Detection of transcriptome-wide microRNA–target interactions in single cells with agoTRIBE

Vaishnovi Sekar[1], Emilio Mármol-Sánchez [1,2,5], Panagiotis Kalogeropoulos[1,5], Laura Stanicek[1], Eduardo A. Sagredo [1], Albin Widmark[3], Evangelos Doukoumopoulos[4], Franziska Bonath[1], Inna Biryukova [1] ✉ & Marc R. Friedländer [1] ✉

MicroRNAs (miRNAs) exert their gene regulatory effects on numerous biological processes based on their selection of target transcripts. Current experimental methods available to identify miRNA targets are laborious and require millions of cells. Here we have overcome these limitations by fusing the miRNA effector protein Argonaute2 to the RNA editing domain of ADAR2, allowing the detection of miRNA targets transcriptome-wide in single cells. miRNAs guide the fusion protein to their natural target transcripts, causing them to undergo A>I editing, which can be detected by sensitive single-cell RNA sequencing. We show that agoTRIBE identifies functional miRNA targets, which are supported by evolutionary sequence conservation. In one application of the method we study microRNA interactions in single cells and identify substantial differential targeting across the cell cycle. AgoTRIBE also provides transcriptome-wide measurements of RNA abundance and allows the deconvolution of miRNA targeting in complex tissues at the single-cell level.

miRNAs are small noncoding RNAs that posttranscriptionally regulate the expression of protein-coding genes[1]. Mechanistically they guide Argonaute effector proteins to messenger RNA targets, allowing Argonaute and cofactors to inhibit translation and/or promote degradation of target mRNAs[2–5]. miRNAs are found in virtually all multicellular animals and plants and play important roles in numerous biological processes, including development, formation of cell identity and human diseases such as cancer[1,6–8]. The human genome harbors hundreds of distinct miRNA genes[9], each of which can putatively regulate hundreds of target genes. The function of each individual miRNA is defined by its specific target repertoire and thus, to understand the function of a given miRNA, it is necessary to map its targets. The current state-of-the-art method to do so is crosslinking-immunoprecipitation-sequencing

(CLIP-seq), which applies ultraviolet light to crosslink the Argonaute protein to its mRNA targets in cells, then isolates the protein using antibodies and uses next-generation sequencing to profile the bound RNA targets[10,11]. This method has brought many new insights to the miRNA field yet it has some inherent limitations. First, because isolation with antibodies is inefficient, it requires as input in the order of millions of cells, making it unsuited for samples with limited material—not to mention in single cells. By averaging over millions of cells, CLIP-seq masks potential cell-to-cell variation. Second, the method is laborious and requires many specialized protocol steps, including ultraviolet crosslinking and immunoprecipitation.

We here present our method agoTRIBE, which circumvents these limitations. We show that our method yields results that are consistent

[1]Science for Life Laboratory, Department of Molecular Biosciences, The Wenner-Gren Institute, Stockholm University, Stockholm, Sweden. [2]Centre for Palaeogenetics, Stockholm, Sweden. [3]Department of Molecular Biosciences, The Wenner-Gren Institute, Stockholm University, Stockholm, Sweden. [4]Department of Medicine, Karolinska Institutet, Solna, Sweden. [5]These authors contributed equally: Emilio Mármol-Sánchez, Panagiotis Kalogeropoulos. ✉e-mail: inna.biryukova@scilifelab.se; marc.friedlander@scilifelab.se
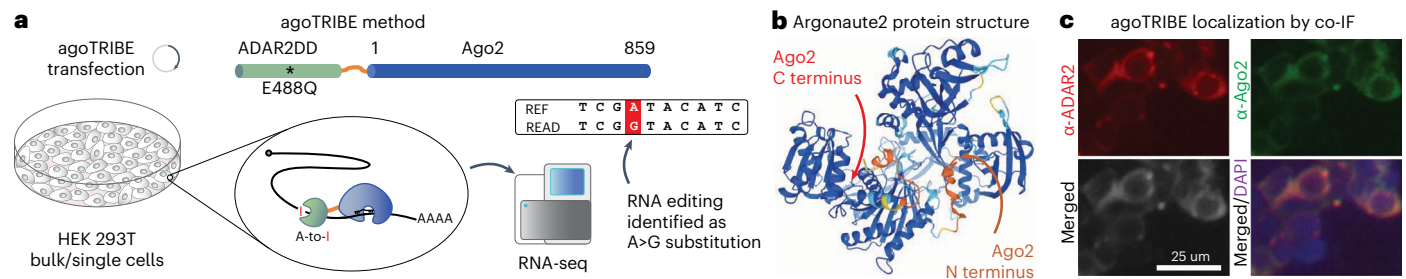
Fig. 1 | agoTRIBE fuses AGO2 and the editing domain of ADAR2. a, Schematic representation of the agoTRIBE method. The agoTRIBE approach fuses human Argonaute2 with the adenosine deaminase domain of human ADAR2, carrying a hyperactive mutation (E488Q) and depositing edits on the targeted transcripts. The edited nucleotides can be detected as A>G substitutions by either standard or single-cell RNA-seq. Top: schematic representation of the N-terminal tagging of human Argonaute2. **b**, Human Argonaute2 protein structure prediction using AlphaFold2. Tagging of the C terminus, which is embedded in the protein structure, could result in a misfolded protein unable to load miRNAs. **c**, Immunofluorescence staining to visualize agoTRIBE. AGO2 (green) and ADAR2 (red) immunofluorescence costaining (co-IF) were used to detect agoTRIBE, while DAPI (magenta) was used for nuclear staining. Each microscopy experiment was performed in at least two replicates and a representative experiment is shown. Colocalization in cytoplasmic foci suggests that agoTRIBE is present in P-bodies.

with the more laborious CLIP-seq method. The reported miRNA targets are supported by evolutionary sequence conservation and are subject to functional miRNA repression. In addition, we show that agoTRIBE can be applied to the detection of miRNA–target interactions in human single cells, and to deconvolution of miRNA targeting in distinct phases of the cell cycle without the need for physical cell sorting.

Results

To develop our method we leveraged on the TRIBE approach[12], in which an RNA-binding protein of interest (in our case, Argonaute2) is fused to the RNA editing domain of ADAR2. The RNA-binding protein part leads the fusion protein to its natural targets while the editing domain deaminates adenosines to inosines (A>I) in the RNA target, in effect leaving nucleotide conversions that can be detected by sequencing as A>G substitutions (Fig. 1). These substitutions can in principle be detected by single-cell RNA sequencing (RNA-seq), and the method avoids lossy isolation because it does not use immunoprecipitation. To tailor agoTRIBE for Argonaute proteins we made three modifications to the original TRIBE approach: (1) we used a hyperactive version of the ADAR2 deaminase domain, in which a E488Q substitution results in increased editing[13,14]; (2) we connected the Argonaute2 and ADAR2 domains with a 55-amino-acids-long flexible linker; and (3) we fused the ADAR2 domain to the N terminus of Argonaute2, because the protein structure of Argonaute2 indicates that fusion to the C terminus would be detrimental to the loading of guide miRNAs[15–19] (Fig. 1b). We confirmed that tagging Argonaute2 with the ADAR2 editing domain does not change its cytoplasmic localization (Fig. 1c), nor its colocalization with TNRC6B, a P-body marker (Supplementary Fig. 1a). In particular, we found that agoTRIBE partly locates to cytoplasmic foci that are similar to P-bodies, which are known to be interconnected with miRNA function[20–22] (Fig. 1c).

When we transiently expressed agoTRIBE in ~50,000 human HEK-293T cells (Methods), we observed that A>G nucleotide substitutions—as expected by ADAR2-mediated editing—increased substantially compared with control cells where editing is low, consistent with ADAR2 being undetectable in this cell type (Fig. 2a, Supplementary Tables 1–4 and Supplementary Fig. 1b–d). In contrast, cells transfected with only the ADAR2 deaminase domain without Argonaute2—henceforward referred as 'ADAR-only'—increased only moderately the number of A>G substitutions, suggesting the importance of miRNA guidance for newly detected editing (Fig. 2a). Of note, other types of nucleotide substitution remained largely unchanged among the analyzed conditions, indicating specific ADAR2-mediated editing. In addition, we observed that editing in mRNA exonic regions specifically

increased while editing in intronic regions and noncoding transcripts such as long noncoding RNAs and pseudogenes remained constant (Fig. 2b,c). This is consistent with miRNAs targeting mature mRNAs in the cytoplasm, while there is little evidence of miRNAs targeting non-translating sequences such as introns or lncRNA transcripts[23], which are most commonly located in the nucleus. We note that the agoTRIBE construct did not appear to substantially change the miRNA profile of transfected cells (Supplementary Fig. 2). In summary, we observed highly increased editing in mRNA transcripts that likely correspond to natural miRNA targets.

We next compared agoTRIBE with state-of-the-art CLIP-seq methods. First, to discern miRNA-guided editing from background editing—including that of endogenously expressed ADAR2—we compared editing patterns following agoTRIBE transfection relative to ADAR-only E488Q controls (Fig. 2d). We assumed that transcript editing specific to agoTRIBE is guided by miRNAs while editing specific to ADAR-only represents editing background activity. To do so, we compared the total number of editing events per gene in the two conditions—agoTRIBE versus ADAR-only—and focused on the top 1,000 putative mRNA targets showing specifically increased editing following agoTRIBE transfection (Fig. 2d). We then compared our list with the top 1,000 targets identified by different variations of the Argonaute CLIP-seq methodology[24–26]. We found that the agoTRIBE top 1,000 targets have substantial overlap with CLIP-seq targets and do not represent an outlier group with an excess of unique targets not found by CLIP-seq methods (Fig. 2e). In fact, the enhanced crosslinking and immunoprecipitation (eCLIP) method reports more targets (624 unique targets) not shared with any other method than does ago-TRIBE (476 unique targets; Fig. 2e). Comparing the similarity between the target sets reported by each of the four methods using the Jaccard index, we found that high-throughput sequencing of RNA isolated by crosslinking and immunoprecipitation (HITS-CLIP) and photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation (PAR-CLIP) resemble each other the most (Jaccard index = 0.25; Fig. 2f), while agoTRIBE also has strong similarity to both of these methods (Jaccard index = 0.21 and 0.18, respectively), eCLIP in contrast had less resemblance to the other two CLIP methods (Jaccard index = 0.14 and 0.10, respectively). More than half of the targets (525) reported by agoTRIBE were supported by one or more CLIP-seq methods (Fig. 2e). We considered that the overlap between methods could be due to undetected biases—for instance, that highly expressed transcripts might be more efficiently detected by both CLIP-seq and agoTRIBE. We therefore compared the overlap between the top 1,000 targets from agoTRIBE and HITS-CLIP with that between our ADAR-only controls
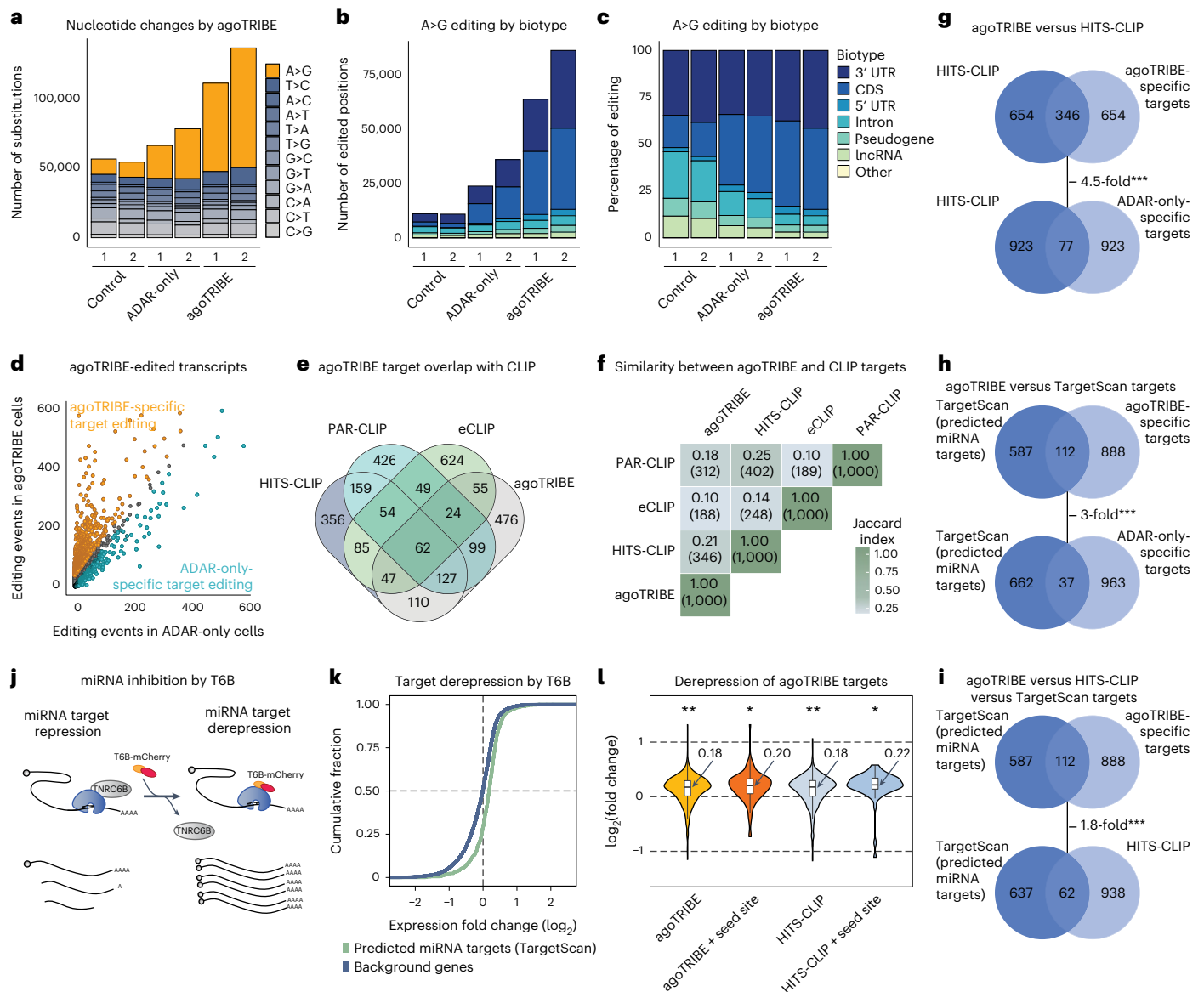
Fig. 2 | agoTRIBE detects miRNA targets through RNA editing. a, A>G editing (colored orange), indicative of ADAR2 editing, specifically increased in agoTRIBE-transfected cells compared with both control cells transfected with eGPF-Ago2 and cells transfected with the ADAR2 editing domain without Argonaute2 (ADAR-only controls). **b**, Editing increased specifically in 3′ UTR and coding sequences (CDS) and remained constant in other transcript types. **c**, As in **b**, but represented as percentages of editing. **d**, Editing in agoTRIBE-transfected versus ADAR-only-transfected cells. Each dot corresponds to one gene. Genes colored orange have substantially more editing in agoTRIBE cells while those colored light blue have substantially more editing in ADAR-only cells (Methods). Genes colored gray have comparable editing in both conditions. **e**, Venn diagram of top 1,000 miRNA targets reported individually by agoTRIBE, HITS-CLIP, PAR-CLIP and eCLIP. **f**, Jaccard similarity index values of top target sets reported by each of the four methods. **g**–**i**, Venn diagrams of overlaps between agoTRIBE and HITS-CLIP targets (**g**); agoTRIBE and TargetScan targets (**h**); and agoTRIBE, HITS-CLIP and TargetScan targets (**i**). **j**, Schematic representation of global miRNA inhibition by T6B. **k**, Derepression of miRNA targets predicted by TargetScan following T6B-mCherry transfection. In total, 609 TargetScan targets and 14,515 background transcripts were profiled. **l**, Increase in expression of agoTRIBE and HITS-CLIP targets following T6B transfection. 'Seed site' indicates that the transcript harbors a conserved binding site for one of the ten most abundant miRNAs in HEK-293T cells. Boxes indicate median and 25th and 75th quantiles, and whiskers indicate lowest and largest values. Significance was calculated using the Wasserstein distance (Methods); P values, left to right: 0.00025, 0.021, 0.00025, 0.012. *$P < 0.05$; **$P < 0.001$. In total, 1,000 agoTRIBE targets (112 with seed sites) and 996 HITS-CLIP targets (62 with seed sites) were profiled.

and HITS-CLIP (Fig. 2g). We found that the overlap between agoTRIBE and HITS-CLIP (346 targets) was 4.5-fold higher ($P < 0.001$, binomial test) than that between our ADAR-only controls and HITS-CLIP (77 targets), indicating that the consistency between methods depends on Argonaute-guided editing and is not due to unspecific biases. Last, it is well established that Argonaute CLIP can identify miRNA binding events at high resolution, in some cases at the level of individual nucleotides[25]. We overlapped the editing positions following agoTRIBE transfection

with reported eCLIP–seq binding sites, finding little consistency in the positional information conferred by the two methods—suggesting that our method might give less precise positional information than does eCLIP (Supplementary Fig. 3). Overall, our results show that agoTRIBE in its reported target repertoire resembles a CLIP-seq method even though it uses a completely distinct, antibody-free approach.

Besides experimental identification of miRNA–target interactions, computational methods that predict target sites are also available and

commonly used[27-29]. These methods typically detect sequence motifs that could confer binding by specific miRNAs, and can also integrate their evolutionary conservation across species. Computational prediction inherently has high false-positive and -negative rates but, on the other hand, predicted miRNA binding sites that are conserved through evolution are likely functional bona fide miRNA target sites. In this way, we next compared our agoTRIBE top targets with transcripts predicted to be regulated by miRNAs according to the widely used TargetScan prediction database[28], finding an overlap of 112 transcripts (Fig. 2h and Methods). This was threefold higher ($P < 0.001$) than the overlap with our ADAR-only control (37 transcripts), indicating that the convergence between agoTRIBE and the computational target prediction is due to Argonaute guidance. Additionally, we found that there was a 1.8-fold higher overlap ($P < 0.001$) between agoTRIBE and TargetScan predictions (112 transcripts) than between HITS-CLIP and TargetScan targets (62 transcripts). This provides sequence conservation evidence that agoTRIBE is more likely to report functional miRNA targets than this CLIP-seq method (Fig. 2i).

To investigate the functionality of agoTRIBE targets further, we performed global inhibition of miRNA action in human HEK-293T cells to detect derepressed targets. Because the TNRC6B protein is an essential cofactor in miRNA-driven, posttranscriptional repression, we performed inhibition of miRNA function by overexpressing the artificial T6B peptide, which effectively occupies the TNRC6B protein-binding pocket on Argonaute and causes global derepression of targets[30,31] (Fig. 2j and Supplementary Fig. 4). We indeed observed a substantial derepression of TargetScan-predicted miRNA targets, suggesting that our perturbation experiment was successful (Fig. 2k and Supplementary Table 2). We next investigated the changes in gene expression of targets reported by agoTRIBE and HITS-CLIP and found that both target sets showed an 0.18 $\log_2$-fold increase in expression when miRNA function was inhibited (Fig. 2l and Supplementary Fig. 5). This became 0.20 and 0.22 $\log_2$-fold increases for the respective methods when only targets supported by sequence motif conservation were considered (Methods). These ~15% increases in expression are consistent with previous studies reporting ~30% derepression at the RNA and protein levels when miRNAs are strongly overexpressed or genetically deleted[32,33]. Furthermore, miRNAs have been proposed to have subtle functions in canalization of gene expression by buffering expression noise[34,35]. In summary, our experiment demonstrates that agoTRIBE predicts functional miRNA targets as efficiently as a state-of-the-art CLIP method.

The agoTRIBE approach tested here expresses an ADAR2 deaminase domain with Argonaute2 as a fusion protein from a single construct. To test the robustness of agoTRIBE we also benchmarked two distinct designs (Supplementary Fig. 6). The first relies on simultaneous expression and dimerization in living cells of an enhanced green fluorescent protein (eGFP)-tagged ADAR2 editing domain and a GFP nanobody fused with Argonaute2 (ref. 36). The second design consists of synthetic coiled-coil E3- and K3-tags, which enable heterodimer formation when Argonaute2 and the ADAR2 editing domain are individually coexpressed in living cells[37]. We found that both approaches increased endogenous editing relative to ADAR-only controls, demonstrating the robustness of agoTRIBE (Supplementary Fig. 7). In particular, these types of modular design give flexibility to experiments and allow an easy exchange of modifying domains (for example, TRIBE[12], hyperTRIBE[13,14] and STAMP[38,39]) and RNA-binding proteins of interest.

To test the sensitivity limits of our method we subjected agoTRIBE-transfected individual cells to Smart-seq3 single-cell RNA-seq[40]. In total, we profiled the transcriptomes of 703 agoTRIBE-transfected cells and 26 control HEK-293T cells following stringent quality controls (Methods and Supplementary Tables 5–7). Because the agoTRIBE construct includes an artificial linker region that is sequenced along with the transcriptomes, we could use sequence reads that map to the linker as an estimate of transfection efficiency of individual cells (Fig. 3a). As expected, many linker reads were detected in agoTRIBE-transfected cells while few linker reads—probably the result of mapping artifacts—were detected in control cells (Fig. 3b, left). Overall, transcriptome-wide editing increased substantially in agoTRIBE-transfected cells, with an average of ~32,500 editing events in each cell compared with ~1,900 in control cells (Fig. 3b, right). We also found that agoTRIBE cells with few linker reads tend to accumulate little editing (Fig. 3c), suggesting that some cells might not be efficiently transfected. However, these cells could easily be computationally identified and discarded, leaving a total of 540 efficiently transfected and edited cells for downstream analyses (Fig. 3c, dashed lines). We found that the transcriptional profiles of the agoTRIBE-transfected cells overall resemble those of control cells (Fig. 3d, light and dark gray, respectively). In contrast, control cells belonging to the fast-growing HEK-293FT cell line, but sequenced with the same protocol, clearly clustered separately from our control and agoTRIBE cells (Fig. 3d, light blue). These results suggest that editing by our fusion protein does not substantially alter transcriptome composition—even in measurements with sensitive single-cell methods such as the Smart-seq3 protocol.

We observed that specific miRNA targets increase strongly in editing in single cells. For instance, *SBNO1* has three binding sites for miR-92, which is a highly expressed miRNA in HEK-293T cells (Fig. 3e, blue arrows). The density plots of sequenced transcript parts show that the sensitive Smart-seq3 protocol yields transcript information for much of the 3′ untranslated region (UTR) in a control cell and an agoTRIBE-transfected cell (Fig. 3e, blue densities). Editing is virtually absent in the control cell but prevalent in the agoTRIBE cell (Fig. 3e, editing colored orange). Besides *SBNO1*, other high-confidence miRNA targets also have substantially increased editing in single cells

Fig. 3 | miRNA targeting in single cells. a, The agoTRIBE transcript includes an artificial linker that is detected in RNA-seq and can be used to estimate agoTRIBE levels. **b**, Levels of linker sequencing reads (left) and editing events (right) detected in 703 agoTRIBE-transfected cells and 26 control cells profiled by Smart-seq3 single-cell RNA-seq. Boxes indicate median and 25th and 75th quantiles, and whiskers indicate lowest and highest values. Significance was estimated using the Wilcoxon rank-sum test, with $P < 0.0001$. ***$P < 0.0001$. **c**, Normalized number of linker reads versus number of editing events for individual cells. The number of linker reads was normalized to the sequencing depth of each cell. Dashed lines indicate thresholds for agoTRIBE cells used for downstream analyses. **d**, Principal component (PC) analysis of agoTRIBE-transfected cells and control cells versus control cells from a different human embryonic kidney cell line (HEK-293FT). Cells were positioned based on their Smart-seq3 transcription profiles. **e**, Editing patterns of the *SBNO1* transcript in a control cell versus an agoTRIBE-transfected cell. **f**, Overview of 540 agoTRIBE-transfected single cells assigned to cell cycle stages using their Smart-seq3 profiles. Each cell cycle stage is divided into an early and late phase (Methods).

Dimensionality reduction was performed with the UMAP algorithm. **g**, Gene expression for cells assigned to the G1 cell cycle stage. The cells originate from two distinct replicate plates of single cells. Each dot represents one gene. **h**, Transcript editing for cells assigned to the G1 cell cycle stage. **i**, Overview of estimated overall miRNA targeting during the cell cycle. Expression values and editing events were normalized to spike-ins and to the number of cells in each cell cycle stage (Methods). **j–u**, Examples of transcript expression and estimated miRNA targeting during the cell cycle: *CCNB1* (**j**), *CCNA2* (**k**), *CDK1* (**l**), *CCNE2* (**m**), *CDK2* (**n**), *ACTB* (**o**), *GAPDH* (**p**), *CENPA* (**q**), *CDK6* (**r**), *RMND5A* (**s**), *IRF2BP2* (**t**), *SPDYA* (**u**). Transcript expression is colored blue and miRNA targeting yellow. **v**, miR-144 and miR-451 increased in expression during K562 differentiation into erythroid precursor cells. **w**, Pseudotime trajectory of the differentiation process. Each dot indicates a single cell profiled by Smart-seq3. **x,y**, Editing in agoTRIBE-transfected cells during differentiation: EIF2S3 (**x**), TOR1AIP2 (**y**). **z**, Expression (read counts) and targeting (editing events) in TOR1AIP2. Each dot represents a single cell.
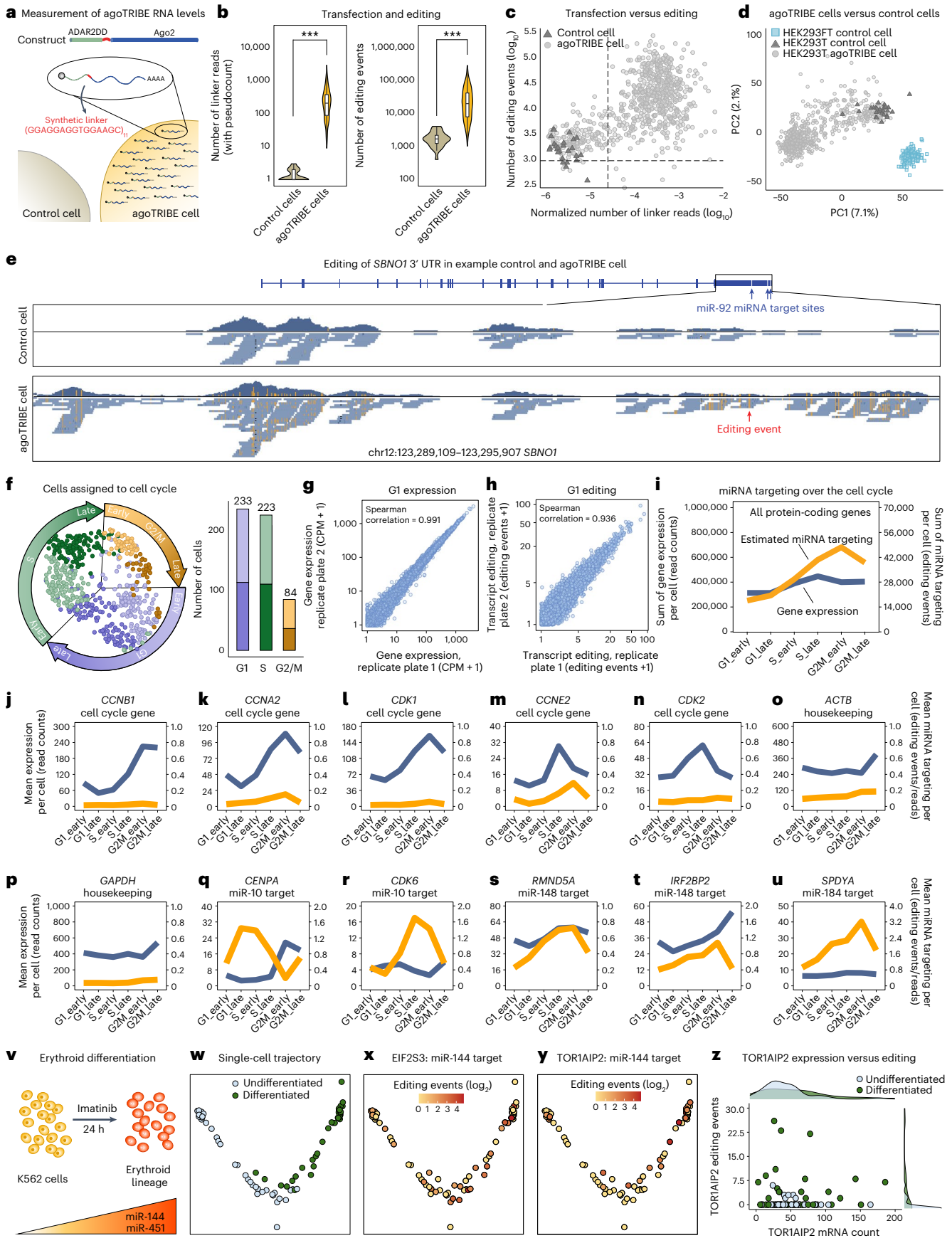
a Measurement of agoTRIBE RNA levels

b Transfection and editing

c Transfection versus editing

d agoTRIBE cells versus control cells

e Editing of *SBNO1* 3′ UTR in example control and agoTRIBE cell

chr12:123,289,109–123,295,907 *SBNO1*

f Cells assigned to cell cycle

g G1 expression

h G1 editing

i miRNA targeting over the cell cycle

j *CCNB1* cell cycle gene

k *CCNA2* cell cycle gene

l *CDK1* cell cycle gene

m *CCNE2* cell cycle gene

n *CDK2* cell cycle gene

o *ACTB* housekeeping

p *GAPDH* housekeeping

q *CENPA* miR-10 target

r *CDK6* miR-10 target

s *RMND5A* miR-148 target

t *IRF2BP2* miR-148 target

u *SPDYA* miR-184 target

v Erythroid differentiation

w Single-cell trajectory

x EIF2S3: miR-144 target

y TOR1AIP2: miR-144 target

z TOR1AIP2 expression versus editing

Table 1 | Comparison between Argonaute CLIP-seq methods and agoTRIBE

Method	Argonaute CLIP-seq	agoTRIBE
Starting input	Millions of cells	Down to single cells
Approach	Immunoprecipitation in lysed cells	Editing in living cells
Antibodies	Needed	Not needed
Specialized laboratory equipment	Needed	Not needed
Duration of protocol (excluding cell work)	>4 days	1.5–2.0 days
Exact miRNA binding positions	Available	Not available
Extra information	–	Provides RNA-seq data

transfected with agoTRIBE, showing the generality of our observations (Supplementary Fig. 8). As a proof of principle of the biological applications of our methods, we next applied Seurat[41] to computationally sort the 540 agoTRIBE-transfected single cells into the G1, S and G2/M stages of the cell cycle (Fig. 3f). Even when stratifying cells into distinct cell cycle stages, our measurements of gene expression (Fig. 3g) and editing (Fig. 3h) were still highly reproducible. We found that overall normalized gene expression remained constant over the cell cycle while miRNA targeting events, as measured by global editing patterns, increased throughout the cell cycle from G1 to M (Fig. 3i). This is consistent with previous observations that miRNA repression is weakest in the G1 phase[42], but the increase in editing could also represent accumulated miRNA targeting over the cell cycle that is then diluted by new transcription in the G1 phase[43,44]. While transcripts with longer half-lives tend to accumulate slightly more editing events than do shorter-lived transcripts (Supplementary Fig. 9), both groups display similarly increased editing throughout the cell cycle (Supplementary Fig. 10), lending some evidence to the former hypothesis that genuine increased miRNA repression is involved. We found that the agoTRIBE transcript is constant across the cell cycle while encoded protein decreases in G2/M, suggesting that our measurements of repression at this stage may in fact be underestimated (Supplementary Figs. 11 and 12). We found that known cell cycle-specific genes behaved as expected in our single-cell data (Fig. 3j–n). For instance, *CCNB1* has the highest expression in the G2/M stage, consistent with its role in promoting transition from G2 to the mitosis phase of the cell cycle (Fig. 3j). Similarly, *CDK2* is predominantly expressed in the S stage, consistent with its role in progression through the G1–S checkpoint (Fig. 3n). These specific cell cycle genes do not appear to be regulated by miRNAs, as evidenced by their low levels of editing across the entire cell cycle. In contrast, we found numerous genes differentially targeted by miRNAs across the cell cycle (Supplementary Table 8). For instance, the transcript of the centromeric protein CENPA appears to be strongly targeted by miRNAs during the G1 and S stages, but this targeting appears alleviated in G2/M where expression increased strongly, consistent with its role in mitosis (Fig. 3q). The *CDK6* gene has important roles in the G1–S transition and we found it to be specifically targeted in the S stage, where its expression might not be required (Fig. 3r). These examples serve as a proof of principle that agoTRIBE can detect miRNA targeting both in single cells and across distinct populations of single cells.

To test agoTRIBE on populations comprising distinct cell types we induced K562 cells to differentiate to erythroid precursor cells and applied Smart-seq3 to single cells (Supplementary Tables 9–11). It is well established that the cotranscribed miRNAs miR-144 and miR-451 are abundant and have important functions in erythrocytes[45] (Fig. 3v). We found that both undifferentiated and differentiated cells separated

well in a pseudotime trajectory, consistent with the linear and unidirectional nature of the differentiation (Fig. 3w). We found in total seven high-confidence targets of miR-144 that displayed increased targeting during this process. The protein TOR1AIP2 has functions in the endoplasmic reticulum and EIF2S3 is involved in translational initiation (Fig. 3x,y). Their repression during the differentiation process would be consistent with mature erythrocytes lacking an endoplasmic reticulum and having little or no translation. miRNA targeting appears to be heterogenous during this stage of erythroid differentiation, with cells either showing robust editing or no editing at all (Fig. 3x,y). We found that single cells that avoid targeting were still robustly detected in the sequencing (Fig. 3z), providing evidence that this heterogeneity in targeting is biological rather than technical in nature. In summary, we show that agoTRIBE can be applied to mixed-cell populations to give insights into heterogeneity in miRNA targeting.

Discussion

In summary, here we present a method for detection of miRNA–target interactions transcriptome-wide in single cells. In a comparison with current state-of-the-art Argonaute CLIP-seq methods, we found that agoTRIBE has several advantages (Table 1). First, agoTRIBE does not require the use of antibodies but rather simple transfection, thus reducing the cost and time of the required experimental procedures by several days. Second, our method uses either ordinary bulk or single-cell RNA-seq to detect editing events, meaning that transcriptome-wide measurements of RNA levels are also provided as part of the protocol. Indeed, agoTRIBE transfection is so straightforward that it could be applied to any given standard RNA-seq experiment to provide transcriptome-wide miRNA–target interaction information at little additional cost or effort. Third, while Argonaute CLIP-seq requires millions of cells, agoTRIBE can be applied to individual cells.

Conversely, our agoTRIBE method is currently limited in the resolution with which it detects individual target sites, making it more useful as a method for detection of miRNA target transcripts rather than specific binding sites. This is in many ways similar to the first CLIP-based methods, which also often identified binding fragments hundreds of nucleotides long; this limitation may be alleviated by improved fusion protein design.

AgoTRIBE will allow us to study the heterogeneity of miRNA targeting in homogenous cell populations, about which little is currently known. It will also allow us to study miRNA targeting in complex cell compositions in cell culture—for instance, in organoids or during induced cell differentiation. The agoTRIBE fusion protein could, with some effort, be placed under an inducible promoter in a living organism, which would allow the profiling of miRNA targeting in individual cell types of a complex tissue such as mouse brain. Because editing events would be detected with single-cell RNA-seq, individual cells could be sorted into cell types using computational approaches and would not require any physical sorting. In conclusion, we foresee that agoTRIBE has numerous applications that will benefit the wider community and that it will facilitate the entry of the miRNA field into the single-cell era.

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/s41587-023-01951-0.

References

1. Bartel, D. P. Metazoan microRNAs. *Cell* **173**, 20–51 (2018).
2. Valencia-Sanchez, M. A., Liu, J., Hannon, G. J. & Parker, R. Control of translation and mRNA degradation by miRNAs and siRNAs. *Genes Dev.* **20**, 515–524 (2006).

3. Chekulaeva, M. & Filipowicz, W. Mechanisms of miRNA-mediated post-transcriptional regulation in animal cells. *Curr. Opin. Cell Biol.* **21**, 452–460 (2009).

4. Jonas, S. & Izaurralde, E. Towards a molecular understanding of microRNA-mediated gene silencing. *Nat. Rev. Genet.* **16**, 421–433 (2015).

5. Gebert, L. F. R. & MacRae, I. J. Regulation of microRNA function in animals. *Nat. Rev. Mol. Cell Biol.* **20**, 21–37 (2019).

6. Bartel, D. P. & Chen, C. Z. Micromanagers of gene expression: the potentially widespread influence of metazoan microRNAs. *Nat. Rev. Genet.* **5**, 396–400 (2004).

7. Flynt, A. S. & Lai, E. C. Biological principles of microRNA-mediated regulation: shared themes amid diversity. *Nat. Rev. Genet.* **9**, 831–842 (2008).

8. Sun, K. & Lai, E. C. Adult-specific functions of animal microRNAs. *Nat. Rev. Genet.* **14**, 535–548 (2013).

9. Fromm, B. et al. MirGeneDB 2.1: toward a complete sampling of all major animal phyla. *Nucleic Acids Res.* **50**, D204–D210 (2022).

10. Hausser, J. & Zavolan, M. Identification and consequences of miRNA-target interactions–beyond repression of gene expression. *Nat. Rev. Genet.* **15**, 599–612 (2014).

11. Hafner, M. et al. CLIP and complementary methods. *Nat. Rev. Methods Primers* **1**, 19 (2021).

12. McMahon, A. C. et al. TRIBE: hijacking an RNA-editing enzyme to identify cell-specific targets of RNA-binding proteins. *Cell* **165**, 742–753 (2016).

13. Xu, W., Rahman, R. & Rosbash, M. Mechanistic implications of enhanced editing by a HyperTRIBE RNA-binding protein. *RNA* **24**, 173–182 (2018).

14. Rahman, R., Xu, W., Jin, H. & Rosbash, M. Identification of RNA-binding protein targets with HyperTRIBE. *Nat. Protoc.* **13**, 1829–1849 (2018).

15. Elkayam, E. et al. The structure of human argonaute-2 in complex with miR-20a. *Cell* **150**, 100–110 (2012).

16. Nakanishi, K., Weinberg, D. E., Bartel, D. P. & Patel, D. J. Structure of yeast Argonaute with guide RNA. *Nature* **486**, 368–374 (2012).

17. Schirle, N. T. & MacRae, I. J. The crystal structure of human Argonaute2. *Science* **336**, 1037–1040 (2012).

18. Baek, M. et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871–876 (2021).

19. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

20. Liu, J., Valencia-Sanchez, M. A., Hannon, G. J. & Parker, R. MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. *Nat. Cell Biol.* **7**, 719–723 (2005).

21. Pillai, R. S. et al. Inhibition of translational initiation by Let-7 MicroRNA in human cells. *Science* **309**, 1573–1576 (2005).

22. Pauley, K. M. et al. Formation of GW bodies is a consequence of microRNA genesis. *EMBO Rep.* **7**, 904–910 (2006).

23. Biasini, A. et al. Translation is required for miRNA-dependent decay of endogenous transcripts. *EMBO J.* **40**, e104569 (2021).

24. Hafner, M. et al. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* **141**, 129–141 (2010).

25. Kishore, S. et al. A quantitative analysis of CLIP methods for identifying binding sites of RNA-binding proteins. *Nat. Methods* **8**, 559–564 (2011).

26. Patel, R. K., West, J. D., Jiang, Y., Fogarty, E. A. & Grimson, A. Robust partitioning of microRNA targets from downstream regulatory changes. *Nucleic Acids Res.* **48**, 9724–9746 (2020).

27. Bartel, D. P. MicroRNAs: target recognition and regulatory functions. *Cell* **136**, 215–233 (2009).

28. Agarwal, V., Bell, G. W., Nam, J. W. & Bartel, D. P. Predicting effective microRNA target sites in mammalian mRNAs. *eLife* **4**, e05005 (2015).

29. Riolo, G., Cantara, S., Marzocchi, C. & Ricci, C. miRNA targets: from prediction tools to experimental validation. *Methods Protoc.* **4**, 1 (2020).

30. Hauptmann, J. et al. Biochemical isolation of Argonaute protein complexes by Ago-APP. *Proc. Natl Acad. Sci. USA* **112**, 11841–11845 (2015).

31. La Rocca, G. et al. Inducible and reversible inhibition of miRNA-mediated gene repression in vivo. *eLife* **10**, e70948 (2021).

32. Baek, D. et al. The impact of microRNAs on protein output. *Nature* **455**, 64–71 (2008).

33. Selbach, M. et al. Widespread changes in protein synthesis induced by microRNAs. *Nature* **455**, 58–63 (2008).

34. Hornstein, E. & Shomron, N. Canalization of development by microRNAs. *Nat. Genet.* **38**, S20–S24 (2006).

35. Schmiedel, J. M. et al. Gene expression. MicroRNA control of protein expression noise. *Science* **348**, 128–132 (2015).

36. Prole, D. L. & Taylor, C. W. A genetically encoded toolkit of functionalized nanobodies against fluorescent proteins for visualizing and manipulating intracellular signalling. *BMC Biol.* **17**, 41 (2019).

37. Fernandez-Rodriguez, J. & Marlovits, T. C. Induced heterodimerization and purification of two target proteins by a synthetic coiled-coil tag. *Protein Sci.* **21**, 511–519 (2012).

38. Meyer, K. D. DART-seq: an antibody-free method for global m(6)A detection. *Nat. Methods* **16**, 1275–1280 (2019).

39. Brannan, K. W. et al. Robust single-cell discovery of RNA targets of RNA-binding proteins and ribosomes. *Nat. Methods* **18**, 507–519 (2021).

40. Hagemann-Jensen, M. et al. Single-cell RNA counting at allele and isoform resolution using Smart-seq3. *Nat. Biotechnol.* **38**, 708–714 (2020).

41. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587 (2021).

42. Vasudevan, S., Tong, Y. & Steitz, J. A. Cell-cycle control of microRNA-mediated translation regulation. *Cell Cycle* **7**, 1545–1549 (2008).

43. Rodriques, S. G. et al. RNA timestamps identify the age of single molecules in RNA sequencing. *Nat. Biotechnol.* **39**, 320–325 (2021).

44. Hsiung, C. C. et al. A hyperactive transcriptional state marks genome reactivation at the mitosis-G1 transition. *Genes Dev.* **30**, 1423–1439 (2016).

45. Rasmussen, K. D. et al. The miR-144/451 locus is required for erythroid homeostasis. *J. Exp. Med.* **207**, 1351–1358 (2010).

Methods

Ago2-tagging DNA construct and plasmids

DNA-encoding ADAR2DD_E488Q (human ADAR2 adenosine deaminase domain 316–701-amino acid (aa) hyperactive mutant E488Q, the ADAR-only control) and ADAR2DD_E488Q in-frame with 55-aa flexible linker and Ago2 (agoTRIBE) were chemically synthetized and inserted in pcDNA3.1(+) vector (GeneArt, Invitrogen/Thermofisher). The synthetic coiled-coil (EIAALEK)3, E3- and (KIAALKE)3, K3-tags were described previously[37]. The E3-GGSG linker-Ago2 and K3-GGSG linker-ADAR2DD_E488Q constructs were synthetized and inserted in pcDNA3.1(+) vector (GeneArt, Invitrogen/Thermofisher). Plasmids encoding EGFP-Ago2 and FLAG/HA-TNRC6B protein were described previously[46,47]. eGFP_ADAR2DD_E488Q was synthetized and cloned in pcDNA3.1(+) vector (GeneArt, Invitrogen/Thermofisher). The plasmid encoding GFP nanobody, pGNb-mCherry[36] was used as pGNb-vector backbone. GNb-Ago2 and GNb-ADAR2DD_E488Q constructs were made by PCR from the corresponding complementary DNA, followed by insertion into the pGNb-vector lacking mCherry using Gibson assembly cloning (NEB). The following primers were used: Ago2-forward (GGGGGATCTGGATCCATGTACTCGGGAGCCGGCCCCGC);

Ago2-reverse (GCCCTCTAGACTCGAGTCAAGCAAAGTACATG GTGCGC);

ADAR2DD-forward (GGGGGATCTGGATCCATGCAGCTGCATT TACCGCAGG);

ADAR2DD–reverse (GCCCTCTAGACTCGAGTCAGGGCGTGAGT GAGAAC).

The DNA-encoding human T6B peptide (TRNC6B 599–683 aa[30]) and T6B fusion with GSG linker and mCherry were synthetized and cloned in pcDNA3.1(+) vector (GeneArt, Invitrogen/Thermofisher). The main constructs have been submitted to the Addgene repository with ID nos. 205598 (pcDNA3.1_ADAR2DD_E488Q), 205599 (pcDNA3.1_ ADAR2DD_E488Q_Ago2) and 205600 (pcDNA3.1_T6B-mCherry).

Cell culture and transfection

HEK-293T cells were grown in DMEM medium (Sigma) supplemented with 10% FBS (ThermoFisher) in the presence of penicillin/streptomycin (Sigma) under standard conditions at +37 °C in 5% CO_2. Before transfection, cells were seeded in 12-well plates (Corning) at 60–70% confluency. Cells were transfected with 0.5–2.0 µg of the recombinant plasmid DNA in two biological replicates 18–24 h after seeding using Lipofectamine 3000 (Invitrogen) according to the manufacturer's instructions.

Immunofluorescence staining and microscopy

HEK-293T cells were seeded on four-well glass slides (Nunc Lab-TekII) at 60–70% confluency and transfected with 0.5–1.0 µg of the recombinant plasmid DNA using Lipofectamine 3000 (Invitrogen). The 48-h posttransfected cells were fixed with 3.7% formaldehyde (Sigma), permeabilized with PBS Tween containing 0.1% Triton X-100 (Sigma) and blocked in PBS (Gibco) with 1% bovine serum albumin (Sigma) and 0.1% Tween 20 (VWR). The primary antibodies against Ago2 and ADAR2 were added for 1 h at room temperature. After washing, secondary antibodies were added for 1 h in the dark. Cells were mounted using Vectashield mounting medium (Vector Laboratories) with 0.1 µg ml^{-1} DAPI (ThermoScientific) for DNA counterstaining. Cells were imaged with a ZOE fluorescent cell imager (Bio-Rad). The following antibodies were used: rabbit polyclonal human ADAR2 (no. GTX54916, GeneTex), mouse monoclonal Ago2 (no. ab57113, abcam), mouse monoclonal anti-Flag (no. F3165, Sigma), goat anti-rabbit IgG (H + L) Alexa 594 (no. A11012, ThermoFisher) and donkey anti-mouse IgG (H + L) Alexa Fluor Plus 488 (no. A32766, ThermoFisher).

Bulk and single-cell RNA-seq

For standard bulk RNA-seq, cells were harvested 24 and 48 h after transfection. For bulk RNA library preparation, total RNA was isolated using the Quick-RNA miniprep kit (Zymo Research). Quality and integrity of RNA samples were assessed using a Bioanalyzer instrument and the RNA 6000 Nano kit (Agilent Technologies). One microgram of total RNA was used for RNA library preparation with the TruSeq Stranded mRNA kit (Illumina). Sequencing was carried out on an Illumina Next-Seq 500. For single-cell RNA library preparation using the Smart-seq3 protocol, cells were harvested 48 h after transfection, stained with 1 µg ml^{-1} propidium iodide (Invitrogen) and sorted into 384-well plates using a BD Bioscience FACS Aria fusion instrument. Single-cell Smart-seq3 libraries were prepared at the Eukaryotic Single-Cell Genomics facility and sequenced on an Illumina NovaSeq 6000 platform at the NGI facility (SciLifeLab) as previously described[40].

Bulk and single-cell RNA-seq data preprocessing and mapping

Both bulk and single-cell RNA-seq data were first processed for quality and adapter trimming using cutadapt v.4.0, allowing a minimum length of 36 base pairs (bp) after trimming (with parameters -q 30 -m 36), and then mapped to the hg38 reference genome using STAR v.2.7.2b (ref. [48]) with parameters detailed on GitHub (see link below). Putative PCR duplicates were removed using Picard using default options. The six samples used for estimation of transcriptome-wide editing levels (Supplementary Table 1) were subsampled to a sequencing depth of 18 million mapped reads using the Sambamba tool[49] with the following command line options: view -f bam -t 20–subsampling-seed=3. We used the featureCounts tool[50] for quantification of the expression level of genes based on Ensembl Hg38 genome annotation (https://www.ensembl.org/info/data/ftp/index.html, release v.99). Multimapping reads and/or reads overlapping by more than one feature (genes in our case) were accounted for using a fractional assignment (-M -O–fraction).

Bulk and single-cell RNA editing analysis

Alignment mismatches were called using a custom variant-calling pipeline (available on GitHub: https://github.com/vaishnoviS/agoTRIBE.git) with two main filtering steps: (1) Phred scale-based quality filtering and (2) read depth-based filtering. All mismatches supported by Phred score >30 and at least two reads supporting the nucleotide substitutions were reported. Single-nucleotide substitutions were then intersected with positions in the single-nucleotide polymorphism (SNP) database (NCBI dbSNP v.151)[51], and all known SNPs were excluded from further analyses. The remaining A-to-G and T-to-C (on reverse strand) substitutions were further compared with editing events compiled in the REDIportal database[52] and known editing sites were excluded. Strand-aware A-to-G substitutions were then reported as editing events. For the purpose of identification of agoTRIBE-specific targets, the number of editing events was computed for each annotated 3′ UTR in protein-coding genes. Nonsense-mediated decay, nonstop decay, processed transcripts, retained introns and unprocessed pseudogene isoforms were removed. Regression analysis was performed with the editing events in agoTRIBE-transfected samples compared with ADAR-only control samples to identify 3′ UTRs having higher editing. Standard residual values were calculated based on linear regression applied on editing events per gene using the lm function from the R stats package. The top 1,000 3′ UTRs with highest residual values were considered as highly edited 3′ UTRs by agoTRIBE. Conversely, the 1,000 3′ UTRs with lowest residual values were considered as background ADAR-only targets.

Small RNA-seq

Total RNA was isolated using the Quick-RNA miniprep kit (no. R1054, Zymo Research). The quality and integrity of RNA samples were assessed using a Bioanalyzer instrument and the RNA 6000 Nano kit (no. 5067-1511, Agilent Technologies). Small RNA libraries were prepared using the NextFlex small RNA library v.3 kit (no. 5132, Bio-Scientific, Perkin Elmer Applied Genomics). The following amounts

of total RNA were used as input: 0.25–1.00 μg of imatinib-treated K562 samples, 0.5 μg of agoTRIBE-transfected and untransfected HEK-293T control and 0.1 μg of cell cycle-specific populations (G0/G1, S, G2/M), and both agoTRIBE-transfected and untransfected HEK-293T control.

Small RNA-seq analysis
Bulk small RNA-seq data were analyzed using miRTrace to trim adapter sequences. Processed reads from miRTrace were then quantified using quantify.pl from the miRDeep2 package. To identify differentially expressed miRNAs in control and agoTRIBE-transfected HEK-293T samples (Supplementary Fig. 2), DESeq2 was used on raw counts for miRNA and all miRNAs with $P_{adj.} < 0.01$ were considered significant. For cell cycle analysis (Supplementary Fig. 13), reads per million (RPM) values obtained from quantify.pl were used to identify the top ten miRNAs in each stage and were plotted as a line plot.

Overlaps between miRNA targets reported by agoTRIBE, CLIP-seq methods and TargetScan
Those genes with 3′ UTRs that were specifically edited in agoTRIBE samples (above) were compared with the top target genes from publicly available HITS-CLIP[25], PAR-CLIP[24] and eCLIP[26] data. HITS-CLIP (Ago2 CLIP-Seq) and PAR-CLIP (AGO1234 PAR-CLIP) binding information was downloaded in BED format from the Dorina database[53] and eCLIP data (GSE140367, GSM4247216, empty vector control rep 1 (AGO-eCLIP-seq)) were downloaded from NCBI GEO[54]. Targeted genes from HITS-CLIP and PAR-CLIP data were sorted based on total peak coverage, while genes from the eCLIP target was sorted based on P values reported in the GEO supplementary file. For each set, the top 1,000 genes were considered for further comparison with our agoTRIBE top 1,000 target genes. For the purpose of identifying top TargetScan targets, the ten most highly expressed miRNAs in HEK-293T cells were identified (miR-10a-5p, miR-92a-3p, miR-10b-5p, miR-191-5p, miR-148a-3p, miR-16-5p, miR-378a-3p, miR-182-5p, miR-222-3p and miR-186-5p, in descending order by expression) using our miRNA-seq data (Supplementary Table 13). The top 100 targets for each of these miRNAs (based on the cumulative context ++ score) were downloaded from the TargetScan database[28]. Overall, 699 unique high-confidence target genes were compiled for the top ten most highly expressed miRNAs in HEK-293T cells (given that some targets were shared between miRNAs and were redundant).

Identification of target derepression following T6B transfection
Gene expression was estimated as described above. Unexpressed genes and those with <50 counts following normalization in any of the considered samples were removed. Between-sample normalization of expression was performed using the trimmed mean of the M-values method[55]. Fold change values between control and T6B-mCherry samples were calculated with the foldchange function from the gtools R package and visualized with their empirical cumulative distribution function. The top 699 TargetScan targets were defined as above, while the remaining expressed genes were defined as background.

Identification of positional information in target editing
To identify whether agoTRIBE editing reflects miRNA binding positional information, we compared agoTRIBE editing positions with those reported by AGO-eCLIP–seq data (GSM4247216). To do so we focused on genes with monoexonic 3′ UTRs and their largest isoforms that showed one eCLIP editing event. Genes with 3′ UTRs shorter than 40 bp were discarded. We then calculated the distance (in bp) between agoTRIBE and eCLIP editing events if detected in the same gene. As background control we defined a random iteration of uniform distributions ($n = 100$) to generate dummy editing positions within each targeted 3′ UTR. Background distances were then computed using eCLIP positional data and randomly distributed editing sites.

Single-cell RNA-seq cell filtering
Smart-seq3 data were preprocessed and mapped as described above. As a preliminary quality control step we filtered out cells with <9,000 detected genes or <150,000 deduplicated reads. To assess the efficiency of transfection we mapped our reads to the linker sequence with bowtie2 v.2.3.5.1 (ref. 56) using local alignment (-local) while allowing for one mismatch (−N 1), and quantified the aligned reads with the featureCounts v.2.0.0 tool[50]. For HEK-293T cells we then calculated normalized linker counts per cell as raw linker counts divided by the total number of reads, and plotted these against the editing events in 3′ UTRs. A pseudocount of 1 was added to all linker values to avoid division by zero. By visual inspection of the resulting plot we decided to exclude 163 cells with $\log_{10}$-normalized linker below −4.5 and below 1,000 editing events, leaving 566 cells in total. For K562 differentiated and undifferentiated cells, individual cells with a minimum of one linker count were considered positively transfected and used for further analysis. The K562 single-cell trajectory was constructed using transcriptomics counts with Monocle v.2.28.0 in R, with the top 1,000 genes differentially expressed between stages.

Cell cycle assignment for single cells
Cell cycle annotation was based on the Seurat v.4.0 pipeline[41]. Cells were count per million normalized using the NormalizeData function from Seurat, with the normalization method of relative counts (normalization.method = 'RC') and a scale factor of 10^6 (scale.factor = 1,000,000). We then annotated cells into cell cycle phases. To do this, we customized the CellCycleScoring function from Seurat to better fit Smart-seq3 data and used it with the default markers provided by Seurat. Specifically, cells with S-score and G2/M-score (as defined by the Seurat software) >50 were assigned to G1. These same marker genes were considered for scaling of data and calculation of the first 50 principal components with the Seurat functions ScaleData and RunPCA and default parameters. Finally, uniform manifold approximation and projection (UMAP) dimensionality reduction on the first ten dimensions was performed with the RunUMAP function. Based on this graphical representation, every cell cycle phase was split into 'early' and 'late' substages, with both substages containing approximately the same number of cells. For the purpose of identifying changes in gene expression and miRNA targeting across the cell cycle, sequence data from all cells belonging to a given cell cycle substage were pooled and mapped to Thermofisher ERCC92 spike-in sequences using bowtie, with strict parameters (-v 1) and a pseudobulk approach. The number of reads mapping to the spike-in was divided by the corresponding number of cells in the specific subphase to derive a scaling factor. This scaling factor was then used to normalize expression values of the pooled cells mapping to each cell cycle substage.

Transcript half-life analyses
The top 20% of stable and unstable transcripts was obtained from ref. 57 (GEO: GSE49831). For the violin plot (Supplementary Fig. 9), the per-transcript sum of expression and sum editing over all agoTRIBE-transfected cells (540) were calculated. For cell phase analysis (Supplementary Fig. 10), cells from each cell cycle phase were pooled and expression and editing were calculated. Normalization was performed as in Fig. 3i.

Erythroid differentiation from K562 cells
K562 cells ($1.0–1.3 \times 10^5$ ml^{-1}) were cultured in 12-well plates containing RPMI (no. 21875034, Gibco) with 10% fetal bovine serum and penicillin/streptomycin supplemented with 5 mM imatinib (no. 9084 S, Cell Signaling) for erythroid differentiation. Cells cultured in RPMI supplemented with 10% FBS, penicillin/streptomycin and 5 mM DMSO (no. 8418, Sigma) were used as an undifferentiated control. The Neon transfection system (no. MPK 1096, Invitrogen) and 1 μg of agoTRIBE and ADAR-only constructs were used for K562 electroporation. Cells

harvested 24 h after electroporation and corresponding DMSO (no. D8418, Sigma) controls were used for Smart-seq3 preparation.

Cell sorting by cell cycle stage

An asynchronous population of un- and transfected HEK-293T cells was used for cell cycle analysis and cell sampling for small RNA-seq. HEK-293T cells were transfected with 2.5 µg of agoTRIBE- and ADAR-only constructs using Lipofectamine 3000 and harvested 48 h after transfection. Following cell cycle analysis and cell sorting, cells were stained with Vybrant DyeCycle Green (no. V35004, Invitrogen) and propidium iodide. A BD Bioscience FACS Aria fusion cytometer was used for cell sorting in the G1, S and G2/M phases as determined by Vybrant DyeCycle Green staining.

Cell cycle profiling of agoTRIBE protein levels

HEK-293T cells were transfected with 2.5 µg of the agoTRIBE recombinant plasmid DNA using Lipofectamine 3000, then 48-h posttransfected cells were harvested. The suspended cells were washed with PBS (no. 10010-031, Gibco), fixed with 3.7% formaldehyde (no. 47608, Sigma), permeabilized with PBS Tween containing 0.1% Triton X-100 (no. T8787, Sigma) and blocked in PBS with 10% FBS, 0.36 µM IgG (no. 026202, Life Technologies) and 0.1% Tween 20 (no. 0777, VWR). The primary antibody against ADAR2 (no. GTX54916, GeneTex) was added for 30 min at room temperature and the secondary antibody, goat anti-rabbit IgG Alexa 594 (no. A11012, ThermoFisher), was added for 20 min in the dark. After washing with PBS, cells were resuspended in 400 µl of PBS and stained with Vybrant DyeCycle Green and 30 nM DAPI (no. 62248, ThermoScientific). A negative control with IgG and without ADAR2 primary antibody was used for Alexa 594 compensation, and the agoTRIBE Alexa 594-positive cell population was used for cell cycle analysis. G0/G1, S and G2/M phases were determined using Vybrant DyeCycle Green staining and a BD Bioscience FACS Aria fusion cytometer. agoTRIBE protein expression was determined as mean Alexa 594 fluorescence intensities normalized to G0/G1, S and G2/M event counts.

Statistical analyses

Differences in fold change derepression following T6B transfection in the top 1,000 edited genes by agoTRIBE and in miRNA targets according to HITS-CLIP, compared with background unedited and/or nontargeted genes, were tested with the DTS test, which calculates a reweighted integral of the distance between two empirical cumulative distributions, available in the dts_test function from the twosamples R package. Overlap significance between agoTRIBE targets, CLIP targets and TargetScan targets was estimated using binomial statistics. Differences in long linker reads and in editing events comparing control samples and samples following transfection with agoTRIBE were assessed with a nonparametric approach using the Wilcoxon rank-sum test.

AI protein structure predictions. The human Ago2 (Q9UKV8) three-dimensional protein structure was predicted by AlphaFold2 DB v.1 (ref. 19). K3/E3 peptides, eGFP and GNb and anti-GFP nanobody (LaG-16) structure predictions were generated using the Robetta, RoseTTAFold[18] modeling method available at https://robetta.bakerlab.org/.

Reporting summary

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

Data availability

Sequencing data have been deposited at SRA under accession no. PRJNA994505 (ref. 58).

Code availability

The source used in this study is available at GitHub: https://github.com/vaishnoviS/agoTRIBE (ref. 59).

References

46. Leung, A. K. L., Calabrese, J. M. & Sharp, P. A. Quantitative analysis of Argonaute protein reveals microRNA-dependent localization to stress granules. *Proc. Natl Acad. Sci. USA* **103**, 18125–18130 (2006).
47. Meister, G. et al. Identification of novel argonaute-associated proteins. *Curr. Biol.* **15**, 2149–2155 (2005).
48. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
49. Tarasov, A., Vilella, A. J., Cuppen, E., Nijman, I. J. & Prins, P. Sambamba: fast processing of NGS alignment formats. *Bioinformatics* **31**, 2032–2034 (2015).
50. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
51. Sherry, S. T., Ward, M. & Sirotkin, K. dbSNP-database for single nucleotide polymorphisms and other classes of minor genetic variation. *Genome Res.* **9**, 677–679 (1999).
52. Picardi, E., D'Erchia, A. M., Lo Giudice, C. & Pesole, G. REDIportal: a comprehensive database of A-to-I RNA editing events in humans. *Nucleic Acids Res.* **45**, D750–D757 (2017).
53. Blin, K. et al. DoRiNA 2.0-upgrading the doRiNA database of RNA interactions in post-transcriptional regulation. *Nucleic Acids Res.* **43**, D160–D167 (2015).
54. Edgar, R., Domrachev, M. & Lash, A. E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* **30**, 207–210 (2002).
55. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11**, R25 (2010).
56. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
57. Schueler, M. et al. Differential protein occupancy profiling of the mRNA transcriptome. *Genome Biol.* **15**, R15 (2014).
58. Sekar, V. et al. Detection of transcriptome-wide microRNA-target interactions in single cells with agoTRIBE. Datasets. *NCBI Sequence Read Archive (SRA)* https://www.ncbi.nlm.nih.gov/sra/PRJNA994505 (2023).
59. Sekar, V. et al. Detection of transcriptome-wide microRNA-target interactions in single cells with agoTRIBE. *GitHub* https://github.com/vaishnoviS/agoTRIBE (2023).

Acknowledgements

Author contributions

M.R.F. and I.B. conceived the study. I.B. designed and developed the agoTRIBE fusion and modular versions and performed protein structure prediction and all laboratory experiments except where specified below. V.S. performed all computational analyses, except where specified below. E.M.-S. and P.K. performed T6B-mCherry

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41587-023-01951-0.

Correspondence and requests for materials should be addressed to Inna Biryukova or Marc R. Friedländer.

Peer review information *Nature Biotechnology* thanks Alexander van Oudenaarden and Alistair Forrest for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

nature portfolio

Corresponding author(s): Marc R. Friedländer, Inna Biryukova

Last updated by author(s): 8/8/2023

Reporting Summary

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

Statistics

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

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

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

Software and code

Policy information about availability of computer code

Data collection	No software was used for data collection.
Data analysis	cutadapt v. 4.0 STAR v. 2.7.2b Sambamba tool v.0.7.1 featureCounts v2.0.0 bowtie2 v 2.3.5.153 Seurat v4.0 AlphaFold2 DB v1 Monocle v2.28.0 Various R packages as described in the methods section Exact command lines used are available at: https://github.com/vaishnoviS/agoTRIBE

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

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

Data availability
Sequencing data have been deposited at SRA under this accession number: PRJNA994505

Code availability
The source used in this study is available at GitHub: https://github.com/vaishnoviS/agoTRIBE

Public data and databases utilized in the study:
NCBI dbSNP v.151
REDIportal database
hg38 reference genome from NCBI
hg38 reference annotation from Ensembl database (https://www.ensembl.org/info/data/ftp/index.html, release version 99)
HITS-CLIP (Ago2 CLIP-Seq) and PAR-CLIP (AGO1234 PAR-CLIP) targets downloaded as BED format from Dorina database
eCLIP data (GSE140367, GSM4247216, Empty vector control rep 1 [AGO-eCLIP-seq]) was downloaded from NCBI GEO database
Predicted miRNA target information was retrieved from TargetScan Human database.

All of the intermediate data files with gene expression values, editing events etc are available as Supplementary Tables.

Human research participants

Reporting on sex and gender	No human research participants were included in this study.
Population characteristics	-
Recruitment	-
Ethics oversight	-

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	Sample size was chosen based on the single cell technology available. Smart-seq3 was run on two 384 well plates, covering control cells and agoTRIBE-transfected cells, for both HEK293T cells and K562 cells.
Data exclusions	Individual cells were excluded based on poor quality statistics, as described in the Methods section.
Replication	Reproducibility of agoTRIBE in single cells was tested by comparing gene expression and editing events in the distinct 384 well plates (Figure 3g-h).
Randomization	Samples were allocated into experimental groups based on the treatment, for instance control cells vs. agoTRIBE-transfected cells.
Blinding	We did not use blinding in our analyses, since the computational scientists were aware of the cell treatments before starting the analyses. Understanding the nature of the cell treatments is necessary for conducting the analyses.

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

Methods

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

Antibodies

Antibodies used	Rabbit polyclonal human ADAR2 (GeneTex, GTX54916), lot number 822003857 and 822102256. Mouse monoclonal Ago2 (Abcam, ab57113), clone 2E12-1C9, lot number gr3186789-1. Mouse monoclonal anti-Flag (Sigma, F3165), clone M2, lot number SLBL1237V. Mouse monoclonal anti-α-Tubulin (Sigma, T5168), clone B-5-1-2, lot number 115M4828V.
Validation	GTX54916 validation by manufacturer https://www.genetex.com/Product/Detail/ADAR2-antibody/GTX54916 and in our study, Sekar et al, Supplementary Figure 1b-d and Figure 1c. ab57113 validation by manufacturer https://www.abcam.com/products/primary-antibodies/argonaute-2-antibody-2e12-1c9-bsa-and-azide-free-ab57113.html ab57113 has been referenced in 150 publications https://www.abcam.com/products/primary-antibodies/argonaute-2-antibody-2e12-1c9-bsa-and-azide-free-ab57113.html F3165 validation by manufacturer http://www.sigmaaldrich.com/catalog/product/sigma/f3165?lang=en&region=SE, references are available from https://www.sigmaaldrich.com/SE/en/product/sigma/f3165. T5168 enhanced validation by manufacturer https://www.sigmaaldrich.com/SE/en/product/sigma/t5168, references are available from https://www.sigmaaldrich.com/SE/en/product/sigma/t5168

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	Standard HEK-293T cells were used, obtained from the D. Kanellis/J. Bartek lab. Standard K562 (ATCC® CCL-243TM) were used, obtained from G. Baselli (Vicent Pelechano lab)
Authentication	HEK-293T cell line was authenticated using cell line authentication test by Eurofins Genomics. The K562 cell line was not authenticated.
Mycoplasma contamination	Both HEK-293T and K562 cells were specifically tested for Mycoplasma contamination.
Commonly misidentified lines (See ICLAC register)	We have not worked with any commonly misidentified cell lines.