MOLECULAR BIOLOGY

NAT10 and N^4-acetylcytidine restrain R-loop levels and related inflammatory responses

Turja K. Debnath[1], Nathan S. Abell[1]†, Yi-Ru Li[1,2]†, Sravan K. Devanathan[1], Enrique Navedo[1], Blerta Xhemalçe[1,2]*

N^4-acetylcytidine (ac^4C) is deposited on diverse RNAs by *N*-acetyltransferase 10 (NAT10), a protein with high biological relevance for aging and cancer. We performed a comprehensive survey of ac^4C using metabolic labeling, sodium cyanoborohydride chemical treatment coupled to next-generation sequencing (NGS), and ac^4C antibody–based cell and molecular biology techniques. Our analysis shows that NAT10-dependent ac^4C-acetylation is robust in rRNA and specific tRNAs but low/spurious in mRNA. It also revealed an inflammatory signature and mutagenesis at transcriptionally active sites in NAT10-KO cells. This finding led us to explore the role of NAT10 in R-loops, which were recently linked to APOBEC3B-mediated mutagenesis. Our analysis showed that R-loops are ac^4C-acetylated in a NAT10-dependent manner. Furthermore, NAT10 restrains the levels of R-loops at a subset of differentially expressed genes in a catalytic activity–dependent manner. Together with cellular biology data showing ac^4C-modified RNA in endosomal structures, we propose that increased levels of ac^4C-unmodified RNAs, likely derived from R-loops, in endosomal structures induce inflammatory responses.

INTRODUCTION

Acetylation of biomolecules by specific enzymes is an essential process for proper cellular function. *N*-acetyltransferase 10 (NAT10) is a particularly important acetyltransferase responsible for both protein and RNA acetylation in mammalian cells. In RNA, NAT10 introduces an acetyl group to the free amine group at the N^4-position of cytosine (C) to form N^4-acetylcytidine (ac^4C) (*1*). This acetyl-group transfer is an adenosine triphosphate (ATP)–driven reaction, with the acetyl–coenzyme A (CoA) cofactor supplying the required acetyl group for ac^4C formation (*1*). This conserved RNA modification has been observed in transfer RNAs and ribosomal RNAs, the most abundant RNAs in cells (*2*). In humans, ac^4C has been detected in helix 34 (C1337) and helix 45 (C1842) of 18*S* ribosomal RNA (rRNA), as well as at position 12 of both $tRNA^{Ser}$ and $tRNA^{Leu}$ (*1*, *3*, *4*). An adaptor protein called THioUridine synthases, RNA Methylases and Pseudouridine synthases (THUMP) domain containing 1 (THUMPD1) is essential for ac^4C deposition on tRNA but is not necessary for 18*S* rRNA acetylation (*4*). In addition, the C/D box small nucleolar RNA SNORD13 facilitates rRNA acetylation by exposing the modification site on pre-rRNA through base-pairing (*4*, *5*). Recently, ac^4C in mRNAs was reported to fine-tune mRNA translation in a position-dependent manner (*6*, *7*), although its prevalence in human mRNA has been debated (*8*, *9*). NAT10 is a multifunctional enzyme that also mediates histone and microtubule acetylation and is involved in nuclear architecture (*10*–*12*).

NAT10 has been linked to aging and cancer. NAT10 inhibition or depletion reverts the aberrant nuclear shape of cells derived from the premature aging disorder Hutchinson Gilford progeria syndrome (HGPS) (*12*). Consistently, preclinical data demonstrated that targeting NAT10 in vivo, either via chemical inhibition or genetic manipulation, significantly enhanced the health span in an LmnaG609G HGPS mouse model (*13*). Furthermore, NAT10 is also up-regulated in several cancers, including hepatocarcinoma (*14*, *15*). However, the biological significance of RNA ac^4C acetylation in the aging and cancer functions of NAT10 remains unclear, mainly due to the lack of methods to robustly analyze this emerging RNA modification.

As mentioned above, sodium borohydride– and sodium cyanoborohydride–based approaches have been reported for base-resolution detection of ac^4C with seemingly disparate results [(*3*, *7*, *16*) and reviewed in (*17*)]. Therefore, we used multiple chemical and molecular biology approaches to unbiasedly decipher acetylation of RNA and its biological importance (Fig. 1). In our first approach, we used metabolic labeling with sodium 4-pentynoate to generate a "clickable" acetyl-CoA analog in cells, aiming to pull down acetylated or acylated RNA. Unfortunately, this method was not successful in detecting ac^4C RNAs (Fig. 1), as reported by another group while our paper was under review (*18*). In a second approach, we applied improved sodium cyanoborohydride ($NaCNBH_3$) chemistry–based transcriptome-wide sequencing methods to detect acetylation levels in diverse RNAs from HeLa wild type (WT) and NAT10-KO cell lines. In a third approach, we used ac^4C-specific immunofluorescence (IF) experiments to detect the spatial distribution of ac^4C RNAs. Together, our analyses suggest a molecular pathway through which NAT10 may regulate gene expression and maintain genome stability by acetylating R-loops and promoting their resolution. R-loops are non-B DNA structures formed by an RNA:DNA hybrid and the displaced ssDNA of the original DNA duplex (*19*), primarily occurring on GC-rich sequences during transcription. Their timely resolution is crucial for resuming productive transcription and preventing genome instability. Furthermore, we observe acetylated R-loop–derived products in endosomes, where they colocalize with the single-stranded RNA sensor Toll-like receptor 7 (TLR7) (*20*, *21*). In the absence of acetylation in NAT10-KO cells, the TLR7 pathway is activated, suggesting that NAT10 and RNA acetylation may suppress TLR7 activation. Our findings may have important biological implications in cancer, inflammation, and aging.

[1]Department of Molecular Biosciences, University of Texas at Austin, 2500 Speedway, Austin, TX 78712, USA. [2]Winship Cancer Center and Department of Biochemistry, Emory University School of Medicine, Wayne O Rollins Research Center, 1510 Clifton Rd NE, Atlanta, GA 30322, USA.
*Corresponding author. Email: blerta.xhemalce@emory.edu
†These authors contributed equally to this work.

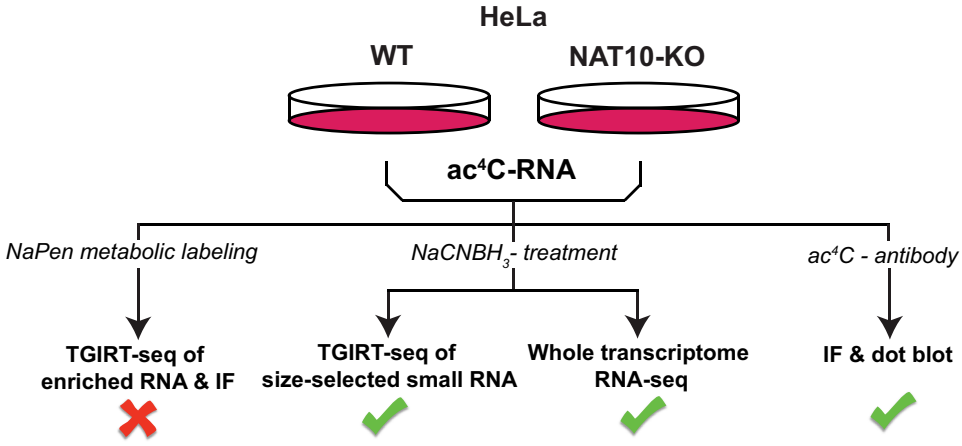

Fig. 1. Schematics representing our strategies to uncover the identity and function of diverse ac^{4}C RNAs in human cells.

RESULTS

Exploring ac^{4}C sites by NaCNBH$_3$-seq

To explore ac^{4}C at single-base resolution in RNA, we applied the sodium cyanoborohydride ($NaCNBH_3$) chemistry developed by the Meier group (*16*). Under acidic pH, $NaCNBH_3$ can reduce the C5═C6 bond of ac^{4}C, resulting in C → T misincorporation during RT-PCR (reverse transcription polymerase chain reaction).

We treated total RNA extracted from HeLa WT and NAT10-KO cells with $NaCNBH_3$ and coupled it to two different methods of next-generation sequencing: (i) thermostable group II intron reverse transcriptase sequencing (TGIRT-seq), which is optimal for sequencing tRNA (*22*), and (ii) whole-transcriptome RNA sequencing (RNA-seq; with post-treatment rRNA depletion), which is optimal for sequencing longer RNAs and mRNAs, regardless of their polyadenylation status (Fig. 2A). Upon $NaCNBH_3$ treatment using our optimized conditions (see Materials and Methods), the RNA was not majorly degraded and the ac^{4}C1842 site on helix 45 of the 18*S* rRNA was very efficiently reduced, leading to near-complete C → T conversion (Fig. 2, B and C).

To explore the ac^{4}C sites in tRNA, we treated total RNA extracted from HeLa WT and NAT10-KO cells with mock or $NaCNBH_3$ and then migrated them on a denaturing urea polyacrylamide gel to size-select full-length tRNAs (Fig. 2C). Subsequently, we used TGIRT-seq to sequence these gel-extracted tRNAs (*23*). The TGIRT can go past many post-transcriptional RNA modifications often giving rise to modification-specific mismatch signatures during sequencing, which has been particularly useful for tRNA sequencing (*22*). To unbiasedly identify all mismatches dependent on $NaCNBH_3$, NAT10-KO and their interaction, we computed Bhattacharya coefficients for all pairs (see Materials and Methods and fig. S1). We were able to identify ac^{4}C at position 12 in CCG (middle C is ac^{4}C) motif in tRNA^{Ser} and tRNA^{Leu}, as well as two sites present (C1337 and C1842) in 18*S* rRNA (table S1 and fig. S2). These findings were consistent with the previous reports (*1*, *3*, *4*, *7*). However, we did not observe any significant C → T conversion in any other CCG motif in the entire tRNA pool (table S1).

We also noted several other base modifications that were affected by either knockdown of NAT10 or $NaCNBH_3$ treatment in our TGIRT-seq data. The mitochondrial tRNA^{Met} (mt-Met-CAT) incorporates f^{5}C at C34 position, which was reduced by $NaCNBH_3$, giving rise to a strong C → T misincorporation signal in WT and NAT10-KO but was absent in all mock-treated samples (fig. S3) (*24*). In mitochondrial tRNA^{Glu} (mt-Glu-TTC), a strong T > A mismatch was observed at U33 in both treated and mock NAT10-KO samples but not in either mock or treated WT samples (fig. S4). Of note, mitochondrial tRNA^{Glu} (mt-Glu-TTC) contains a τm^{5}s^{2}U modification at the 34 wobble position, but no modification has been reported so far at U33. It is possible that another modifying enzyme is activated upon loss of NAT10 function and modifies the mitochondrial tRNA^{Glu} (mt-Glu-TTC) at its U33 position. It is also possible that this modification is an indirect consequence of a potential mitochondrial defect in NAT10-KO cells.

In parallel with TGIRT-seq of size-fractionated tRNAs (Fig. 2, C and D), we aimed to detect ac^{4}C in mRNAs, regardless of their poly(A) status, and in noncoding RNAs. To this end, after mock and $NaCNBH_3$ treatment as above (Fig. 2, A and B), we depleted ribosomal RNAs and performed whole-transcriptome RNA-seq libraries (see Materials and Methods). As shown in the volcano plots in Fig. 2E, the $NaCNBH_3$ treatment did not significantly affect the levels of detected RNAs in either WT or NAT10-KO samples (tables S2 and S3). As above, instead of assessing only C → T conversions, we analyzed all the possible mismatches induced by $NaCNBH_3$ treatment, NAT10-KO, or both. We failed to detect any statistically significant ac^{4}C sites on mRNAs, i.e., C → T mismatches consistently induced by $NaCNBH_3$ treatment in WT but not NAT10-KO cells (note that manual curation showed that the sole significant C → T site mapped to a unique rRNA site; Fig. 2F). NAT10-KO cells showed an elevated level of C → T mismatches compared to WT cells (Fig. 2G), regardless of whether the RNA was treated with $NaCNBH_3$ or not (Fig. 2H). Furthermore, most of these C → T mismatches in NAT10-KO cells were at ~50% of HeLa WT cells, suggesting that they may arise at the DNA level, on one of the two alleles (Fig. 2I).

NAT10 regulates R-loop levels

Our further analysis of the above data showed that the observed C → T predominantly occur within CCG triplets [CCG in the (+) strand of the reference genome and GGC in the (−) strand of the reference genome; Fig. 3A]. CCG corresponds to the NAT10-modifying sequence in rRNA and tRNA. Moreover, our RNA-seq showed that APOBEC3B and APOBEC3C, which deaminate C to U (resulting in C → T mismatches in DNA), are significantly increased at the transcriptional level. The fact that the

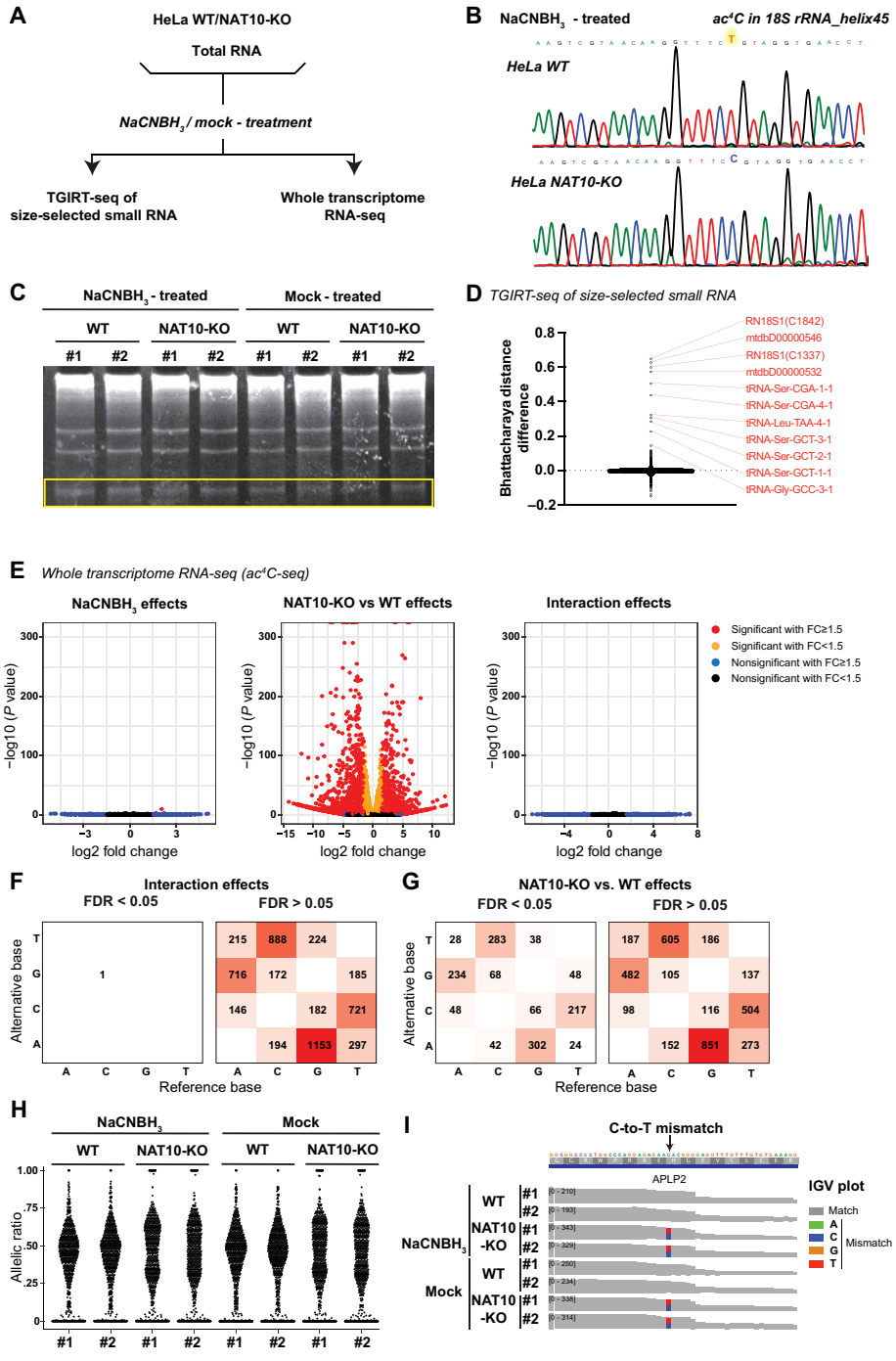

Fig. 2. NGS results from total and transfer RNAs, after treatment with mock or sodium cyanoborohydride ($NaCNBH_3$) in human cells. (**A**) Schematic representing our NGS strategy. (**B**) Validation of $NaCNBH_3$ treatment efficiency and specificity by Sanger sequencing of RT-PCR product, showing ~100% C → T conversion after $NaCNBH_3$ treatment in WT but not NAT10-KO cells at the known ac^4C site at the 18*S* rRNA helix 45 region. (**C**) Image of the urea polyacrylamide gel stained with SYBR used to size-select tRNAs (yellow box). (**D**) Analysis of mismatches induced by $NaCNBH_3$ treatment in the TGIRT-seq data of size selected tRNAs using Battacharya coefficient distances (see also fig. S1 for explanation of analysis and figs. S2 to S4 for IGV plots of select sites). (**E**) Volcano plots showing $NaCNBH_3$ versus mock treatment, NAT10-KO versus WT, or interaction effects on differential gene expression in the ac^4C-seq data of total RNA. *x* axis shows the fold change (FC) of normalized RNA read counts and the *y* axis shows the false discovery rate (FDR) ($\log_{10}$). Red dots indicate significantly down- or up-regulated genes with a fold change FC ≥ 1.5 and FDR ≤ 0.05, while the other colors indicate the other combinations of FC and significance, as indicated in the legend in top right of the volcano plot. (**F**) Heatmaps showing interaction effects on all possible mismatches observed in the ac^4C-seq data of total RNA. (**G**) Heatmaps showing NAT10-KO versus WT effects on all possible mismatches observed in the ac^4C-seq data of total RNA. (**H**) Graph representing raw allelic ratios (alternative count/total count) of all C → T and G → A variants by sample, showing highly different distributions between WT and NAT10-KO samples, but no difference based on $NaCNBH_3$ treatment. (**I**) IGV plot showing a typical $NaCNBH_3$ treatment-independent C → T mismatch observed in NAT10-KO cells.

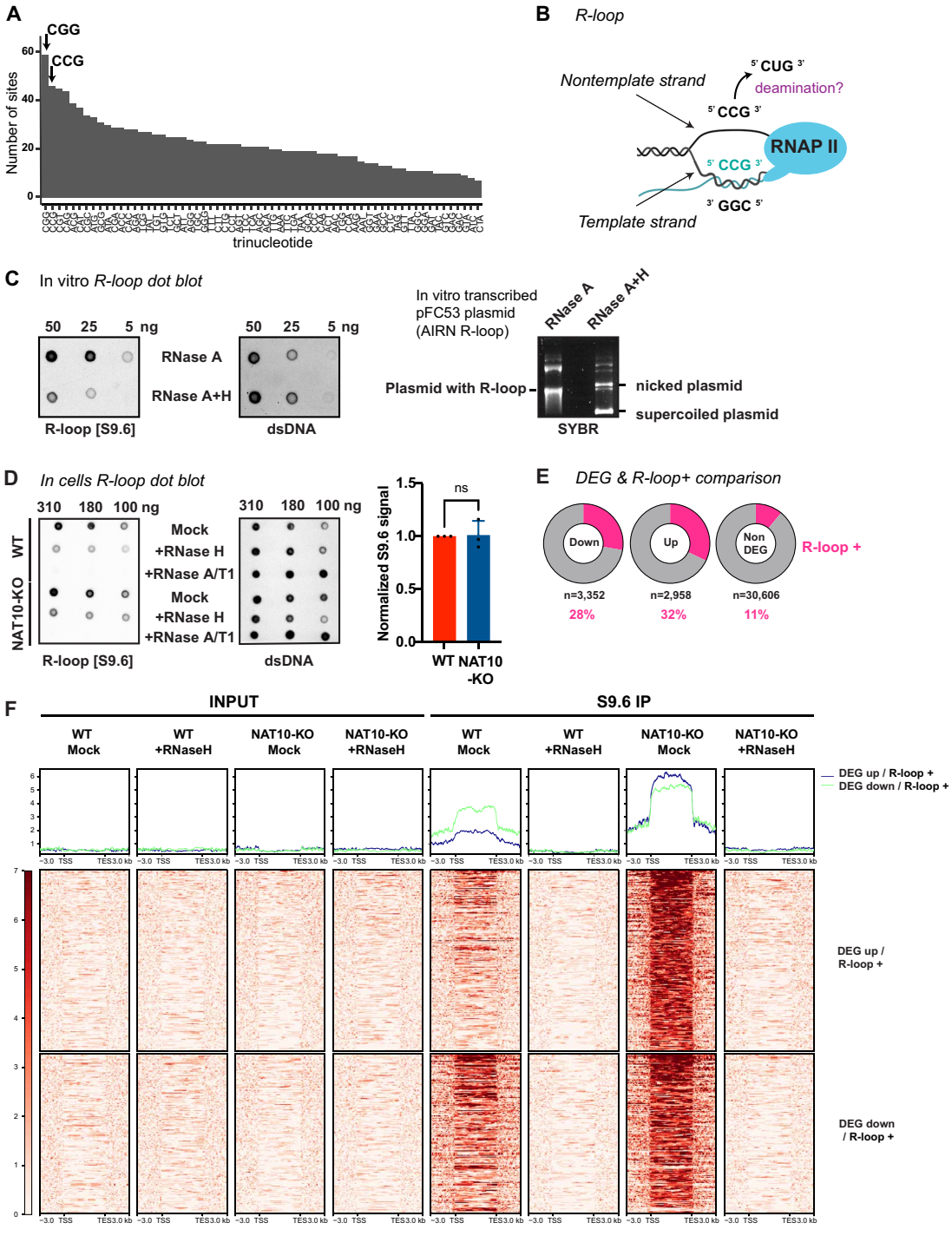

Fig. 3. NAT10 regulates R-loop levels. (**A**) Triplet sequence analysis showing C ⟶ T mismatches in NAT10-KO cells. (**B**) Schematic of a cotranscriptional R-loop, showing the single-stranded DNA strand that is also the coding strand and the DNA/RNA hybrid, where the RNA is base-paired with the template strand. An example of a CCG triplet getting deaminated is also shown for clarity. (**C**) Verification of the specificity of the S9.6 antibody with a dot blot using an in vitro produced R-loop. Left: Dot blot with the S9.6 and anti-dsDNA antibodies with in vitro produced R-loop, treated with only RNase A that cleaves single-stranded RNA, or with RNase A and RNase H that cleaves the RNA strand within an RNA/DNA hybrid. Right: Validation of in vitro produced R-loop used in the dot blot by agarose gel electrophoresis. (**D**) Dot blot with S9.6 and anti-dsDNA antibodies with the indicated amount of DNA from HeLa WT and NAT10-KO cells, treated with mock, RNase H, or excess of RNase A and RNase T1. The graph shows the quantification of the S9.6 signal in HeLa WT and NAT10-KO cells (mock sample) over the dsDNA signal (RNase A + T1 sample) (mean ± SD, $n = 3$ biological replicates, paired t test). (**E**) Donut chart showing the portion of R-loop–forming genes among DEG, either down- or up-regulated in NAT10-KO versus WT HeLa cells, versus non-DEG. The R-loop–forming genes list was retrieved from RR-ChIP-seq (*40*). Note that genes with dRNase H1 peaks only in their introns were excluded from the list of R-loop + genes. n = total number of genes in each category. (**F**) Heatmaps showing the enrichment of R-loops in DEG, from TSS to TES, ±3 kb on each side. The genes are separated in DEGs that either up-regulated (up) or down-regulated (down) in NAT10-KO versus WT cells.

C → T mutations are mostly within NAT10 consensus sequence, instead of a random sequence, is intriguing. APOBEC3B has recently been implicated in resolving R-loops formed in a cotranscriptional manner (*25*). As mentioned in the introduction, R-loops are formed by an RNA/DNA hybrid and the displaced ssDNA of the original DNA duplex (*19*) (Fig. 3B). APOBEC3B is thought to initiate R-loop resolution by deaminating the displaced ssDNA, which at transcribed regions corresponds to the nontemplate strand. The fact that we observe more C → T mismatches than G → A mismatches in mRNAs in NAT10-KO cells suggests that APOBEC3B may have acted on the nontemplate or coding strand preferentially. Consistent with this, 55% of the identified mismatches overlap with high-confidence R-loop regions (table S4). A quarter of these mismatches are found in clusters of two or more mismatches within the same R-loop region, which is indicative of enzymatic deamination activity (*25*). Together, these observations led us to hypothesize that NAT10 might be involved in R-loop resolution.

To functionally assess the effect of NAT10 depletion on R-loop formation, we first performed dot blots with the RNA/DNA hybrid S9.6 mouse monoclonal antibody. Our antibody was validated with an in vitro R-loop system consisting of the pFC53 plasmid (*26*) in vitro transcribed and treated with ribonuclease (RNase) A to remove free RNA or RNase A + H to additionally degrade the R-loop (Fig. 3C). Genomic DNA was purified from HeLa WT and NAT10-KO cells using a gentle DNA extraction method that preserves R-loops (*27*) and was analyzed with dot blot with the S9.6 antibody using both a RNase H treatment control and a cocktail of excess RNase A + T1 control (in low-salt concentration) to remove any residual RNA (Fig. 3D). This analysis showed that NAT10-KO does not significantly affect global R-loop levels, at least not at the resolution provided by dot blot analysis (Fig. 3D). However, given that differentially expressed genes (DEGs) in NAT10-KO versus WT are enriched in R-loop–positive genes (Fig. 3E), we analyzed the effect of NAT10-KO on R-loop levels at specific genes with DNA-RNA hybrid immunoprecipitation coupled to next-generation sequencing or DNA-RNA immunoprecipitation sequencing (DRIP-seq) (*27*). As shown on Fig. 3F, NAT10-KO substantially increased the levels of R-loops at R-loop–positive genes that are differentially expressed (DEGs). The simplest explanation for the increase in R-loop levels in up-regulated genes is that their transcription is increased upon NAT10-KO. However, the increased R-loop levels in down-regulated genes are more intriguing and indicate that NAT10 may help resolve R-loops to promote productive transcription. Thus, it is likely that at least a part of the effect of NAT10-KO on the transcriptome may be related to a defect in R-loop resolution. In addition to the effects on gene expression, and consistent with an increase in R-loop levels, NAT10-KO cells display increased levels of endogenous DNA damage, as marked by γ-H2AX (fig. S5).

NAT10 acetylates R-loops and promotes their resolution in a catalytic activity–dependent manner

Our results showing that NAT10-KO increases R-loop levels at specific genes suggest that NAT10 could directly affect R-loop resolution by acetylating the RNA strand. To test this hypothesis, we first performed dot blots on gently extracted DNA from HeLa WT and NAT10-KO cells with a specific ac^4C antibody (*6*) as described above (Fig. 3D). As shown on Fig. 4A, HeLa WT cells display an ac^4C antibody signal that is significantly reduced by treatment with RNase H, down to the level observed in NAT10-KO cells, suggesting that NAT10 may acetylate the RNA portion of R-loops. To determine whether the acetyltransferase activity of NAT10 is functionally important for R-loop resolution, we constructed HeLa Flp-In T-REx cells with an inducible NAT10-FLAG (NAT10f) insert, either WT or an R628A catalytic mutant (Fig. 4B, top), and induced NAT10 expression for 24 hours (Fig. 4B, bottom) before performing DRIP-seq with the S9.6 antibody. Impressively, inducing NAT10f-WT overexpression for 24 hours resulted in considerably lower levels of R-loops at the R-loop positive DEGs (Fig. 4C). These data are consistent with the results in NAT10-KO cells; while NAT10-KO increases R-loop levels, NAT10 overexpression reduces them at NAT10-dependent DEGs. However, this was not observed with the NAT10f-R628A catalytic mutant, despite this mutant being expressed at the same level as NAT10f-WT (Fig. 4B). This is a key result that strongly suggests that NAT10's catalytic activity is crucial for its role in R-loop resolution.

ac^4C antibody reveals cellular RNA targets localized at endosomal structures

In parallel to the above studies, we also used the ac^4C antibody to detect ac^4C RNA targets by IF (see Materials and Methods). Notably, HeLa WT cells displayed bright fluorescent foci (30 to 300 dots per cell) that were highly reduced in number and intensity in NAT10-KO cells (0 to 60 dots per cell) (Fig. 5A and fig. S6). Moreover, these foci were also significantly reduced by RNase A treatment (Fig. 5A), confirming their RNA origin. Under our IF conditions, the ac^4C foci were mostly cytoplasmic, but they were also visible in the nucleus and nucleolus (Fig. 5A). To investigate the nature of the cytoplasmic ac^4C foci, we performed a small screen for colocalization of these foci with established markers of cytoplasmic bodies and organelles.

First, we performed colocalization studies between ac^4C and the small ribosomal subunit RPS6. Although RPS6 exhibited the expected cytoplasmic staining, it did not colocalize with the ac^4C foci (fig. S7). This result suggested that the ac^4C modification on the intact small subunit of the ribosome (within the 18*S* rRNA) is likely inaccessible under our IF conditions. This also suggests that the IF with the ac^4C antibody is likely detecting other types of RNAs.

To check colocalization of ac^4C foci with P-bodies, lysosomes, endoplasmic reticulum (ER), mitochondria, and peroxisomes, we used transfection of plasmids expressing fluorescent marker proteins or endogenous antibodies: Decapping MRNA 1A–green fluorescent protein (DCP1A-GFP) (P-body marker, fig. S8), Lysosomal Associated Membrane Protein 1 (LAMP1; lysosomal marker, fig. S9), CALNEXIN (ER marker, fig. S10), cpVenus-GFP (mitochondrial reporter, fig. S11), and Peroxisomal Biogenesis Factor 14 (PEX14-GFP; peroxisome marker, fig. S12). We observed no substantial colocalization of ac^4C signals with DCP1A-GFP, CALNEXIN, cpVenus-GFP, and negligible colocalization of ac^4C signals with LAMP1 and PEX14-GFP (figs. S7 to S12). While the majority of PEX14 foci were on close proximity to ac^4C foci, they were not colocalized. In contrast to these cellular bodies and organelles, we observed notable overlap of the ac^4C foci with multiple endosomal markers, such as Ras-related protein RAB5 (early endosomes), RAB7, and RAB11 (mature endosomes) (Fig. 5B and fig. S12). Moreover, in the RAB5-GFP/ac^4C images, we clearly observed distinct stages of ac^4C foci, residing on or inside GFP-positive circular organelles (Fig. 5B).

Given our results that NAT10-KO increases the levels of R-loops, we wondered whether certain R-loop processing products were transported out of the nucleus and packaged into endosomal structures. We observed that overexpression of RNase H1 significantly increased the number of ac^4C foci, suggesting that at least a portion

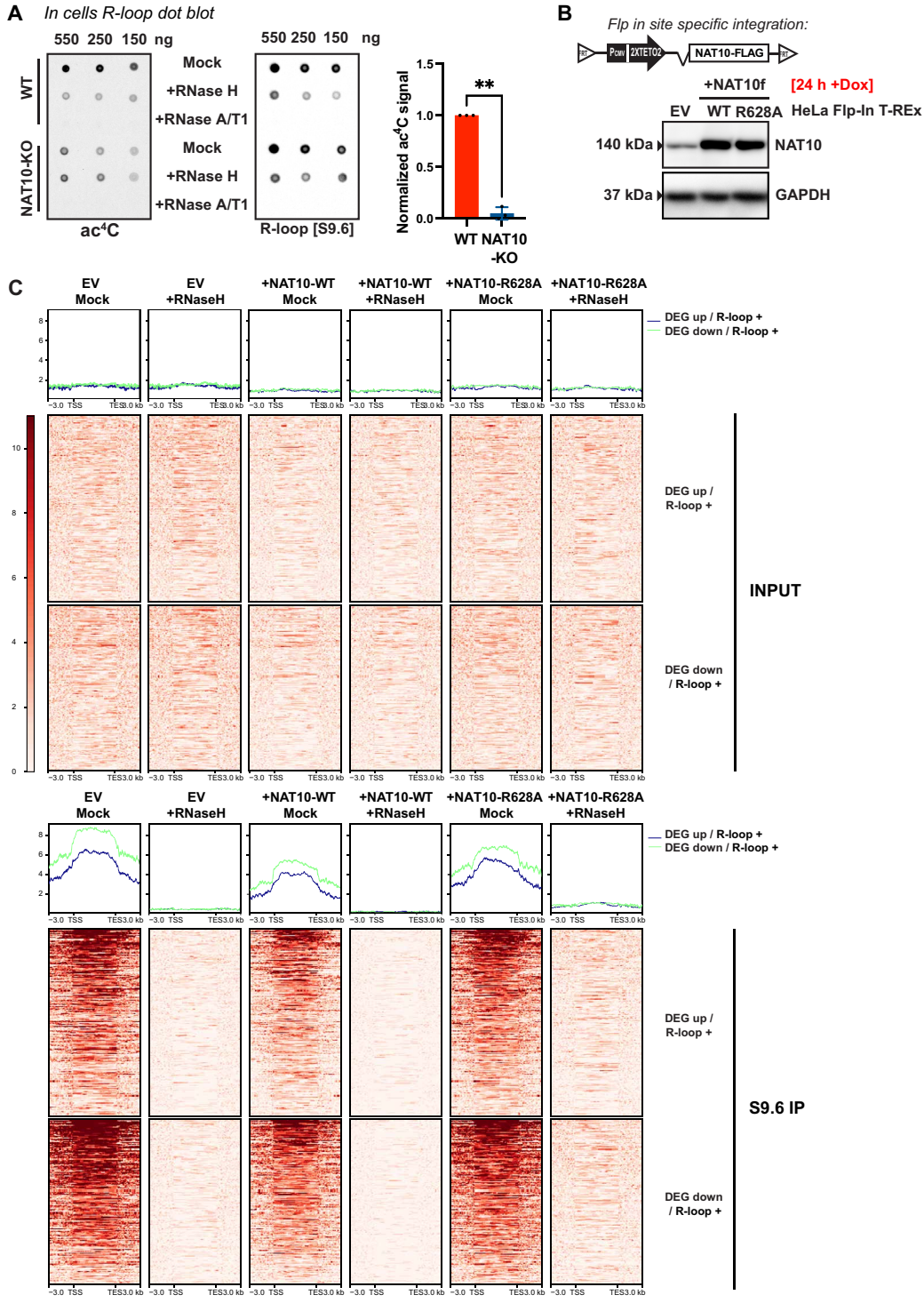

Fig. 4. NAT10's catalytic activity is required to regulate R-loop levels. (**A**) Dot blot with the ac^{4}C and S9.6 antibodies with the indicated amount of gently extracted DNA from HeLa WT and NAT10-KO cells, treated with mock, RNase H, or excess of RNase A and RNase T1. The graph shows the quantification of the ac^{4}C signal in HeLa WT and NAT10-KO cells mock sample over the +RNase H sample (mean ± SD, $n = 3$ biological replicates, paired t test). (**B**) Top: Schematic of the FRT locus in HeLa Flp-In T-REx cells with an inducible NAT10-FLAG (NAT10f) insert. Bottom: Western blots of whole-cell extracts (WCE) of HeLa Flp-In T-REx empty vector (EV), and FLAG-tagged NAT10f-WT or R628A catalytic mutant after induction for 24 hours with doxycycline (2 μg/ml) with antibodies detecting NAT10 and glyceraldehyde-3-phosphate dehydrogenase (GAPDH; loading control). The size of the protein in kDa is also shown on the left of each Western blot. (**C**) Heatmaps showing the enrichment of R-loops in HeLa Flp-In T-REx EV, and FLAG-tagged NAT10f-WT or R628A catalytic mutant after induction for 24 hours with doxycycline [2 μg/ml; see (B)], analyzed as in Fig. 3F, using the set of DEGs that either up-regulated (up) or down-regulated (down) in NAT10-KO versus WT cells.

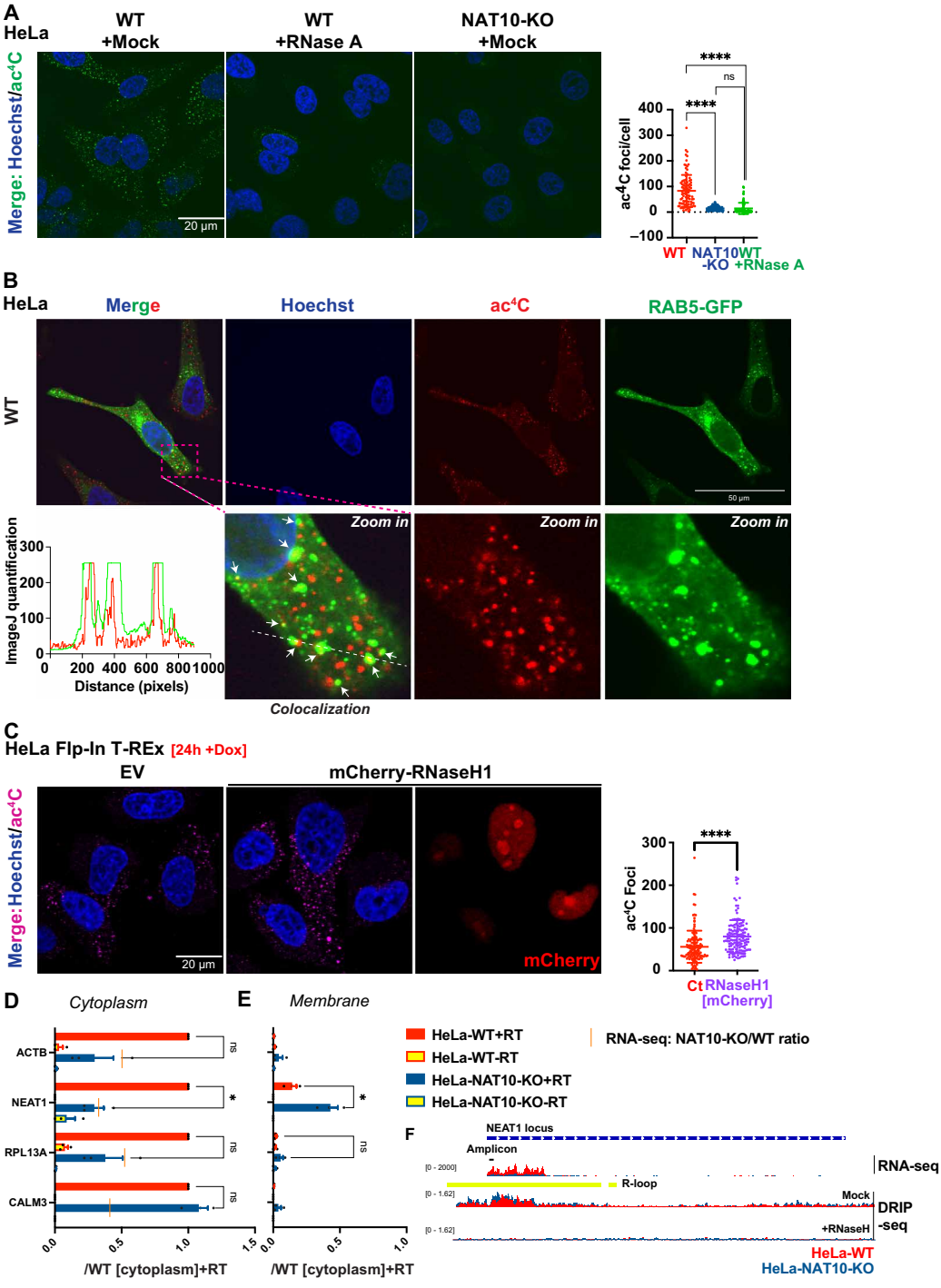

Fig. 5. NAT10 acetylates single stranded RNAs transported into endosomal structures. (**A**) NAT10 acetylates RNAs sensitive to degradation by RNase A. Left: Representative image of IF with the ac^{4}C antibody in HeLa WT and NAT10-KO cells treated with mock, and in HeLa WT treated with RNase A. Right: Quantification of the number of ac^{4}C foci per cell [$n = 100$ cells, ordinary one-way analysis of variance (ANOVA) test]. (**B**) Representative image of IF with the ac^{4}C antibody in HeLa WT cells transfected with RAB5-GFP. The white arrows on the zoomed-in image show colocalization between ac^{4}C and RAB5 (in yellow). The graph shows the raw Image J profile analysis of the line shown on the zoom-in merged image for each channel. (**C**) Nuclear RNase H1 overexpression increases the number of ac^{4}C foci. Left: Representative image of IF with the ac^{4}C antibody in HeLa Flp-In T-REx WT empty vector (EV) or human nuclear mCherry-RNase H1 induced for 24 hours with doxycycline (2 µg/ml). Right: Quantification of the number of ac^{4}C foci per cell ($n = 138$ cells, Welch's t test). Only red RNase H1–positive cells were used for ac^{4}C foci quantification. (**D** and **E**) Quantitative PCR analysis after a reverse transcription reaction with mock (−RT) or with reverse transcriptase (+RT) from cytoplasmic- or membrane-bound nucleic acids at the indicated R-loop amplicons. The results are normalized to the WT[Cytoplasm] + RT samples for each amplicon (mean ± SD, $n = 4$ biological replicates, multiple paired t test). The yellow line shows the normalized level of expression of the corresponding gene in the RNA-seq analysis. (**F**) UCSC browser image of the RNA-seq and DRIP-seq results at the *NEAT1* nuclear long noncoding RNA locus. The position of the NEAT amplicon is also shown.

of these ac^{4}C foci are R-loop–derived RNAs (Fig. 5C). In addition, we performed cellular fractionation of HeLa WT and NAT10-KO cells (fig. S5C), purified nucleic acids from the cytoplasmic and membrane fractions, and checked several R-loop amplicons from the DEG down category with and without including a reverse transcription step. We found that the measured levels of the amplicon in the cytoplasm were similar to the levels of the mature RNA (Fig. 5D) for *ACTB*, *NEAT1* and *RPL13A*, but not *CALM3*. For the *NEAT1* amplicon, we found that its levels were highly up-regulated in the membrane fraction in the NAT10-KO cells (Fig. 5E). This may also be the case for *ACTB* and *RPL13A*, but in those cases, the signal coming from the mature mRNA may dominate over the signal coming from the R-loop. No or a negligible signal was obtained for all amplicons in the samples without a reverse transcription step, showing that only the RNA portion of the R-loop amplicon was present in the cytoplasm/membrane fractions (Fig. 5, D and E).

NAT10-KO induces inflammatory TLR7 signaling

Analysis of RNA-seq in HeLa WT and NAT10-KO using ingenuity pathway analysis (IPA) revealed that NAT10-KO induces gene expression patterns triggered in inflammation, viral response, cardiac hypertrophy, hepatic fibrosis, and osteoarthritis. Recent research from the Meier group reported that substitution of cytidine with ac^{4}C in transfected RNA can decrease inflammatory gene activation (*28*). Therefore, it is possible that the loss of acetylation in cellular RNA triggers the inflammatory gene expression in NAT10-KO cells. IPA predicts that TLR7-related pathway is induced upon NAT10 knockout (Fig. 6A). TLR7 is an endosomal innate immune sensor capable of detecting single-stranded RNAs (*20*, *21*, *29*), including of small size (*29*–*31*), such as those produced by RNase H1 (*32*). As shown in Fig. 6B, TLR7 levels are highly increased in HeLa NAT10-KO cells. Moreover, TANK-binding kinase 1 (TBK1), a major downstream marker of inflammatory pathways, is hyperphosphorylated, especially in the nuclear fraction (Fig. 6C). Moreover, consistent with the lack of detected R-loop DNA in the cytoplasm (Fig. 5, D and E), P-STING (stimulator of interferon genes phosphorylated at Ser^{366}) was not activated (Fig. 6B). TLR7 protein levels are also induced by siNAT10 (Fig. 6D), showing that this effect can be induced rapidly and is not a consequence of NAT10-KO cell line generation. When we used a TLR7 antibody for IF imaging, we observed that many TLR7 cytoplasmic foci colocalized with the ac^{4}C antibody foci (Fig. 6E). This result suggested that TLR7 may interact with single-stranded RNA that is acetylated by NAT10 to prevent its activation. In NAT10-KO cells, these single-stranded RNAs levels may be increased and/or hypo-acetylated, triggering TLR7 activation and inflammatory responses. Given that our analysis shows that the RNA counterpart of R-loops is acetylated in a NAT10-dependent manner (Fig. 4A), we decided to test whether these single-stranded RNAs are derived from R-loops in the cytoplasm. To this end, we overexpressed RNase H1 in WT cells and observed its effect on the formation of TLR7 foci. This experiment showed that TLR7 foci increased in response to RNase H1 overexpression and colocalized with ac^{4}C foci (Fig. 6F and fig. S14). To functionally test whether R-loop–derived RNAs are responsible for TLR7 activation, we generated HeLa Flp-In T-REx NAT10-KO cells with inducible expression of either human nuclear-only RNase H1 WT or a catalytic mutant (dRNase H1) from a single-site insertion at the FRT locus. As shown in Fig. 6G, overexpression of nuclear RNase H1 WT did not reduce TLR7 activation in NAT10-KO cells; however, expression of nuclear dRNase H1 drastically decreased the levels of TLR7 (Fig. 6G). This result is consistent with nuclear RNase H1 being responsible for producing the nuclear R-loop–derived RNAs in the cytoplasm, while nuclear dRNase H1 traps them in the nucleus, leading to a decrease of R-loop-derived RNAs in the cytoplasm and a reduction in TLR7 signaling. Thus, overall, our results suggest that NAT10 may acetylate the RNA strand within R-loops, rendering R-loop–derived RNAs less inflammatory.

DISCUSSION

In this paper, we presented our findings from several biochemical approaches aiming to comprehensively decipher NAT10-dependent RNA acetylation in human cells (Fig. 1). We initially sought to use metabolic labeling with sodium 4-pentynoate (NaPen). While this approach labeled proteins as expected (*33*), it did not label known ac^{4}C sites. This may be due to NAT10's inability to transfer longer or bulkier acyl chains, such as those derived from NaPen, to target RNA acetylation sites.

Using $NaCNBH_3$ chemistry coupled with two RNA sequencing methods (Fig. 2A), we reproduced all known sites in tRNAs and rRNAs (Fig. 2, B to D) but observed no confident acetylated RNA in mRNA (Fig. 2, E and F). This is in agreement with the Sas-Chen *et al.* (*3*) paper but also with the data from Arango *et al.* (*7*) paper. Arango *et al.* reported ac^{4}C sites after pooling biological replicates, whereas we only reported sites that were reproduced at similar rates in both biological replicates separately. This lack of reproducibility between biological replicates may be due to technical issues connected to the labile nature of ac^{4}C or possibly to the spurious or reversible nature of mRNA acetylation sites.

Our unbiased analysis of all mismatches to the reference genome in both mock- and $NaCNBH_3$-treated cells revealed an increased rate of C → T conversions in NAT10-KO cells compared to WT cells (Fig. 2, G and H), many of which behaved like mutations occurring at the level of DNA (Fig. 2I). This, together with the fact that APOBEC3B activity, which produces C → T mutations through cytosine deamination, has recently been shown to act on the single-stranded DNA part of cotranscriptional R-loops (*25*), led us to explore a potential connection between R-loops and NAT10. We found that NAT10 knockout increased R-loops levels at many genes, including at genes with reduced expression levels in NAT10-KO compared to WT (Fig. 3F). Thus, it is likely that a portion of the NAT10's effects on the transcriptome may be due to defects in R-loop processing. We also found that the RNA counterpart of R-loops is acetylated by NAT10 (Fig. 4A). This result suggests that NAT10 may directly interact with R-loops, which is consistent with the analysis of R-loop interactome from the Gromak lab (*34*). Moreover, when overexpressed for a short period of time (24 hours), NAT10-WT, but not NAT10-R628A catalytic mutant, highly reduced R-loop levels at R-loop–positive DEGs. This demonstrates that NAT10's catalytic activity is involved in R-loop resolution. Given that NAT10 acetylates the RNA portion of R-loops in cells, it is likely that its RNA acetylation activity is required for NAT10's effect on R-loop resolution, though we do not exclude the possibility that NAT10's protein acetylation activity may also be involved. This activity may be required for transcriptional regulation, as well as during responses to DNA damage, as recently observed for other RNA modifications within R-loops (*35*). Thus, given that NAT10 is relocalized from the nucleolus to the nucleoplasm upon DNA damage (*15*), it

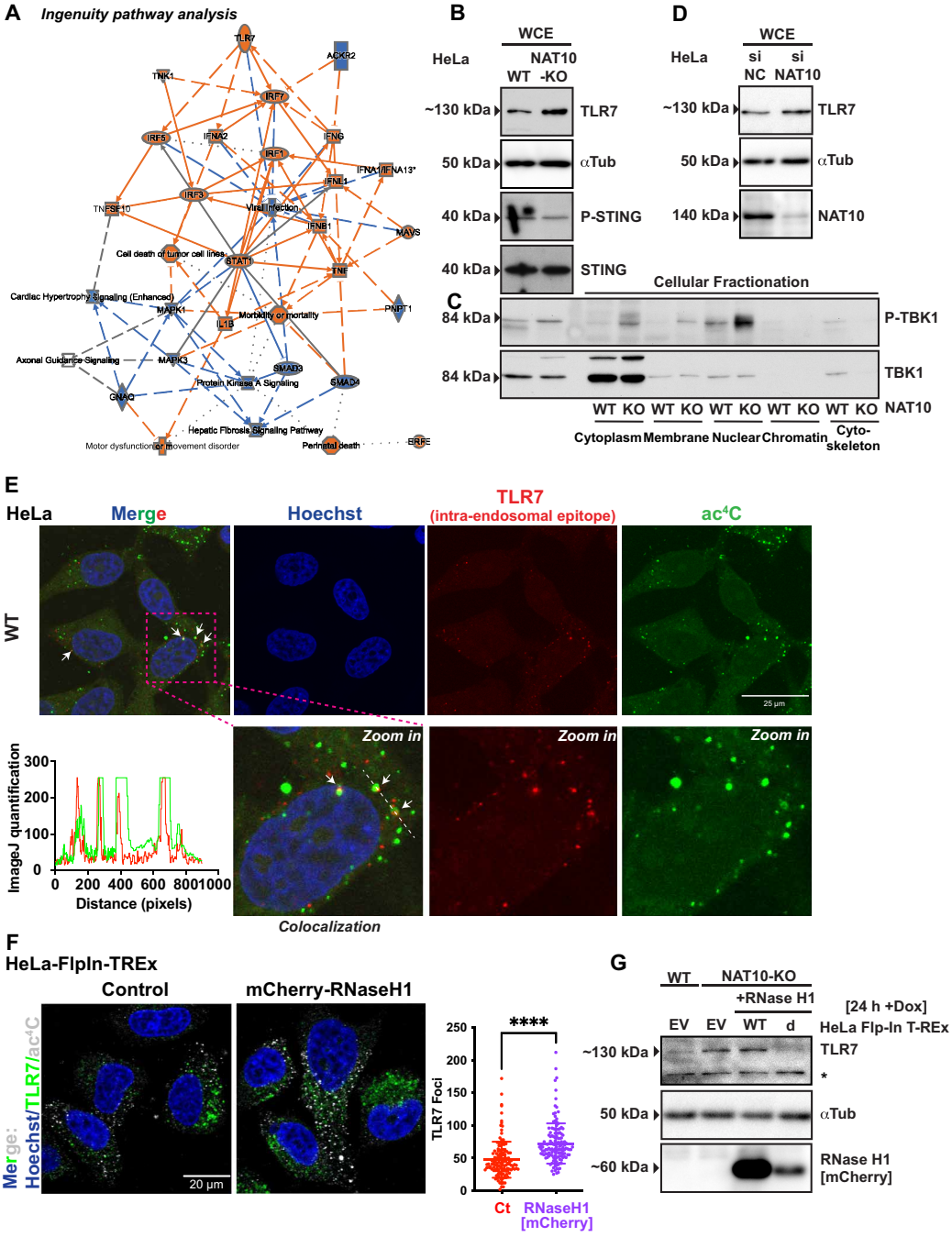

Fig. 6. NAT10 knockout induces the TLR7 ssRNA innate immune sensor in an R-loop–sensitive manner. (**A**) Graphical summary of pathways predicted to be differentially regulated in NAT10-KO versus WT cells by IPA. (**B**) Western blots of whole-cell extracts (WCE) of HeLa WT and NAT10-KO cells with antibodies detecting TLR7, α-tubulin (loading control), P-STING (STING phosphorylated at Ser^{366}) and unmodified STING. The size of the protein in kDa is also shown on the left of each blot. (**C**) Western blot of WCE and cellular fractions of HeLa WT and NAT10-KO cells with antibodies detecting P-TBK1 (TBK1 phosphorylated at Ser^{172}) and unmodified TBK1. See also fig. S5 for full analysis of cellular fractionation. (**D**) Western blots of WCE of HeLa siNC and siNAT10-KO cells with antibodies detecting TLR7, α-tubulin (loading control), and NAT10. (**E**) Representative image of IF with the ac^{4}C antibody and TLR7 in HeLa WT cells. The white arrows on the zoomed-in image show colocalization between ac^{4}C and TLR7 (in yellow). The graph shows the raw ImageJ profile analysis of the line shown on the zoom-in merged image for each channel. (**F**) RNase H1 overexpression increases the number of TLR7 foci. Left: Representative image of IF with the ac^{4}C and TLR7 antibodies in HeLa Flp-In T-REx WT empty vector (EV) or mCherry-RNase H1 induced for 24 hours with doxycycline (2 μg/ml; same as in Fig. 5C). Right: Quantification of the number of TLR7 foci per cell (n = 138 cells, Welch's t test). Only red RNase H1–positive cells were used TLR7 foci counting. (**G**) Western blots with antibodies detecting TLR7, α-tubulin (loading control), and RNase H1 of WCE of HeLa-FlpIn-Trex WT and NAT10-KO cells containing a single-copy insertion of empty vector, human nuclear RNase H1-mCherry WT, or human nuclear dRNase H1-mCherry (d for deactivated, D210N), after induction for 24 hours with doxycycline (2 μg/ml). RNase H1 and dRNase H1 mRNAs are induced at the same level, but the dRNase H1 protein is less stable, explaining the lower levels of induction compared to WT RNase H1. Asterisk indicates a nonspecific band.

will be of importance to determine whether NAT10-mediated R-loop regulation also plays a role during DNA damage.

By using an ac^{4}C-specific antibody (*6*), we revealed a substantial number of NAT10-dependent acetylated RNA in the cytoplasm (Fig. 5A). Given that NAT10 is mostly nuclear, it is likely that NAT10 acetylates these RNAs in the nucleus, before their trafficking to the cytoplasm. These ac^{4}C RNAs did not colocalize with the small ribosomal subunit (fig. S7), but colocalized with endosomal structures (Fig. 5B and fig. S13). These ac^{4}C RNAs may be mostly single stranded, as they are efficiently degraded by RNase A (Fig. 5A and fig. S6). Given that NAT10 acetylates rRNA and tRNA within double-stranded RNA regions, these single-stranded ac^{4}C RNAs may derive from misfolded or fragmented rRNA and tRNA. In addition, they may also derive from mRNA (*6*, *7*) and from the RNA counterpart of R-loops. In support of this latter RNA category, RNase H1 overexpression resulted in a significant increase of ac^{4}C foci (Fig. 5C). The ac^{4}C RNAs detected by IF also colocalized with TLR7, an RNA immune sensor specific for single-stranded RNA (Fig. 6E), and importantly, TLR7 levels also increased upon RNase H1 overexpression and colocalized with ac^{4}C foci (Fig. 6F). As predicted by IPA (Fig. 6A), TLR7 protein levels are increased upon NAT10 knockout, along with increased levels of TBK1 phosphorylation, a major downstream marker of inflammation (Fig. 6B). In addition, TLR7 levels decreased in NAT10-KO cells within 24 hours of inducing a catalytic mutant of RNase H1 (dRNase H1) (Fig. 6G), which stabilizes R-loops in the nucleus. Overall, we speculate that in WT cells, RNase H1 fragments from the RNA part of R-loops, are exported to the cytoplasm, where they are protected from triggering TLR7 through ac^{4}C modification. In NAT10-KO cells, R-loop levels are not only increased but also are no longer acetylated, leading to TLR7 activation (Fig. 7). As mentioned previously, this is in agreement with acetylation rendering transfected mRNAs less immunogenic (*28*). Future studies will be required to decipher the characteristics of these cytoplasmic acetylated RNAs and their roles in cancer and inflammatory human diseases. In addition, the visualization method of acetylated RNAs we introduced here may be used as a biomarker in patient cells and tissues.

MATERIALS AND METHODS

Cell lines

The HeLa NAT10-KO cell line was a gift from S. Oberdoerffer. The HeLa Flp-In T-REx cell line was a gift from M. Hentze and M. Price.

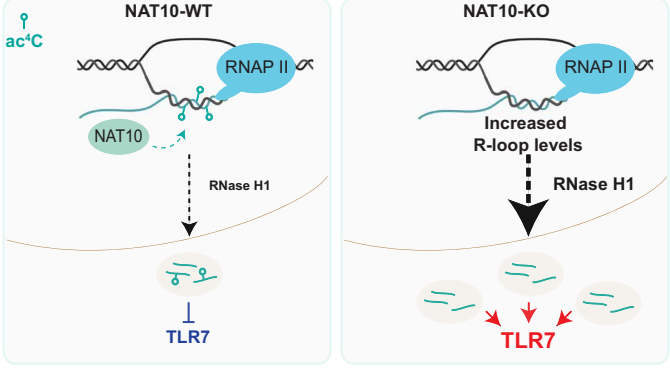

Fig. 7. Schematic of a simplified model of how NAT10 and ac^{4}C might suppress TLR7 activation.

HeLa-Flp-In T-REx NAT10-KO cells were constructed using the Santa Cruz Biotech sc-406713 and sc-406713-HDR vectors. HeLa Flp-In T-REx WT with an empty vector, NAT10-WT-FLAG, NAT10-R628A-FLAG; HeLa Flp-In T-REx WT, and NAT10-KO with an empty vector (EV), RNase H1, and dRNase H1 (D210N) were constructed according to the manufacturer's instructions (Thermo Fisher Scientific). The corresponding plasmids will be deposited on Addgene.

RNA/protein extraction

HeLa WT, NAT10-KO cells were grown in Dulbecco's minimum essential medium (DMEM) supplemented with 10% fetal bovine serum (FBS), penicillin (100 U/ml), streptomycin (100 μg/ml), and 2 mM L-glutamine (PSQ). HeLa WT (5×10^4 cells/ml) and HeLa NAT10-KO (7.5×10^4 cells/ml) cells were plated in 35-mm plates and cultured for 2 days. Since the growth of NAT10-KO is slower than WT, 1.5 times more NAT10-KO cells were seeded for all experiments, unless otherwise mentioned. Cells were washed with 37°C-heated phosphate-buffered saline (PBS) two times, lysed for 5 min in SKP buffer, and frozen at −80°C until RNA and protein purification using a Norgen RNA/Protein Plus kit (product no. 48200) was performed according to the manufacturer's instructions.

HeLa Flp-In T-REx WT empty vector cells were grown in DMEM + FBS + PSQ supplemented with blasticidin (10 μg/ml; Invivogen, ant-bl-1) and hygromycin (400 μg/ml; Invivogen, ant-hg-5). HeLa Flp-In T-REx NAT10-KO empty vector, RNase H1, and dRNase H1 cells were grown in DMEM + FBS + PSQ supplemented with blasticidin (10 μg/ml; Invivogen, ant-bl-1), hygromycin (400 μg/ml; Invivogen, ant-hg-5), and puromycin (1 μg/ml; GIBCO, Fisher, no. A11138-03). 1×10^5 cells/ml were seeded, and 24 hours postplating, RNase H1 expression was induced by treating the cells with doxycycline (2 μg/ml; Sigma-Aldrich, no. D5207-1G) for 24 hours. RNA/protein extraction procedure was same as mentioned above.

RT-PCR/qPCR analysis

One hundred to 500 ng of RNA (chemically treated or untreated) was reverse transcribed into cDNA with the SuperscriptIII First-Strand Synthesis System for RT-PCR kit (Invitrogen) using both oligo(dT)$_{20}$ and random hexamers. For the PCR, 1/20th of each reaction was used according to the GoTaq Flexi DNA Polymerase (no. M8295) protocol, loaded onto 1% agarose gel to confirm the PCR product and purified using QIAquick PCR Purification Kit (Qiagen). The purified PCR product was sent to GENEWIZ for Sanger sequencing. For quantitative PCR (qPCR), 1/20th of each reaction was used for real-time PCR with PowerUp SYBR Master Mix and specific primers on a StepOne Plus system.

NaCNBH$_3$-treatment coupled to next-generation sequencing

NaCNBH$_3$-treatment was performed according to the protocol described in Sas-Chen *et al.* (*3*). Three micrograms of total RNA from WT HeLa and NAT10-KO cells was used for NaCNBH$_3$ or mock treatment. On the same day as the treatments were performed, a fresh solution of 1 M sodium cyanoborohydride (Sigma-Aldrich, no. 42077) was prepared in ultrapure RNase-free water. In a 1.5-ml Eppendorf tube, 3 μg of total RNA in 25 μl of water was mixed with 10 μl 1 M HCl solution, followed by 55 μl of water, and 10 μl of the freshly prepared 1 M NaCNBH$_3$ solution (final concentration of 100 mM). For the mock treatment, the conditions were exactly the same, except that 10 μl of water was used instead of the NaCNBH$_3$ solution. Each tube was gently vortexed at a low speed for 30 s and incubated at 20°C

for 20 min. Immediately after the reaction, first 30 µl of 1 M tris-HCl (pH 8.0) and then 70 µl of water was added to each tube to scale up the volume to 200 µl. The reaction mixture was then cleaned up using RNeasy MinElute kit (Qiagen, no. 74204) with a modified protocol that allows recovery of RNAs of all sizes (*36*). The conversion efficiency was confirmed by Sanger sequencing (Fig. 2B), taking ac^4C1842 site (in helix 45 of 18*S* rRNA) as a positive control. One microgram of each treated/purified total RNA was directly used for ribosomal RNA depletion without fragmentation, followed by RNA-seq library preparation (TruSeq stranded total RNA library preparation kit from Illumina). Libraries were sequenced on a NextSeq 500 instrument [75 nucleotides (nt), paired-end reads] at the Next Generation Sequencing Facility at MD Anderson Science Park.

RNA gel electrophoresis for size-selected small TGIRT-library preparation

After mock and $NaCNBH_3$ treatment and clean up with the Qiagen RNeasy MinElute Cleanup Kit, 1 µg of RNA was mixed with an equal volume of Gel Loading Buffer II (AM8547), denatured for 15 min at 70°C, and placed on ice. The RNA was migrated on 15% denaturing 8 M urea polyacrylamide gel electrophoresis gel in a Criterion midi vertical electrophoresis system in 1× tris-borate EDTA (TBE). The gel was stained for 5 min with 50 ml of 1× TBE containing 5 µl of 10,000X SYBR Safe and a picture was taken with the SynGene G-Box system. The tRNA bands were cut out from the gel, eluted for 24 hours in 100 µl of TBE at 4°C, followed by purification using RNeasy MinElute kit as above. The purified RNAs were run on Pico Bioanalyzer (Agilent) before TGIRT-library preparation.

TGIRT-seq libraries

TGIRT-seq libraries were prepared as in (*22*). Reverse transcription with TGIRT-III (InGex) was initiated from a DNA primer (5′-GTGACTGGAGTTCAGACGTGTGCTCTT CCGATCTTN-3′) encoding the reverse complement of the Illumina Read2 sequencing primer binding site (R2R) annealed to a complementary RNA oligonucleotide (R2) such that there is a single-nucleotide 3′ DNA overhang composed of an equimolar mixture of A, G, C, and T. The RNA oligonucleotide is blocked at its 3′ end with C3Sp (IDT) to inhibit template switching to itself. Reactions contained purified RNAs, reaction medium [20 mM tris-HCl (pH 7.5), 450 mM NaCl, and 5 mM $MgCl_2$], 5 mM dithiothreitol (DTT), 100 nM starting annealed molecule and 1 µM TGIRT-III. Reactions were pre-incubated at room temperature for 30 min and cDNA synthesis was initiated by addition of 1 mM deoxynucleotide triphosphates (an equimolar mix of dATP, 2′-deoxyguanosine 5′-triphosphate, 2′-deoxycytidine 5′-triphosphate, and 3′-deoxythymidine 5′-triphosphate). Reactions were incubated at 60°C for 15 min and were terminated by adding 5 N NaOH to a final concentration of 0.25 N and incubated at 95°C for 3 min to degrade RNAs and denature proteins. The reactions were then cooled to room temperature and neutralized with 5 N HCl. cDNAs were purified by using a Qiagen MinElute Reaction Cleanup Kit and then ligated at their 3′ ends to a DNA oligonucleotide encoding the reverse complement of the Illumina Read1 primer binding site (R1R) using Thermostable 5′ AppDNA/RNA Ligase [New England Biolabs (NEB)]. Ligated cDNAs were repurified with MinElute Reaction Cleanup Kit as above and amplified by PCR for 12 cycles using Phusion DNA polymerase (Thermo Fisher Scientific) with overlapping multiplex and barcode primers that add sequences necessary for Illumina sequencing. PCR reactions were cleaned up with AMPure XP beads (Beckman Coulter) to remove adapter dimers. Libraries were sequenced on a NextSeq 500 instrument (75 nt, paired-end reads) at the Next Generation Sequencing Facility at MD Anderson Science Park.

Bioinformatics

Analysis of TGIRT-seq small RNA libraries

TGIRT-seq small RNA libraries were analyzed as in (*22*). For all samples, reads were adapter trimmed using Cutadapt version 3.2 from both the 3′ end (adapter AGATCGGAAGAG) and 5′ end (adapter GATCGTCGGACT), and requiring a minimum read size of 18 base pairs (bp). Read pairs were merged using Flash version 1.2.11 requiring a minimum of 5 bp of overlap, and only merged fragments were used for alignment. We assembled a custom human small RNA reference from several published databases, including nuclear tRNAs from tRNAdb, mitochondrial tRNAs from mitotRNAdb, piwiRNAs from piRNAdb, microRNAs from miRBase, and ribosomal RNA from annotations in the UCSC Genome Table Browser. We padded both ends of each sequence with 10 N's (to capture nontemplated end modifications), concatenated the resulting FASTA files, and built a reference genome using STAR version 2.5.2b with default parameters, except for the following: --genomeSAindexNbases 7 --genomeChrBinNbits 10 (to accommodate the high count and short length of human sRNA sequences). We mapped all processed libraries against this reference using STAR version 2.5.2b with default parameters, except for the following: --outSAMprimaryFlag AllBestScore --seedSearchStartLmax 15 --outFilterScoreMinOverLread 0.25 --seedSearchLmax 15. Counts per sRNA were obtained using SAMtools idxstats. Allelic counts were obtained using the pysamstats utility (https://github.com/alimanfoo/pysamstats) and detected variants were filtered to require a depth of at least 50 reads in each sample. As described in fig. S7, the allelic proportions were used to compute Bhattacharya coefficients, in-group and out-group distances, and lastly, Bhattacharya distance differences. As seen in Fig. 2D, this method worked well to find known and unknown sites responsive to $NaCNBH_3$, NAT10-KO, or both.

Analysis of ac^4C-seq libraries

For all samples, reads were adapter trimmed using Cutadapt version 3.2 from both the 3′ end (adapter AGATCGGAAGAGCACACGTC) and 5′ end (adapter AGATCGGAAGAGCGTCGTG), and requiring a minimum read size of 18 bp. Trimmed reads were aligned against the hg38 reference genome using STAR version 2.5.2b with default parameters, sorted, and indexed using SAMtools. Gene counts were generated against the Gencode version 21 annotation set using featureCounts only counting fragments identified as Primary by STAR. Allele counts were generated using bcftools mpileup and required a read to be uniquely aligned and have base quality >20. Allelic counts were filtered to remove indels and exported to tab-delimited text files using bcftools filter, bcftools norm, and bcftools query.

Differentially expressed genes. DEGs were identified from gene counts using DESeq2 with default parameters except that a DEG was required to have both an absolute fold change of 1.5 or greater and a statistically significant false discovery rate–adjusted *P* value <0.05. We analyzed effects that are NAT10 dependent, effects that are $NaCNBH_3$ dependent, and those that are both (Fig. 2E) by performing contrasts on the same DESeq2 model fit. All final results were exported to Excel and all downstream plotting and analysis was performed with custom scripts in R with the ggplot2 graphics package, or IPA as indicated.

Allelic count data. To handle allelic count data, we separated the counts per variant site by reference and alternate allele count. The

resulting dataset was filtered to require that (i) variants are single-base substitutions, (ii) the total depth per site must be greater than 20, and (iii) a site must pass this threshold in at least four samples. Differential allelic expression analyses were performed with DESeq2 as described in https://rpubs.com/mikelove/ase, resulting in a set of estimated changes in allelic ratio between groups (as when comparing NAT10-KD to WT) or that are specific to the interaction of certain groups (as when testing for allelic expression specific to samples that are both $NaCNBH_3$ treated and WT). As with differential gene expression, these were computed with contrasts on the same DESeq2 model fit and processed with custom scripts in R. All the datasets have been uploaded to Gene Expression Omnibus (GEO) and can be accessed with the following accession number: GSE274993.

In vitro R-loop preparation

The in vitro R-loops were prepared by in vitro transcription of the pFC53-mAIRN plasmid with the T3 RNA polymerase. Reactions (45.5 μl) containing 3 μg of pFC53-mAIRN plasmid, 1× transcription optimized buffer from Promega [no. P1181: 40 mM tris-HCl (pH 7.9), 6 mM $MgCl_2$, 2 mM spermidine, and 10 mM NaCl], 20 mM DTT, 0.05% Tween-20, and 0.25 mM rNTP were set up in 1.5-ml tubes on ice. Transcription was initiated by adding 4.5 μl (160 U) of T3 RNA polymerase (Promega, no. P4024). Samples were split in two tubes of 25 μl each and incubated at 37°C for 30 min and heat inactivated at 65°C for 10 min. One of the tubes was treated with RNase A (0.01 μg/μl), while the other tube was treated with RNase A (0.01 μg/μl) and 2 μl of RNase H (NEB, no. M0297S) at 37°C for 30 min. Both tubes were treated with 2 μl of proteinase K (Roche, catalog no. 03115828001) at 37°C for 30 min. To verify R-loop quality, 4 μl of each reaction was supplemented with 1 μl of 50% glycerol, loaded on a 0.9% agarose, 1× TBE gel, and run at 100 V for 1 hour and stained with 1× SYBR. The gel was imaged using a Syngene G:box. After gel verification of R-loop formation, substrates were purified with a G50 column and quantified using a DeNovix DS-11 spectrophotometer.

R-loop dot blot (S9.6/dsDNA/ac^{4}C)

Five micrograms of total gDNA extracted from HeLa WT and NAT10-KO cells as in (*27*) was either treated with mock or RNase H or excess RNase A/T1 cocktail in a 1.5-ml tube for 30 min on a 37°C heated Eppendorf mixer at 300 rpm. For mock treatment, RNase H buffer was used without adding RNase H. For RNase H treatment, 7 μl of RNase H (NEB, M0297S) was used in provided buffer. For excess RNase A/T1 cocktail treatment, 4 μl of RNase A (10 μg/μl) and RNase T1 (1000 U/μl, no. EN0541, Thermo Fisher Scientific) was used in 1× PBS. Each reaction was scaled up to 100-μl volume. After the reaction, 1 μl of 0.5 M EDTA was added to each tube. Subsequently, gDNAs were purified as in (*27*). The indicated amount of purified gDNAs was blotted on BrightStar-Plus Positively Charged Nylon Membrane. Once the membranes were dried, DNA was cross-linked on the membranes twice with 150 mJ/cm^2 ultraviolet (UV) in a Stratagene UV cross-linker. The membranes were blocked with 5% nonfat milk in PBST (PBS + 0.1% Tween) for 30 min at room temperature on a shaker. Then, the membranes were probed with the S9.6 or the dsDNA or ac^{4}C antibody diluted 1:1000 in PBS + 1% milk and incubated overnight on the rocker at 4°C. The membranes were washed three times with 10 ml of PBST, 5 min each wash. The membranes were incubated with horseradish peroxidase–conjugated appropriate secondary antibody (1:10,000) in PBST + 1% milk, overnight at 4°C. The membranes were washed three times with 10 ml of PBST, each wash for 5 min on the shaker, and lastly detected with ECL reagents (Amersham) in the Syngene G:box or developed on film.

DRIP-seq with S9.6 antibody

Ten milliliters of HeLa WT (5×10^4 cells/ml) and HeLa NAT10 KO (7.5×10^4 cells/ml) cells were plated in 10-cm plates and cultured for 3 days. gDNA was purified and used for S9.6-based DRIP-seq according to the protocol described in Sanz and Chédin (*27*). Forty nanograms of input diluted in 40 μl of TE buffer and 40 μl of DRIP sample from the DRIP experiment described above were subjected to sonication using a Bioruptor for 12 cycles with 15-s ON and 60-s OFF, using the "high" setting. The sonicated samples were used to prepare libraries using the NEBNext Ultra II kit and libraries were quality checked using the bioanalyzer 2100. The barcoded libraries were subjected to sequencing using the NovaSeq S1 instrument to obtain 100-nt-long single-end reads at the Genomic Sequencing and Analysis Facility at UT Austin. The adapter sequences in each raw sequencing file (fastq) were trimmed using Trim Galore (https://github.com/FelixKrueger/TrimGalore). The quality of each processed fastq file was then assessed with FastQC (www.bioinformatics.babraham.ac.uk/projects/fastqc/). Subsequently, the reads were aligned to the reference human genome hg38 using the bwa mem alignment method (https://github.com/lh3/bwa). SAMtools was used to remove low-quality mapped reads by filtering out reads with a mapping quality score of less than 10, as well as PCR duplicates (*37*). RLPipes was applied to generate bw and broadPeak files (https://github.com/Bishop-Laboratory/RLPipes). DeepTools computeMatrix and plotHeatmap were used for generating heatmaps of the defined genomic sites of each sample (*38*). The overlapped regions between two sets of genomic features (bed files) were identified using BEDTools intersect (*39*). All the datasets have been uploaded to GEO and can be accessed with the following accession numbers: GSE273067 and GSE273230.

Cellular fractionation

HeLa WT and NAT10-KO cells were plated in 100-mm plates and collected on the third day. WT cells (2×10^6 cells) and NAT10-KO cells (3×10^6) were used for cellular fractionation. After trypsinization, cells were collected by centrifugation at 200*g* for 5 min at 4°C and washed twice with 5 ml of cold PBS. The cells were then fractionated with the Subcellular Protein Fractionation Kit for Cultured Cells from Thermo Fisher Scientific (no. 78840) according to the manufacturer's instructions. RNA from cytoplasmic fraction was purified using the Qiagen RNeasy MinElute Cleanup Kit with a modified protocol that allows recovery of RNAs of all sizes (*36*). To extract the RNA from the membrane bound fraction, proteinase K (recombinant, PCR Grade, Roche, no. 3115887001) was used (37°C for 30 min), before RNA purification with Qiagen RNeasy MinElute Cleanup Kit as above.

Immunofluorescence

One milliliter of HeLa WT (5×10^4 cells/ml) and HeLa NAT10 KO (7.5×10^4 cells/ml) cells were plated on each well of a 4- or 24-well culture plate containing a BioCoat 12 mm #1 German Glass Coverslip coated with poly-L-lysine (Corning, no. 354085) and then grown for 48 hours in a 37°C, 5% CO_2 incubator. The medium was aspirated, coverslips were washed twice with 1 ml of 37°C heated PBS, and then immediately fixed with 500 μl of prewarmed 2% methanol-free formaldehyde (Polysciences, no. 18814) diluted in PBS with gentle rotation for 10 min. The fixed cells were washed two times with 1 mL of 1 × PBS,

once with 1 ml of PBS + 3% bovine serum albumin (BSA) and then were permeabilized with 1 ml of PBS + 3% BSA + 0.3% Tween-20 for 15 min with gentle rotation. The permeabilized cells were washed twice for 5 min each with 1 ml of 1× PBS. Then, 200 µl of 65°C heated PBS was added to each coverslip and incubated in 65°C heated chamber for 10 min with gentle shaking. In this step, PBS (mock) or ~33 µg of RNase A was used as control. After exactly 10-min incubation, the coverslips were snap-chilled on ice for at least 2 to 3 min. Then, each coverslip was washed with 1 ml of PBS + 3% BSA and incubated for 18 hours with 200 or 250 µl of PBS + 3% BSA + 1/200 ac^4C-specific antibody at 4°C overnight with gentle shaking. The following day, the coverslips were washed two times for 5 min each with 1 ml of 1× PBS, once with 1 ml of PBS + 3% BSA and incubated with 250 µl of PBS + 3% BSA + 1/1000 compatible secondary antibody + Hoechst 33258 at a final concentration of 1 µg/ml for 1 hour in the dark. The cells were washed four times with 1 ml of 1× PBS, mounted using Vectashield and sealed with nail polish. In all cases, washes were performed for 5 min each on a rocking platform. All slides were stored at 4°C before and after imaging, and all images were acquired using an inverted Olympus Fluoview 1000 confocal microscope.

For GFP-plasmid transfection followed by IF experiments, 24 hours after plating as above, cells were transfected with 200 to 250 ng of plasmid DNA using Lipofectamine 3000 reagent according to the manufacturer's instructions (Thermo Fisher Scientific). For IF experiments that did not involve the ac^4C antibody, PBS + 3% BSA + 0.3% Triton X-100 solution was used for permeabilization and the 65°C heating step was not performed. Foci in the IF images were quantified with Fiji software and graphs were plotted using Prism 9.

Supplementary Materials

The PDF file includes:

Legends for tables S1 to S4
Tables S5 and S6
Figs. S1 to S14
References

Other Supplementary Material for this manuscript includes the following:

Tables S1 to S4

REFERENCES AND NOTES

1. S. Ito, S. Horikawa, T. Suzuki, H. Kawauchi, Y. Tanaka, T. Suzuki, T. Suzuki, Human NAT10 is an ATP-dependent RNA acetyltransferase responsible for N4-acetylcytidine formation in 18 S ribosomal RNA (rRNA). *J. Biol. Chem.* **289**, 35724–35730 (2014).
2. P. Boccaletto, F. Stefaniak, A. Ray, A. Cappannini, S. Mukherjee, E. Purta, M. Kurkowska, N. Shirvanizadeh, E. Destefanis, P. Groza, G. Avşar, A. Romitelli, P. Pir, E. Dassi, S. G. Conticello, F. Aguilo, J. M. Bujnicki, MODOMICS: A database of RNA modification pathways. 2021 update. *Nucleic Acids Res.* **50**, D231–d235 (2022).
3. A. Sas-Chen, J. M. Thomas, D. Matzov, M. Taoka, K. D. Nance, R. Nir, K. M. Bryson, R. Shachar, G. L. S. Liman, B. W. Burkhart, S. T. Gamage, Y. Nobe, C. A. Briney, M. J. Levy, R. T. Fuchs, G. B. Robb, J. Hartmann, S. Sharma, Q. Lin, L. Florens, M. P. Washburn, T. Isobe, T. J. Santangelo, M. Shalev-Benami, J. L. Meier, S. Schwartz, Dynamic RNA acetylation revealed by quantitative cross-evolutionary mapping. *Nature* **583**, 638–643 (2020).
4. S. Sharma, J. L. Langhendries, P. Watzinger, P. Kötter, K. D. Entian, D. L. J. Lafontaine, Yeast Kre33 and human NAT10 are conserved 18S rRNA cytosine acetyltransferases that modify tRNAs assisted by the adaptor Tan1/THUMPD1. *Nucleic Acids Res.* **43**, 2242–2258 (2015).
5. M. L. Bortolin-Cavaillé, A. Quillien, S. Thalalla Gamage, J. M. Thomas, A. Sas-Chen, S. Sharma, C. Plisson-Chastang, L. Vandel, P. Blader, D. L. J. Lafontaine, S. Schwartz, J. L. Meier, J. Cavaillé, Probing small ribosomal subunit RNA helix 45 acetylation across eukaryotic evolution. *Nucleic Acids Res.* **50**, 6284–6299 (2022).
6. D. Arango, D. Sturgill, N. Alhusaini, A. A. Dillman, T. J. Sweet, G. Hanson, M. Hosogane, W. R. Sinclair, K. K. Nanan, M. D. Mandler, S. D. Fox, T. T. Zengeya, T. Andresson, J. L. Meier, J. Coller, S. Oberdoerffer, Acetylation of cytidine in mRNA promotes translation efficiency. *Cell* **175**, 1872–1886.e24 (2018).
7. D. Arango, D. Sturgill, R. Yang, T. Kanai, P. Bauer, J. Roy, Z. Wang, M. Hosogane, S. Schiffers, S. Oberdoerffer, Direct epitranscriptomic regulation of mammalian translation initiation through N4-acetylcytidine. *Mol. Cell* **82**, 2797–2814.e11 (2022).
8. H. Beiki, D. Sturgill, D. Arango, S. Relier, S. Schiffers, S. Oberdoerffer, Detection of ac4C in human mRNA is preserved upon data reassessment. *Mol. Cell* **84**, 1611–1625.e3 (2024).
9. J. Georgeson, S. Schwartz, No evidence for ac4C within human mRNA upon data reassessment. *Mol. Cell* **84**, 1601–1610 (2024).
10. J. Lv, H. Liu, Q. Wang, Z. Tang, L. Hou, B. Zhang, Molecular cloning of a novel human gene encoding histone acetyltransferase-like protein involved in transcriptional activation of hTERT. *Biochem. Biophys. Res. Commun.* **311**, 506–513 (2003).
11. Q. Shen, X. Zheng, M. A. McNutt, L. Guang, Y. Sun, J. Wang, Y. Gong, L. Hou, B. Zhang, NAT10, a nucleolar protein, localizes to the midbody and regulates cytokinesis and acetylation of microtubules. *Exp. Cell Res.* **315**, 1653–1667 (2009).
12. D. Larrieu, S. Britton, M. Demir, R. Rodriguez, S. P. Jackson, Chemical inhibition of NAT10 corrects defects of laminopathic cells. *Science* **344**, 527–532 (2014).
13. G. Balmus, D. Larrieu, A. C. Barros, C. Collins, M. Abrudan, M. Demir, N. J. Geisler, C. J. Lelliott, J. K. White, N. A. Karp, J. Atkinson, A. Kirton, M. Jacobsen, D. Clift, R. Rodriguez, Sanger Mouse Genetics Project, C. Shannon, M. Sanderson, A. Gates, J. Dench, V. Vancollie, C. McCarthy, S. Pearson, E. Cambridge, C. Isherwood, H. Wilson, E. Grau, A. Galli, Y. E. Hooks, C. L. Tudor, A. L. Green, F. L. Kussy, E. J. Tuck, E. J. Siragher, R. S. B. McLaren, A. Swiatkowska, S. S. Caetano, C. I. Mazzeo, M. H. Dabrowska, S. A. Maguire, D. T. Lafont, L. F. E. Anthony, M. T. Sumowski, J. Bussell, C. Sinclair, E. Brown, B. Doe, H. Wardle-Jones, N. Griggs, M. Woods, H. Kundi, G. McConnell, J. Doran, M. N. D. Griffiths, C. Kipp, S. A. Holroyd, D. J. Gannon, R. Alcantara, R. Ramirez–Solis, J. Bottomley, C. Ingle, V. Ross, D. Barrett, D. Sethi, D. Gleeson, J. Burvill, R. Platte, E. Ryder, E. Sins, E. Miklejewska, D. von Schiller, G. Duddy, J. Urbanova, K. Boroviak, M. Imran, S. K. Reddy, D. J. Adams, S. P. Jackson, Targeting of NAT10 enhances healthspan in a mouse model of human accelerated aging syndrome. *Nat. Commun.* **9**, 1700 (2018).
14. Q. Li, X. Liu, K. Jin, M. Lu, C. Zhang, X. du, B. Xing, NAT10 is upregulated in hepatocellular carcinoma and enhances mutant p53 activity. *BMC Cancer* **17**, 605 (2017).
15. X. Liu, Y. Tan, C. Zhang, Y. Zhang, L. Zhang, P. Ren, H. Deng, J. Luo, Y. Ke, X. du, NAT10 regulates p53 activation through acetylating p53 at K120 and ubiquitinating Mdm2. *EMBO Rep.* **17**, 349–366 (2016).
16. J. M. Thomas, C. A. Briney, K. D. Nance, J. E. Lopez, A. L. Thorpe, S. D. Fox, M.-L. Bortolin-Cavaille, A. Sas-Chen, D. Arango, S. Oberdoerffer, J. Cavaille, T. Andresson, J. L. Meier, A chemical signature for cytidine acetylation in RNA. *J. Am. Chem. Soc.* **140**, 12667–12670 (2018).
17. T. K. Debnath, B. Xhemalçe, Deciphering RNA modifications at base resolution: From chemistry to biology. *Brief. Funct. Genomics* **20**, 77–85 (2021).
18. S. Yan, Z. Lu, W. Yang, J. Xu, Y. Wang, W. Xiong, R. Zhu, L. Ren, Z. Chen, Q. Wei, S.-M. Liu, T. Feng, B. Yuan, X. Weng, Y. Du, X. Zhou, Antibody-free fluorine-assisted metabolic sequencing of RNA N^4-acetylcytidine. *J. Am. Chem. Soc.* **145**, 22232–22242 (2023).
19. H. Gaillard, A. Aguilera, Transcription as a threat to genome integrity. *Annu. Rev. Biochem.* **85**, 291–317 (2016).
20. S. S. Diebold, T. Kaisho, H. Hemmi, S. Akira, C. Reis e Sousa, Innate antiviral responses by means of TLR7-mediated recognition of single-stranded RNA. *Science* **303**, 1529–1531 (2004).
21. F. Heil, H. Hemmi, H. Hochrein, F. Ampenberger, C. Kirschning, S. Akira, G. Lipford, H. Wagner, S. Bauer, Species-specific recognition of single-stranded RNA via toll-like receptor 7 and 8. *Science* **303**, 1526–1529 (2004).
22. C. W. Reinsborough, H. Ipas, N. S. Abell, R. M. Nottingham, J. Yao, S. K. Devanathan, S. B. Shelton, A. M. Lambowitz, B. Xhemalçe, BCDIN3D regulates tRNAHis 3′ fragment processing. *PLOS Genet.* **15**, e1008273 (2019).
23. H. Xu, R. M. Nottingham, A. M. Lambowitz, TGIRT-seq protocol for the comprehensive profiling of coding and non-coding RNA Biotypes in cellular, extracellular vesicle, and plasma RNAs. *Bio Protoc.* **11**, e4239 (2021).
24. T. Suzuki, Y. Yashiro, I. Kikuchi, Y. Ishigami, H. Saito, I. Matsuzawa, S. Okada, M. Mito, S. Iwasaki, D. Ma, X. Zhao, K. Asano, H. Lin, Y. Kirino, Y. Sakaguchi, T. Suzuki, Complete chemical structures of human mitochondrial tRNAs. *Nat. Commun.* **11**, 4269 (2020).
25. J. L. M. Cann, A. Cristini, E. K. Law, S. Y. Lee, M. Tellier, M. A. Carpenter, C. Beghè, J. J. Kim, A. Sanchez, M. C. Jarvis, B. Stefanovska, N. A. Temiz, E. N. Bergstrom, D. J. Salamango, M. R. Brown, S. Murphy, L. B. Alexandrov, K. M. Miller, N. Gromak, R. S. Harris, APOBEC3B regulates R-loops and promotes transcription-associated mutagenesis in cancer. *Nat. Genet.* **55**, 1721–1734 (2023).
26. R. Stolz, S. Sulthana, S. R. Hartono, M. Malig, C. J. Benham, F. Chedin, Interplay between DNA sequence and negative superhelicity drives R-loop structures. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 6260–6269 (2019).
27. L. A. Sanz, F. Chédin, High-resolution, strand-specific R-loop mapping via S9.6-based DNA-RNA immunoprecipitation and high-throughput sequencing. *Nat. Protoc.* **14**, 1734–1755 (2019).
28. K. D. Nance, S. T. Gamage, M. M. Alam, A. Yang, M. J. Levy, C. N. Link, L. Florens, M. P. Washburn, S. Gu, J. J. Oppenheim, J. L. Meier, Cytidine acetylation yields a hypoinflammatory synthetic messenger RNA. *Cell Chem. Biol.* **29**, 312–320.e7 (2022).

29. Z. Zhang, U. Ohto, T. Shibata, E. Krayukhina, M. Taoka, Y. Yamauchi, H. Tanji, T. Isobe, S. Uchiyama, K. Miyake, T. Shimizu, Structural analysis reveals that Toll-like receptor 7 is a dual receptor for guanosine and single-stranded RNA. *Immunity* **45**, 737–748 (2016).
30. G. J. Brown, P. F. Cañete, H. Wang, A. Medhavy, J. Bones, J. A. Roco, Y. He, Y. Qin, J. Cappello, J. I. Ellyard, K. Bassett, Q. Shen, G. Burgio, Y. Zhang, C. Turnbull, X. Meng, P. Wu, E. Cho, L. A. Miosge, T. D. Andrews, M. A. Field, D. Tvorogov, A. F. Lopez, J. J. Babon, C. A. López, Á. Gónzalez-Murillo, D. C. Garulo, V. Pascual, T. Levy, E. J. Mallack, D. G. Calame, T. Lotze, J. R. Lupski, H. Ding, T. R. Ullah, G. D. Walters, M. E. Koina, M. C. Cook, N. Shen, C. de Lucas Collantes, B. Corry, M. P. Gantier, V. Athanasopoulos, C. G. Vinuesa, TLR7 gain-of-function genetic variation causes human lupus. *Nature* **605**, 349–356 (2022).
31. M. P. Gantier, S. Tong, M. A. Behlke, D. Xu, S. Phipps, P. S. Foster, B. R. G. Williams, TLR7 is involved in sequence-specific sensing of single-stranded RNAs in human macrophages. *J. Immunol.* **180**, 2117–2124 (2008).
32. H. Wu, W. F. Lima, S. T. Crooke, Properties of cloned and expressed human RNase H1. *J. Biol. Chem.* **274**, 28270–28278 (1999).
33. Y. Y. Yang, J. M. Ascano, H. C. Hang, Bioorthogonal chemical reporters for monitoring protein acetylation. *J. Am. Chem. Soc.* **132**, 3640–3641 (2010).
34. A. Cristini, M. Groh, M. S. Kristiansen, N. Gromak, RNA/DNA hybrid interactome identifies DXH9 as a molecular player in transcriptional termination and R-loop-associated DNA damage. *Cell Rep.* **23**, 1891–1905 (2018).
35. B. Xhemalce, K. M. Miller, N. Gromak, Epitranscriptome in action: RNA modifications in the DNA damage response. *Mol. Cell* **84**, 3610–3626 (2024).
36. B. Xhemalce, S. C. Robson, T. Kouzarides, Human RNA methyltransferase BCDIN3D regulates microRNA processing. *Cell* **151**, 278–288 (2012).
37. P. Danecek, J. K. Bonfield, J. Liddle, J. Marshall, V. Ohan, M. O. Pollard, A. Whitwham, T. Keane, S. A. McCarthy, R. M. Davies, H. Li, Twelve years of SAMtools and BCFtools. *GigaScience* **10**, giab008 (2021).
38. F. Ramírez, D. P. Ryan, B. Grüning, V. Bhardwaj, F. Kilpert, A. S. Richter, S. Heyne, F. Dündar, T. Manke, deepTools2: A next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).
39. A. R. Quinlan, I. M. Hall, BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
40. S. M. Tan-Wong, S. Dhir, N. J. Proudfoot, R-Loops promote antisense transcription across the mammalian genome. *Mol. Cell* **76**, 600–616.e6 (2019).
41. Y. Liu, P. Siejka-Zielińska, G. Velikova, Y. Bi, F. Yuan, M. Tomkova, C. Bai, L. Chen, B. Schuster-Böckler, C.-X. Song, Bisulfite-free direct detection of 5-methylcytosine and 5-hydroxymethylcytosine at base resolution. *Nat. Biotechnol.* **37**, 424–429 (2019).
42. S. T. Gamage, A. Sas-Chen, S. Schwartz, J. L. Meier, Quantitative nucleotide resolution profiling of RNA cytidine acetylation by ac4C-seq. *Nat. Protoc.* **16**, 2286–2307 (2021).

Acknowledgments: We thank M. Hentze (EMBL) and M. Price (University of Iowa) for the sharing of the HeLa Flp-In T-REx cell line; S. Oberdoerffer (NCI Intramural Program) for the sharing of the HeLa NAT10-KO cell line and S. Mukhopadhyay (UT Austin) for the sharing of the plasmids expressing GFP-tagged RAB5, RAB7 and RAB11; K. M. Miller (UT Austin) for the use of confocal microscope; and R. Nottingham (Lambowitz lab at UT Austin) for advice on the TGIRT-seq protocol. **Funding:** This work was supported by NIH Grant R01 GM127802, and an Associate Professor Experimental (APX) grant from the office of the Vice President for Research at the University of Texas at Austin, USA. **Author contributions:** T.K.D. and B.X. conceptualized the study; N.S.A., Y.-R.L., and B.X. performed bioinformatic analyses; T.K.D., E.N., S.K.D., and Y.-R.L. performed cloning; S.K.D., Y.-R.L., and B.X. constructed stable cell lines; E.N. generated the in vitro R-loops; T.K.D. and S.K.D. set up R-loop assays; T.K.D. performed all DRIP-seq experiments and their validation by quantitative PCR, as well as IF analyses; T.K.D. and B.X. wrote the paper, with contributions from all authors. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Next-generation sequencing data have been uploaded to GEO and can be accessed with the following accession numbers: GSE274993, GSE273067, and GSE273230.

Submitted 21 August 2024
Accepted 24 February 2025
Published 26 March 2025
10.1126/sciadv.ads6144