & analysis

n/a	Involved in the study
☐	☐ Functional and/or effective connectivity
☐	☐ Graph analysis
☐	☐ Multivariate modeling or predictive analysis

Functional and/or effective connectivity

Graph analysis

Multivariate modeling and predictive analysis