Non-destructive transcriptomics via vesicular export

Niklas Armbrust, Julian Geilenkeuser, Martin Grosshauser, Luisa Stroppel, Sebastian Schmidt, Tobias Panne, Emily Steinmaßl, Florenc Widenmeyer, Niklas Warsing, Oleksandr Berezin, Melike Sabry, Asina Sultanbai, Tobias Santl, Fabian Theis, Julien Gagneur, Dong-Jiunn Jeffery Truong[$], Gil Gregor Westmeyer[$]
[$]Co-correspondence

Transcriptomics is a valuable technique for multiplexed monitoring of cellular states. However, its application has been predominantly limited to assays that necessitate cell fixation or lysis, which hinders real-time, longitudinal tracing of RNA expression in live cells. Here, we present non-destructive transcriptomics by vesicular export (NTVE) to enable multi-time-point monitoring of RNA expression dynamics in living inducible reporter cell lines. Stabilized RNA reporter barcodes can be selectively packaged and exported from cells via virus-like particles (VLPs) containing selective affinity handles for convenient multichannel tracking of co-cultured cells. Using an engineered RNA adapter, NTVE exports cellular transcripts from inducible human and murine reporter cell lines with high concordance to conventional lysate-derived RNA-seq. NTVE allows monitoring of transcriptome changes in response to genetic and chemical perturbations in the same cells over time using standard sequencing workflows. NTVE can also be equipped with fusogens to deliver mRNA-encoded effectors or ribonucleoprotein gene editors from sender cells to activate gene reporters in co-cultured recipient cells. In addition, we demonstrate the utility of NTVE in monitoring hiPSC differentiation through daily NTVE analysis tracking multiple lineage-specific marker genes.

Transcriptomics, enabled by microarray and next-generation sequencing (NGS) technology, has become the key technology for characterizing cellular states. Its unparalleled ability to amplify and multiplex biological samples has unlocked profound insights into cellular processes and molecular interactions.

However, prevailing RNA sampling techniques are destructive and typically require cell fixation or lysis. Repeated sampling is currently only possible through laborious sequential aspiration using nanostraws or microneedles, which limits the number of robust cells and hinders scale-up [1,2].

Thus, creating non-consumptive methodologies that facilitate multiple temporal analyses of transcriptomic modifications within identical cellular populations is invaluable. This approach obviates the need to sacrifice cells, allowing them to be studied within their specific cellular network and environment.

Inspired by the information that can be gained from circulating extracellular vesicles (EVs), researchers overexpressed secreted microRNAs (Sec-miR) to obtain artificial secreted biomarkers for cancer cells [3]. Similarly, the TRACE-seq method employed a protein handle fused to an exosome-resident CD9 to export m6A-modified RNAs to identify correlations between endogenous transcripts and hydrogen peroxide, which acts as a cell stressor [4].

To develop a bio-orthogonal export mechanism with adjustable strength, we had previously constructed export vesicles derived from the HIV-1 gag polyprotein [5–7], to specifically export aptamer-tagged RNA barcode reporters via the PP7 coat protein (PCP) from cells [8]. Horns *et al.* developed a capsid-based export system termed COURIER to secrete engineered barcodes from cultured cells for monitoring cell population dynamics [9]. While the fusion of a nanocage with MS2 coat protein (MCP) resulted in a substantial increase in the export of target RNA-specific barcodes, the authors noted that constructs based on MMLV-Gag resulted in the non-specific export of cellular transcripts. However, this transcriptomic information has not been used for longitudinal monitoring of alterations in cellular states [9].

In this study, we present significant advances over our previous HIV1-Gag-based system [8] for exporting stabilized RNA barcodes and poly(A)-containing transcripts to enable multi-time-point transcriptomics from living cells. We have generated inducible cell lines for simultaneous NTVE analyses of co-cultures, non-invasive monitoring of transcriptomic responses to chemical perturbations, and non-invasive tracking of stem cell differentiation.

Results

HIV-Gag-based export machinery

We utilized HIV-Gag as the main chassis to implement membrane budding inducible by doxycycline (dox) combined with specific adapters for transcript packaging and export and membrane modifications to enable affinity-based enrichment (**Figure 1A-E**). We designed an expression cassette for stable genomic integration of NTVE systems via an ecotropic lentivirus (**Figure 1B**). A PGK promoter-driven TetON3G transactivator allows dox-dependent TRE3G-driven expression of the export vehicle. We inserted either a PP7 coat protein (PCP) or the first two poly(A)-binding domains protein PABPC1 (RRM1+RRM2) downstream of the nucleocapsid (NC) domain of codon-optimized Gag_{L21S} to modulate the export towards PP7-tagged or poly-adenylated RNA, respectively (**Figure 1C,E**).

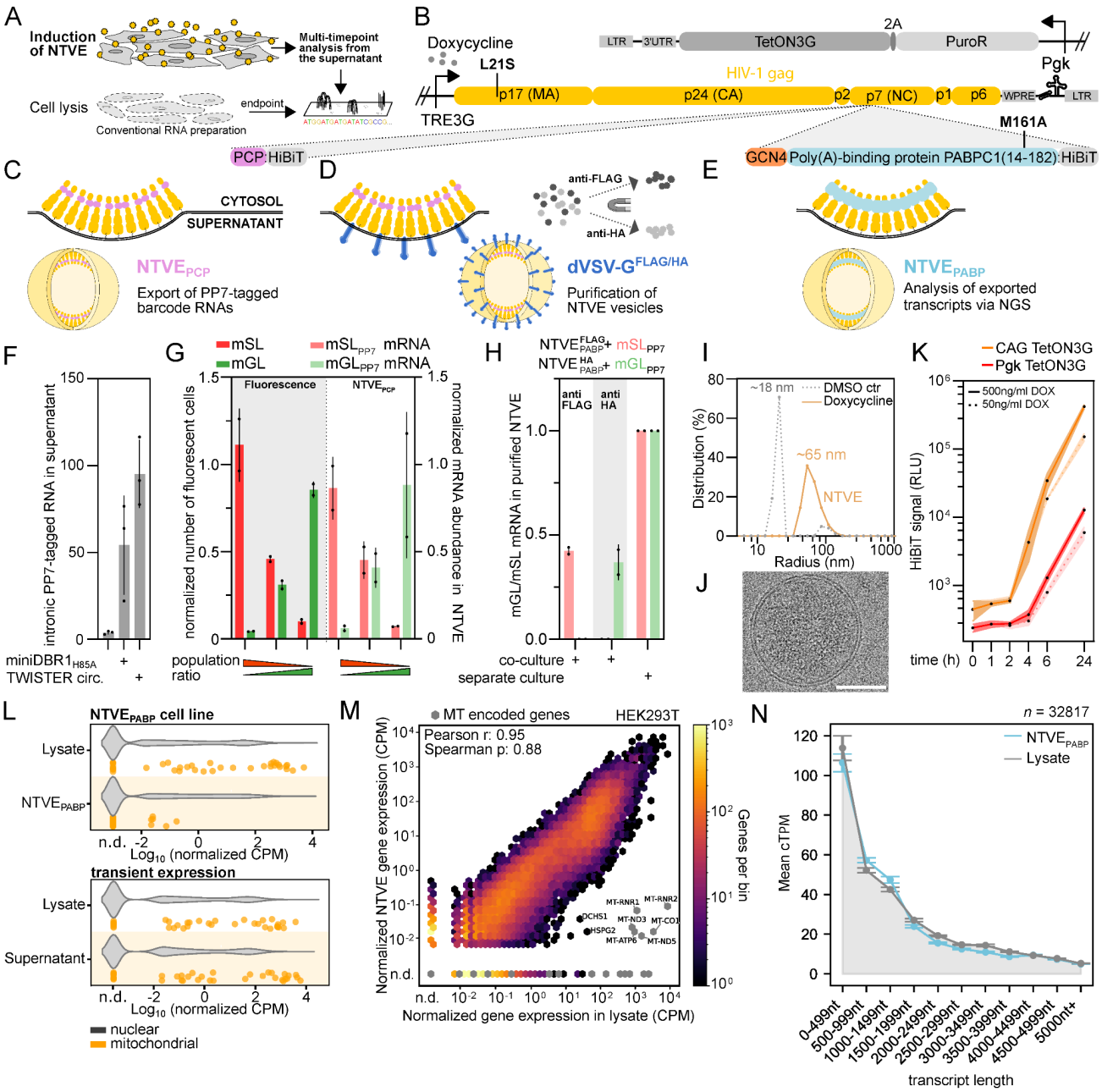

Figure 1. Development of Non-destructive Transcriptomics by Vesicular Export (NTVE)

(A) Inducible expression of cellular RNA-export machinery enables repeated sampling of vesicle-enclosed transcripts from the supernatant of the same cell population. In contrast, conventional transcriptomics requires cell lysis at each time point. **(B)** Schematic of the transfer vector for generating stable NTVE cell lines via ecotropic lentiviral delivery of a modular cassette with a PGK promoter driving the expression of the TetON3G transactivator for doxycycline-dependent expression of non-infectious Gag chimeras leading to membrane budding. The cassette is flanked by two Long terminal repeats (LTR) and stabilized by the Woodchuck Posttranscriptional Response Element (WPRE), a triple helix, and a tRNA. For selection, expression of the puromycin N-acetyltransferase is coupled via a 2A peptide. The NTVE budding module can be equipped with different adapters for packaging RNAs, and affinity handles for purification. A HiBiT tag allows for bioluminescence-based quantification of NTVE vesicles. (**C**) A PP7-coat protein (PCP) serves as a specific adapter to export PP7-tagged barcode RNAs. (**D**) NTVE vesicles can be decorated with purification tags mounted on membrane-anchored deactivated VSV-G (W72A) to enable the separation of distinct VLPs from different cell populations. (**E**) The first two RNA recognition motifs (RRM) from poly(A)-binding protein PABPC1 enable the export of polyadenylated mRNAs, while the GCN4 leucine-zipper domain mediates dimerization. **(F)** To enhance the accumulation of otherwise unstable intronic RNA, *e.g.,* PP7-tagged introns via $NTVE_{PCP}$, intronic RNA was stabilized by either co-expression of a catalytically inactive DBR1 or via the Tornado system, enabling NTVE export of introns. RNA levels were determined via RT-qPCR and normalized to a control condition lacking $NTVE_{PCP}$ expression. **(G)** HEK293T cells with integrated $NTVE_{PCP}$

expression cassette were transfected with plasmids encoding either mScarlet-I (mSL) or mGreenLantern (mGL) with PP7-tagged 3'UTRs. 24 h post-transfection, cells were reseeded in the ratios 10:1, 1:1, and 1:10. The different co-cultured cells were subsequently quantified by fluorescence microscopy (left *y*-axis). In addition, the mGL and mSL mRNA abundance in the supernatant was quantified via RTqPCR (right *x*-axis). Both measures were normalized to separate mono-cultures, exclusively expressing only one of the reporters. **(H)** For purification using the epitope-tagged dVSV-G (W72A), stable $NTVE_{PCP}$ HEK293T cells were transfected with plasmids encoding either mSL or mGL with PP7-tagged 3'UTRs. 24 h post-transfection, cells were reseeded in a ratio of 1:1 and subsequent NTVE induction (500 ng/ml doxycycline). The tagged VLPs of the co-cultured cell lines were pulled down with either anti-FLAG or anti-HA magnetic beads, and mGL/mSL abundance was analyzed via RT-qPCR. (**I**) Normalized distribution of the hydrodynamic radii of NTVE vesicles after doxycycline induction and enrichment by ultracentrifugation from the supernatant compared to a DMSO vehicle control. (**J**) Cryo-electron micrograph of $NTVE_{PABP}$ particles enriched by ultracentrifugation. Scalebar 50 nm. **(K)** Accumulation of NTVE vesicles in the supernatant after doxycycline induction (Continuous line: 500 ng/ml; dashed line: 50 ng/ml) monitored by bioluminescence signals from a HiBiT tag fused to PABPC1. Two stable cell lines that drive the TetON3g transactivator either via PGK or a CAG-promoter are compared. Points indicate the mean of three biological replicates with bounds for s.d.. (**L**) Distribution of expression levels for nuclear-encoded genes (box and violin plot) and mitochondrially-encoded genes (orange dots) from a stable cell line with inducible expression of $NTVE_{PABP}$ (top) and with transient overexpression of the same construct (bottom). Elevated mitochondrial species in the supernatant indicate potential cytotoxicity. **(M)** Normalized counts per million (CPM) identified in NTVE vesicles collected in the supernatant of HEK293T cells 72 hours after doxycycline induction plotted against those detected from the corresponding lysates. Genes encoded on the mitochondrial genome are labeled with their gene symbol and plotted in gray. Genes with less than 1% of the CPM in NTVE compared to the lysate are labeled in black. **(N)** Mean transcripts per million (TPM) of expressed coding transcripts (TPM >1) for bins of transcript lengths exported by $NTVE_{PABP}$ compared to that in the corresponding lysates. Data are plotted as the mean ±s.d. of triplicates.

Cellular export of population RNA barcodes

To enhance the sensitivity of our prior system for exporting PP7-tagged intronic RNA [8], we aimed to compact and stabilize the transcripts using two strategies: protection of intronic transcripts via a deactivated debranching enzyme (DBR1) or ribozyme-mediated circularization via the Tornado system [10].

While endogenous DBR1 debranches introns to initiate degradation, the truncated and inactivated $miniDBR1_{H85A}$ variant likely outcompetes the enzyme, leading to the stabilization and export of substantial amounts of PP7-tagged RNA barcodes (**Supplementary Figure 1)**. Alternatively, circularization of the PP7-tagged intronic RNA barcode using the Tornado system resulted in even more exported RNA barcodes (**Figure 1F, Supplementary Figure 1**).

As a further alternative, we evaluated whether we could effectively export mRNA transcripts tagged with a PP7 aptamer to track the relative abundance of two co-cultured cell populations via exported RNA reporters (**Figure 1G**). Therefore, cells were transfected in duplicate with either mGreenLantern (mGL) or mScarlet-I (mSL) containing a PP7 aptamer in the 3'UTR. Twenty-four hours after transfection, the two cell populations were reseeded with varying ratios (10:1, 1:1, 1:10). Fluorescence was quantified by microscopy two days after NTVE induction (**Figure 1G, left)**. RNA was extracted from the supernatant for subsequent RT-qPCR quantification of relative abundances of the respective transcripts (**Figure 1G, right**), which closely matched the stoichiometries obtained from microscopy. Next, we aimed to export an endogenous transcript of interest and thus inserted five PP7-loops in the 3'UTR of the housekeeping gene *POLR2A* via CRISPR/Cas9. After induction of the $NTVE_{PCP}$ system with doxycycline, we were able to detect a ~3-fold enrichment in comparison to a control cell line, where the *POLR2A* locus was not tagged with PP7 (**Supplementary Figure 2**).

Bio-orthogonal purification of NTVE vesicles

Subsequently, we attached distinct affinity tags to the NTVE particles to conveniently separate disparate cell populations before transcript analysis. Based on the successful use of the VSV-G glycoprotein in the cellular uptake of virus-like particles (VLPs) through LDL receptors, we thought that it could serve as an effective membrane adapter for the incorporation of epitope tags. However, to avoid any fusion between the vesicular stomatitis virus glycoprotein and host cell membranes, we strategically introduced the W72A mutation (dVSV-G) into the hydrophobic fusion loops [11].

We then fused HA ($NTVE^{HA}$) or FLAG epitope tags ($NTVE^{FLAG}$) to the N-terminus of dVSV-G and used magnetic anti-HA and anti-FLAG beads for pull-down. Combining the selective affinity handles with the specific barcode allowed for extracting the correct abundances of the co-cultured populations without carry-over (**Figure 1H**).

Poly(A)-binding adapter for export

Our next objective was to develop an RNA adapter that would enable the export of a broad spectrum of endogenous transcripts, providing more detailed information on the cellular state than selective barcodes. We wondered if we could enrich poly-adenylated RNAs (coding and non-coding) by replacing PCP with the first two RNA recognition motifs (RRMs) from the

ubiquitously expressed poly(A)-binding protein PABPC1 ($NTVE_{PABP}$). These two RRMs of the four in PABPC1 were shown to be sufficient for binding to poly(A) tails [12]. Furthermore, we introduced the mutation M161A to disrupt its binding of eIF4G and minimize the effects on protein translation [13,14].

After inducing $NTVE_{PABP}$ budding with dox and enriching vesicles from the supernatant by ultracentrifugation, we observed a peak hydrodynamic radius of the vesicles of 65 nm (spread of normalized distribution from ~30 to ~90 nm) by dynamic light scattering (DLS). In comparison, only very few scattering events were recorded from the ultracentrifuged supernatant taken from cells treated with DMSO vehicle control, yielding an apparent radius of <20 nm (normalized distributions in **Figure 1I**).

We then visualized the same samples by cryo-electron microscopy and found membrane-enclosed vesicles with luminal contrast enhancement and a median radius of 67 nm (*n*=156*)* (**Figure 1J, Supplementary Figure 3**).

Furthermore, we quantified $NTVE_{PABP}$ export following dox induction through HiBiT bioluminescent signals, comparing moderate (Pgk promoter) and strong expression (CAG promoter) of the TetON3G transactivator. The higher expression level of CAG-TetON3G resulted in a detectable bioluminescence signal ~4 hours after induction, while PGK-TetON3G required approximately six hours to be detected. (**Figure 1K, Supplementary Figure 4)**.

Subsequently, we conducted a first characterization of the exported transcripts and compared supernatants from the $NTVE_{PABP}$ cell line after dox induction against transient lipofection of the same genetic components using a standard RNA-seq analysis workflow (**Supplementary Figure 5**).

While transient overexpression resulted in substantial export of cytosolic transcripts, it also contained mitochondrial transcripts at comparable abundances, which should not be available for export via budding from the plasma membrane. In comparison, mitochondrial transcripts were strongly depleted in the supernatant from the $NTVE_{PABP}$ cell line (**Figure 1L**).

Therefore, we decided to work exclusively with cell lines with genomic integration of dox-inducible $NTVE_{PABP}$ to minimize potential membrane perturbations induced by lipofection.

$NTVE_{PABP}$ for export of endogenous transcripts

We proceeded with a more detailed analysis of the exported transcripts from the inducible $NTVE_{PABP}$ cell line and collected the supernatant 72 hours after dox induction before lysing the cells and subjecting both parallel samples in triplicate to Illumina stranded mRNA library preparation and next-generation sequencing (NGS).

We first assessed the variance in NTVE transcripts and found a coefficient of variation (CV) of 0.57, which was similar to that obtained from the corresponding endpoint transcriptome obtained from the lysate (CV = 0.60, **Supplementary Figure 6**).

We also investigated the effect of doxycycline on the differential expression in the reporter cells. Besides the expected overexpression of the Gag transcript, we only observed upregulation of a few metabolic genes known to be affected by the commonly used inducer doxycycline [15] (**Supplementary Figure 7)**.

Subsequently, we quantified the expression of genes detected with at least one count per million (CPM) from poly(A) enriched samples of three $NTVE_{PABP}$ replicates and found a high correlation to those from matching lysates (Pearson r = 0.95 (p is below numeric precision, pseudocount = 10^{-3} CPM) and Spearman ρ = 0.88, p = 0)) (**Figure 1M**).

In total, we detected 15957 genes in $NTVE_{PABP}$ that were also detected in the lysate. In comparison, only 64 genes were detected exclusively in the lysate, leading to a recovery rate of 99.6% of all expressed genes (**Supplementary Figure 8)**.

As expected, genes encoded by the mitochondrial DNA were depleted in the $NTVE_{PABP}$ sample, including all 13 protein-encoding mitochondrial genes, four mitochondrial tRNA genes, and one rRNA gene (**Figure 1M** and full list in **Supplementary Table 2**).

We then evaluated which transcript lengths were exported by $NTVE_{PABP}$ vesicles and thus quantified transcript abundances using Salmon[16] to plot the average transcripts per million (TPM) for different length bins (**Figure 1N**). The length distribution of transcripts in the NTVE fraction mirrored that of the lysate very closely (Pearson r = 0.993, p = 8.84×10^{-10}).

We then asked whether $NTVE_{PABP}$ showed a read depth bias along the transcripts compared to the lysate samples. We found that the read depth of $NTVE_{PABP}$ decreased as a function of the distance from the 3' terminus, while that of the lysate remained constant over 5000 nt, consistent with the intended binding of the PABPC1-RRMs to the 3' poly(A) tail of transcripts (**Supplementary Figure 9A , top**). In comparison, the number of transcripts at each nt position was not substantially different between the conditions (**Supplementary Figure 9A, bottom**). For comparison to Figure 1M, we also generate a scatter plot of TPM for $NTVE_{PABP}$ *vs.* lysate (**Supplementary Figure 9B**).

To further characterize the effect of the PABPC1-RRMs on the profile of transcripts exported by $NTVE_{PABP}$, we directly compared whole RNA preparations from the supernatant of the $NTVE_{PABP}$ cell line with that of a line expressing Gag without an RRM (Gag) and a reference lysate after nuclear depletion to filter for cytosolic transcripts.

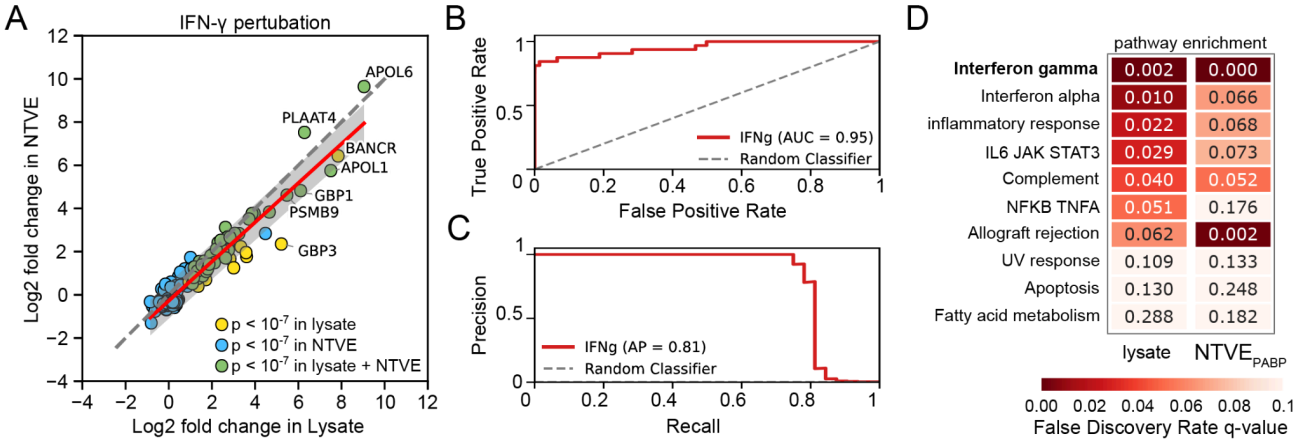

Figure 2. Detecting differential gene expression with NTVE

(A) Log2 fold change (log2FC) after IFN-γ perturbation obtained from NTVE against lysate. The supernatant was sampled 72 hours post perturbation. Colors indicate significant log2FC (p<1E-7) in the lysate (yellow), NTVE (blue), or both (green). Linear fit in red with gray 95% confidence intervals. The dotted line shows the perfect agreement reference. **(B)** Receiver operating characteristic (ROC) and **(C)** Precision-recall (PR) curves were employed to assess the performance of NTVE log2FCs as a classifier for identifying differentially expressed genes in response to IFN-γ stimulation. Genes identified in the lysate giving an adjusted p-value < 0.05 (Wald test in DESeq2 package) for absolute log2FC > 1 or < 1 were designated as ground truth positives or negatives, respectively. The log2FC from NTVE was used as the predictor variable for constructing the PR curve. **(D)** Comparison of positively enriched pathways based on the KEGG database after perturbation with HEK293T cells with 50 ng/ml interferon-gamma (IFN-γ). The pathways are ranked by increasing false discovery rate (FDR) q-values.

For $NTVE_{PABP}$, we again observed that the read depth decreased as a function of the base position relative to the 3' end of the Ensembl transcript annotation as compared to the transcripts in the lysate. In contrast, Gag without RRMs showed a more constant read depth across transcript lengths, which may indicate a less specific binding to RNA than for the PABP adapter (**Supplementary Figure 10A**).

When examining the length distribution of coding transcripts, we found that $NTVE_{PABP}$ reflected the profile in the reference lysates better than Gag-exported transcripts, which displayed a relative overrepresentation of lengths above 3000 nt (**Supplementary Figure 10B**). This difference was also observed when all 47455 Ensembl-derived transcripts were included to calculate the average TPM per length bin, resulting in a high average TPM in the smallest bin (**Supplementary Figure 10C**).

NTVE monitors genetic and chemical perturbations

Having successfully established the correspondence between cellular and NTVE transcripts under normal growth conditions, our next objective was to evaluate the efficacy of NTVE in tracking changes in transcript expression in response to external stimuli.

Initially, we induced the low-expressing gene *ASCL1* via CRISPRa, to evaluate the sensitivity and specificity of NTVE. After $NTVE_{PABP}$ induction, we cultured the cells for 72 hours and collected the supernatant for subsequent analysis of NTVE transcripts via NGS. NTVE correctly detected the only significantly upregulated gene, *ASCL1* (**Supplementary Figure 11**).

We then proceeded to investigate a chemical stimulation by interferon-gamma (interferon-γ), known to activate several essential signaling pathways in mammalian cells. We found a high agreement (Pearson r = 0.96 in log-scale, $p < 10^{-127}$) between the fold changes in the NTVE and lysate samples (**Figure 2A, Supplementary Figures 12**) and confirmed that mitochondrial transcripts were not enriched in the supernatant, after perturbation (**Supplementary Figure 13).**

We then asked how well NTVE can classify significantly up- or downregulated transcripts in the lysate. To this end, we constructed receiver operating characteristic (ROC) and precision-recall (PR) curves by taking the log2 fold changes (log2FC) from NTVE as predictors for the binary classification of all lysate-derived transcripts with absolute log2FC > 1 or <1 and adjusted p-value < 0.05 (Wald test in DESeq2 package) (**Figure 2B,C**).

A pathway analysis of the differentially expressed genes showed that NTVE correctly identified the interferon-γ signaling network (**Figure 2D**).

NTVE analysis in co-cultured cells and genetic inter-cell communication

Next, we were interested in using NTVE in co-cultures of cells. Therefore, we co-expressed the FLAG or HA affinity tags we used before (**Figure 1DH**) on $NTVE_{PABP}$ vesicles and co-cultured human HEK293T cells and murine Neuro-2a cells with the respective handle (**Figure 3A**). We sampled the supernatant after dox induction and purified it with either anti-FLAG or anti-HA magnetic beads. Subsequent NGS analysis revealed a predominant representation of human or mouse

transcriptomes, demonstrating selective transcriptome enrichment of the target cell population (**Figure 3B**).

Based on these successful experiments, we next set out to test whether we could pseudotype NTVE with glycoproteins that can be taken up selectively by receiver cells expressing cognate receptors as the basis for a cell-cell communication system. We, therefore, co-expressed the MLV-Env glycoprotein (Eco) in $NTVE_{PCP}$-expressing HEK293T cells and co-cultured them with cells expressing a murinizedSLC7A1 selective for Eco[17]. We further modified sender cells to package PP7-tagged mRNA coding for the large serine recombinase pa01[18] and equipped the recipient cell with a matching attB/attP fluorescent switch that expresses mGL upon attB/attP recombination (**Figure 3C**). Flow cytometry analysis after dox induction revealed that a substantial number of receiver cells translated the pa01 mRNA, mostly via NTVE uptake (bottom right quadrant) but occasionally also via direct fusion with a sender cell, indicated by simultaneous expression of the sender cell marker mTagBFP2 (upper right quadrant), also visible in confocal fluorescence microscopy (**Figure 3D, Supplementary Figure 14B**).

In addition, we tested whether it is also possible to transmit information by packaging a CRISPR effector as ribonucleoprotein (*i.e.* complexed with respective guide RNA). We chose the improved prime editor (iPE-C[19] and a pegRNA to target a reporter locus of recipient cells, resulting in a blue-to-green switch in fluorescence (**Figure 3E**). For this system, we also replaced the sender cell marker with the far-red fluorescent protein miRFP670nano3 and the MLV-Env glycoprotein with VSV-G to minimize direct membrane fusion of neighboring cells.

Remarkably, we observed a substantial number of editing events in the recipient cells by flow cytometry, suggesting that cell-cell communication can also be engineered by exchanging modular CRISPR effectors delivered as ribonucleoproteins, not just via mRNA.

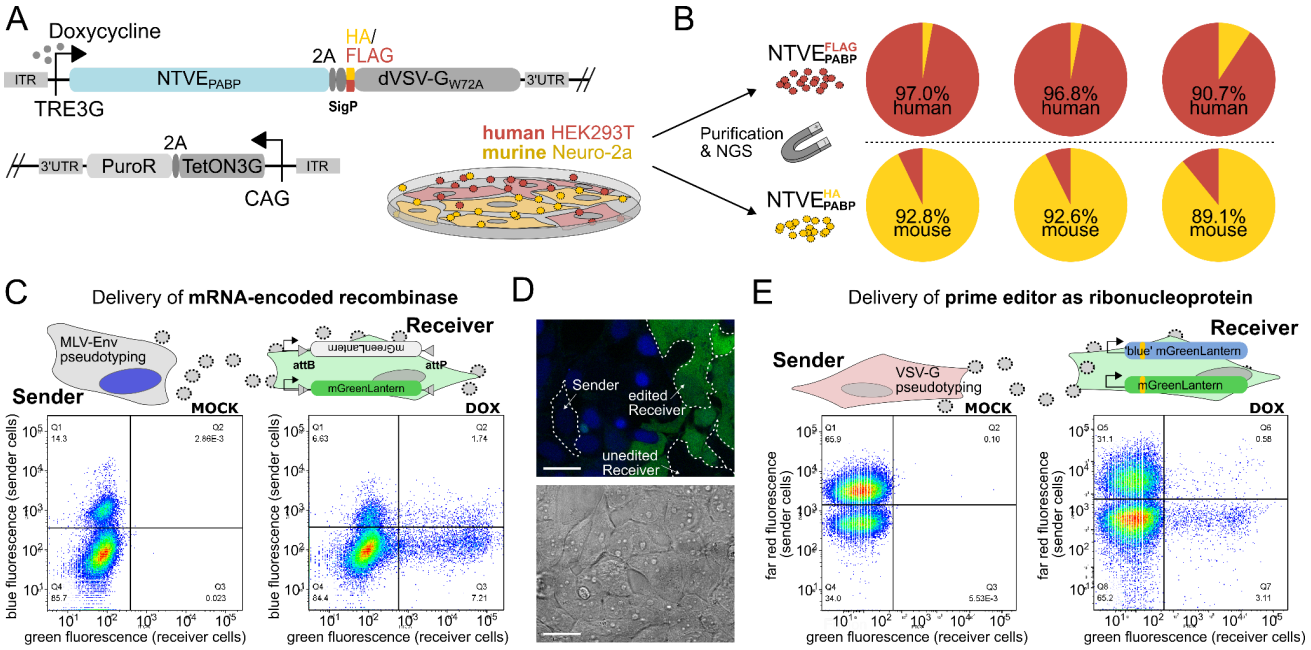

Figure 3. Selective NTVE reporting and cell-to-cell communication in co-cultured cells

(**A**) FLAG- or HA-tags mounted on non-fusogenic dVSV-G(W72A) coupled to TRE3G-driven $NTVE_{PABP}$ expression via 2A skipping peptide. The murine Neuro-2a and the human HEK293T cell lines were co-cultured, and mRNA secretion was induced by adding 500 ng/ml doxycycline. (**B**) NTVE vesicles were purified by the respective anti-Flag or anti-HA magnetic beads, followed by RNA preparation and NGS. Reads were mapped on human and mouse cDNA to quantify the proportion of human and murine transcripts. (**C,D**) Engineered cell-cell communication via vesicular mRNA export from a sender cell encoding either a recombinase (C) or a prime editor ribonucleoprotein (D) and uptake into HEK293T receiver cells. The vesicles are pseudotyped via glycoproteins, for which the receiver cells express the respective receptors. The receiver cell undergoes a genetic edit of a fluorescent reporter locus. (**C**) Vesicular export of a PP7-tagged mRNA coding for the recombinase $pa01_{NLS}$ that activates an mGreenLantern reporter system in receiver cells by site-specific inversion. Vesicles present MLV-Env(ΔC16) glycoproteins for selective entry into receiver cells via a murinized SLC7A1 acting as a receptor. FACS scatter plots show the co-culture population of sender and receiver cells 3 days post (mock) induction of NTVE expression with doxycycline. (**D**) Co-culture from (C) captured with confocal microscopy after 72 hours, scale bar 50µm. (**E**) Delivery of a prime editor as a ribonucleoprotein (RNP) complex, facilitating a shift from blue to green fluorescence upon editing of a reporter locus. In contrast to (C), sender cells co-express the VSV-G glycoprotein to facilitate uptake on the recipient cells via LDL receptors. FACS scatter plots show the co-culture population of sender and receiver cells 3 days post (mock) induction of NTVE expression with doxycycline.

NTVE monitors hIPSC-cell differentiation

Next, we generated a monoclonal hiPSC line via PiggyBac recombinase-mediated insertion of the doxycycline-inducible $NTVE_{PABP}$ cassette (**Figure 4A,B**). We then analyzed Yamanaka factor *OCT4* and *NANOG* in NTVE samples by RT-qPCR which were both expressed at constant levels 24 hours and 48 hours post-induction. The lysate showed a higher apparent abundance of transcripts relative to *ACTB*, but a direct comparison with the NTVE fraction may be confounded by differences in transcript length (**Figure 4C**). $NTVE_{PABP}$ export was confirmed via HiBiT signal to remain strong over this period (**Figure 4D**). Viability and cytotoxicity of hiPSCs were analyzed 48 hours post-induction with varying doxycycline concentrations, showing no substantial alterations (**Figure 4E**).

Subsequently, we differentiated the cells into the three germ layers (ectoderm, endoderm, and mesoderm) to investigate the upregulation of lineage-specific markers[20]. We added lineage-specific culture medium supplemented with doxycycline, freshly prepared each day (4-6 days) (**Figure 4F**). During this time, NTVE was monitored via HiBiT bioluminescence, showing a decreasing intensity of export activity during differentiation towards ectoderm and mesoderm, while it increased for cells differentiated towards mesoderm (**Supplementary Figure 15**). We confirmed that the successful differentiation process can be monitored in the NTVE fraction (daily), consistent with the downregulation of Yamanaka factors also observed in the endpoint measurement from lysates (**Supplementary Figure 17**) and the upregulation of lineage-specific marker genes (**Figure 4G-I**). For all three lineages, the lineage-specific markers were barely detectable via RT-qPCR in the iPSC state, but a strong upregulation of the marker genes could be observed already one day post-differentiation (**Figure 4G-I**).

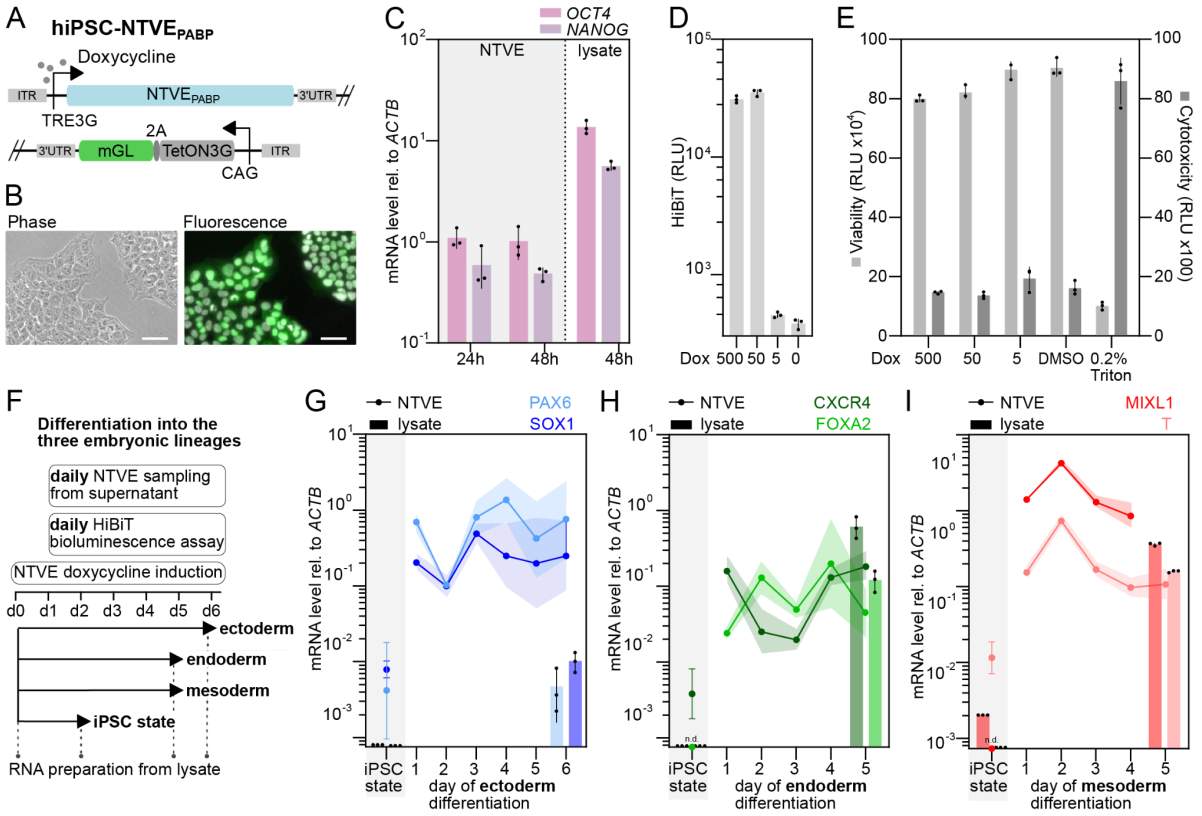

Figure 4. Monitoring gene expression during hiPSC differentiation.

(**A**) Schematic of the PiggyBac insertion cassette for human induced pluripotent stem cells (hiPSC) encoding a CAG-promoter-driven TetON3G to mediate doxycycline-dependent expression of $NTVE_{PABP}$ via a TRE3G promoter. Expression of TetON3G is coupled to that of mGreenLantern via a 2A peptide. **(B)** Phase and fluorescence microscopy of hiPSC:$NTVE_{PABP}$**.** Scale bar for microscopy images: 50µm. **(C)** mRNA levels of the Yamanaka factors *OCT4* and *NANOG* in NTVE 24 and 48 hours post-induction with 500 ng/mL doxycycline compared to the lysate at the end-point. Bars represent the mean of three biological replicates, with error bars indicating s.d. **(D)** Quantification of $NTVE_{PABP}$ budding efficiency at given doxycycline levels by HiBiT measurement from the supernatant 48 hours post-induction. **(E)** Viability and cytotoxicity assay of hiPSC at given doxycycline concentrations 48 hours post-induction. **(F)** Timeline of the trilineage differentiation. NTVE-RNA was prepped daily while NTVE budding was maintained with 500 ng/mL doxycycline during differentiation to endo- and mesoderm (5 media exchanges) and ectoderm (6 media exchanges). RNA from lysates was prepped in the hiPSC state and as endpoint measurement from the respective germ layer. **(G-I)** mRNA levels of lineage-specific markers normalized to *ACTB* in the supernatant over the time course of iPSC differentiation towards ectoderm (**G**), endoderm (**H**), or mesoderm (**I**). The last time point of *MIXL1* was excluded due to an abnormal melting curve. In F-H, the points represent the mean of three biological replicates, with the shading indicating the s.d.

Discussion

We have demonstrated that non-destructive transcriptomic analyses can be performed on live cells over multiple time points to assess responses to perturbations and monitor stem cell differentiation while preserving the cells for subsequent assays.

Since we found that transient transfection may lead to increased membrane permeability and unspecific RNA leakage (**Figure 1L**), we generated stable mammalian cell lines, including hiPSCs, in which the vesicular export can be regulated via dox induction to sample cellular transcripts without affecting cell viability and with high correspondence between the NTVE-exported and lysate-derived transcripts (**Figure 1M, Supplementary Figure 9**).

Gag with RRM domains from PABPC1 installed on the interior of the protein shell achieved the highest read coverage close to the 3' end, consistent with preferential binding to the poly(A) tail of cellular transcripts (**Supplementary Figures 9A,10A)**. This characteristic likely contributes to the high degree of concordance between the length distribution of exported transcripts and the actual length distribution of transcripts inside the cell [21].

In contrast, the transcriptome exported by Gag with its native RNA binding domain (NC) demonstrated a preference towards longer transcripts, likely a consequence of its more uniform binding probability across the transcript bases (**Supplementary Figure 10**).

To monitor NTVE export strength, we installed a HiBiT bioluminescence reporter, which allowed us to conveniently estimate the optimal sampling rate for sufficient read depth. We found that a ~100-fold increase in HiBiT signal over the uninduced condition typically allowed us to identify a sufficiently large set of >10,000 genes (**Supplementary Figure 4c**).

Based on the spike-in of known amounts of a unique transcript into the RNA preparations prior to NGS, we estimate the RNA export rate to be ~760 transcripts per day per HEK293T cell. Higher sensitivity can be achieved using dedicated single-cell RNA sequencing protocols such as the SMARTseqv4 protocol, which also allows for the inclusion of unique molecular identifiers (UMIs), enabling greater sequencing depth from the same sample.

Further engineering of the export machinery, e.g. by generating larger vesicles, may allow preferential sampling of larger transcripts if desired. RNA binding domains other than the PABP and PCP domains used here may allow preferential sampling of other specific subsets of transcripts. In conjunction with the FLAG and HA tags used in this study, additional membrane-exposed affinity handles will facilitate the differentiation of a greater number of cell types.

Microfluidics may then allow continuous sampling of NTVE vesicles during nutrient replenishment in 3D cell cultures, enabling online transcriptomic analysis to optimize differentiation protocols in a closed-loop system and dynamically adjust NTVE induction via HiBiT monitoring.

Regarding new analytical strategies enabled by NTVE, it will be of great interest to train predictive models from learned correlations between NTVE-lysate and NTVE-NTVE sample pairs to reconstruct past transcriptomic trajectories from end-point lysate-derived sequencing data. Similarly, it should be possible to predict future transcriptomic states, e.g. after a perturbation, from end-point lysate-derived transcriptome data based on learned NTVE trajectories.

Furthermore, NTVE may be used to determine which transcriptomic states provide protection against catastrophic perturbations. This could be achieved by comparing the NTVE transcripts before adding a toxin or infectious agent with the NTVE and lysate-derived transcripts of surviving cells.

In addition to obtaining information from cells non-destructively, we have shown that information encoded by mRNA and a CRISPR effector packaged as ribonucleoproteins can be exchanged between cells. This capability opens up future possibilities to collect analogous multi-timepoint omics data of proteins and metabolites and to exchange these biomolecules between cells.

To summarize, NTVE stands as an accessible and versatile technology that empowers live-cell transcriptome-wide profiling of dynamic cellular processes in various biomedical contexts.

Limitations of the study

NTVE cannot currently export nuclear-localized transcripts such as many of the non-coding RNAs or small nucleolar RNAs, but RNA reporters exported via $NTVE_{PCP}$ could provide information about the regulation of the respective promoters (**Supplementary Figure 2)**.

Although we chose to drive expression of the Tet-ON 3G transactivator from the CAG promoter, which provides strong expression levels in hIPSC cells, a sufficient induction strength of NTVE should be confirmed by HiBiT luminescence assay for a specific differentiation protocol of interest (**Supplementary Figure 15**.

While it is possible to export RNA barcodes that can report the relative abundance of co-cultured cell populations, it is not currently feasible to assign a specific exported transcript to a single cell.

Methods

Molecular cloning

PCR

Single-stranded primer deoxyribonucleotides (Integrated DNA Technologies (IDT)) were solubilized at 100 µM in nuclease-free water. PCR reaction with plasmid and genomic DNA templates was performed with Platinum SuperFi II PCR Master Mix (Thermo Fisher Scientific) according to the manufacturer's protocol. PCR reactions were purified by DNA agarose gel electrophoresis and subsequent DNA extraction using Monarch DNA Gel Extraction Kit (New England Biolabs (NEB)).

DNA digestion with restriction endonucleases

Samples were digested with NEB restriction enzymes according to the manufacturer's protocol in a total volume of 40 µl with 1–3 µg of plasmid DNA. Afterward, DNA fragments were purified by agarose gel electrophoresis and subsequent purification using Monarch DNA Gel Extraction Kit (NEB).

Ligation and Gibson assembly

Concentrations of agarose-gel purified DNA fragments were determined by a spectrophotometer (NanoDrop 1000, Thermo Fisher Scientific). Ligations were carried out with 50–100 ng backbone-DNA (DNA fragment containing the ori) in 20 µl volume, with molar 1:1–3 backbone:insert ratios, using T4 DNA ligase (Quick Ligation Kit, NEB) at room temperature for 5–10 minutes. Gibson assemblies were performed with 75 ng backbone DNA in a 15 µl reaction volume and molar 1:1–5 backbone:insert ratios, using NEBuilder HiFi DNA Assembly Master Mix (2x) (NEB) for 20-60 min at 50 °C.

DNA agarose gel-electrophoresis

1% (m/m) agarose (Agarose Standard, Carl Roth) gels were prepared in 1x TAE buffer and 1:10.000 SYBR Safe stain (Thermo Fisher Scientific). Gel electrophoresis was carried out for 20–40 minutes at 100 V. For size determination, 1 kb Plus DNA Ladder (NEB) was used. DNA samples were mixed prior to loading with Gel Loading Dye (Purple, 6x) (NEB).

Bacterial strains (E. coli) for molecular cloning

Chemically competent Stable cells (NEB Stable) were used to transform circular plasmid DNA. For plasmid amplification, carbenicillin (Carl Roth) was used as a selection agent at a final concentration of 100 µg/ml. All bacterial cells were incubated in Lysogeny Broth (LB) medium or on LB agar plates, including the respective antibiotics.

Bacterial transformation with plasmid DNA

Transformation was performed by mixing 1–5 µl of ligation or Gibson reaction with 50 µl thawed, chemically competent cells, which were incubated on ice for 30 minutes before heat shocking at 42 °C for 30 s. Afterward, cells were incubated on ice for 5 minutes and mixed with 450 µl SOC-medium (NEB). Transformed cells were then plated on agar plates containing the respective antibiotics according to the supplier's information. Plates were incubated overnight at 37 °C.

Plasmid DNA purification and Sanger sequencing

E. coli colonies with correct potential constructs were inoculated from agar plates in 2 ml LB medium at 37°C with the respective antibiotics and incubated for at least 6 hours or overnight. Plasmid DNA was extracted with Monarch Plasmid Miniprep Kit (NEB) according to the manufacturer's protocol and sent for Sanger sequencing (GENEWIZ, Azenta Life Sciences). Sanger sequencing validated clones were inoculated in 100 ml LB medium overnight at 37 °C containing the respective antibiotic selection agent. Plasmid DNA was extracted using the Plasmid Maxi Kit (QIAGEN).

Cultivation of mammalian Cell lines

Cells were cultured at 37 °C, 5% CO_2, and H_2O saturated atmosphere.

HEK293T (ECACC: 12022001, Sigma-Aldrich), **Neuro-2a** (ECACC 89121404) and **NIH/3T3** (ECACC 93061524) cells were maintained in advanced DMEM (Gibco, Thermo Fisher Scientific) supplemented with 10% FBS (Gibco, Thermo Fisher Scientific), GlutaMAX (Gibco, Thermo Fisher Scientific), 100 µg/ml penicillin-streptomycin (Gibco, Thermo Fisher Scientific), 10 µg/ml piperacillin (Sigma-Aldrich) and 10 µg/ml ciprofloxacin (Sigma-Aldrich). Cells were passaged at 90% confluence by removing the medium, washing with DPBS (Gibco, Thermo Fisher Scientific), and incubating with Accutase solution (Gibco, Thermo Fisher Scientific). Detachment of the cells was observed after 5–10 minutes. Accutase was subsequently inactivated by adding pre-warmed DMEM, including 10% FBS and all supplements. Cells were then transferred into a new flask at an appropriate density or counted and plated in 96-well, 48-well, or 6-well format for plasmid transfection or doxycycline induction.

Human induced pluripotent stem cells (hiPSCs (WTC11,UCSFi001-A (RRID:CVCL_Y803)) were incubated in Essential 8 (E8) Flex medium (Thermo Fisher Scientific, A2858501) at 37 °C under 5% CO2 saturation on Geltrex-coated (A1413201, Thermo Fisher Scientific) plates. At 70% confluency, the cells were subcultivated by incubating the cells for 5 minutes at room temperature with StemMACS Passaging Solution XF (Miltenyi Biotec). Subsequently, StemMACS Passaging Solution XF was

aspirated, and hiPSC colonies were harvested in E8 flex medium and chopped using a 1000 µl pipette tip. Collected cells were seeded on 96-, 48- or 6-well plates at the appropriate cell densities.

Genetic constructs and generation of stable NTVE cell lines

The NTVE knock-in plasmid is based on random integration via ecotropic lentiviral delivery or PiggyBac transposase activity. For transposition, a NTVE donor plasmid was co-transfected with a plasmid encoding for a hyperactive mutant of the PiggyBac transposase[22,23]. For ecotropic pseudotyping of the lentiviral delivery, the ecotropic MLV glycoprotein (Addgene #35617) was used in combination with psPAX2 (Addgene #12260) and a transfer vector encoding for NTVE. In both cases, the puromycin selection cassette is used to select for multi-copy insertion (4 µg/ml for HEK293T, NIH/3T3 and Neuro-2a; 1 µg/ml for hiPSC). All core plasmids will be available from Addgene.

Ecotropic lentiviral delivery

HEK293T cells were seeded on a 6-well plate in 3 ml advanced DMEM with a concentration of 250,000 cells/ml. The cells were transfected one day post seeding with a total of 1.2 µg plasmid DNA with the following stoichiometry: psPAX2 packaging plasmid (Addgene #12260, 400 ng), pseudotyping plasmid pCMV-Eco (adapted from Addgene #12260); 200 ng) and the transfer plasmids encoding psi-tagged, LTR-flanked NTVE (600 ng). The lipofection was carried out according to the manufacturer's protocol (jetOPTIMUS, Polyplus). 72 hours post transfection, the supernatant (SN) was collected, sterile filtered (0.45 µm), and concentrated (~20x) using a 100 kDa cut-off filter. For transduction, the HEK293T WT cells were seeded on a 48-well plate (75,000 cells in 500 µl advanced DMEM). One day post seeding, cells were treated with the ecotropic booster according to protocol (Takara Bio, 631471) before the supernatant of one 6-well production was applied to three wells. Cells were expanded to a T25 flask three days post ecotropic transduction and selected for multi-copy insertion with 4 µg/ml puromycin.

Random integration via PiggyBac transposase

HEK293T or Neuro-2a cells were seeded with a density of 200,000 cells/ml in 3 ml advanced DMEM for transfection (jetOPTIMUS, Polyplus) of 1.2 µg plasmid DNA with the following stoichiometry: 400 ng of the plasmid encoding a hyperactive mutant of the PiggyBac transposase[22,23] and 800 ng of the NTVE knock-in plasmid. Cells were expanded to a T75 flask three days post ecotropic transduction and selected for multi-copy insertion with 4 µg/ml puromycin.

For nucleofection, 1,000,000 NIH/3T3 cells were centrifuged (200× g, 3 minutes) and resuspended in 100 µl of the SG Cell Line nucleofection Kit (Lonza, V4XC-3024). After the addition of 500 ng plasmid DNA encoding the hyperactive PiggyBac transposase and 1000 ng of the NTVE knock-in plasmid, cells were transferred into a cuvette for subsequent nucleofection with the 4D-Nucleofector (Lonza, program EN-158). After electroporation, cells were resuspended in 5 ml advanced DMEM and transferred into a T25 flask. Three days post nucleofection, cells were treated with 4 µg/ml Puromycin to select for cells with NTVE knock-in.

Generation of a monoclonal, stable hiPSC NTVE cell line

The hiPSC NTVE cell line was generated by multi-copy random integrating via PiggyBac transposase. For seeding, hiPSCs were singularized using Accutase (Sigma, A6964-500ML) for 5 minutes at 37 °C and counted using a Neubauer improved cell counting chamber (Carl Roth, PK361). The cells were seeded in E8 flex medium supplemented with 10 µM Y27632 (Enzo Life Sciences, ALX-270333-M005) on a Geltrex-coated 48-well plate (40,000 cells in 250 µl). 24 hours post seeding, cells were supplemented with 300 µl fresh media without Y27632 and subsequently transfected with 750 ng plasmid DNA using the Lipofectamine Stem Reagent (Thermo Fisher Scientific, STEM00001) according to protocol. Cells were allowed to recover to ~70 % confluency in E8 flex medium. NTVE hiPSCs were clonalized and picked; green fluorescent colonies were expanded.

Generation of stable HEK293T receiver cell lines for delivery of an mRNA-encoded recombinase

The polyclonal recombinase reporter cell line was designed by integrating an inverted mGL codon sequence flanked by inverted attB/P sites downstream of a CAG promoter onto a PiggyBac transposon. A puromycin resistance marker was also placed under the control of the same promoter via an IRES downstream of the reporter (see **Supplementary Figure 14** for a detailed schematic). HEK293T cells were co-transfected with a plasmid containing the entire transposon and a corresponding PiggyBac transposase to stably integrate the transposon into the HEK293T genome (600,000 cells, 1.2 µg plasmid DNA, jetOPTIMUS, polyplus), and selected with 1 µg/ml puromycin for at least 14 days.

Generation of stable HEK293T receiver cell lines for delivery of a prime editor RNP

To monitor prime editing events in the receiver cell, a reporter encoding for the blue-shifted mGL mutant (G65S, Y66H) with a CAG promoter was cloned into flanking homology arms for a safe-harbor locus (see **Supplementary Figure 14** for a detailed schematic). To generate the stable HEK293T reporter cell line, cells were transfected with the donor plasmid as well as a second plasmid encoding Cas9 and the corresponding gRNA according to the manufacturer's protocol (600,000 cells, 1.2 µg DNA, jetOPTIMUS, Polyplus) 24 hours post-seeding on a 6-well plate (600,000 cells in 3 ml per well) in the presence of 0.5 µM of the DNA-PKcs inhibitor AZD7648 (HY-111783; MedChemExpress), and CAG-promoter-driven i53, a 53BP1 inhibitor. The polyclonal population was monoclonalized using limited dilution in 96-well plates, and clones were genotyped to determine homozygosity. To generate susceptibility for NTVE particles pseudotyped with MLV-Env, this cell line was further modified in the endogenous SLC7A1 locus via prime editing to express an engineered variant of the SLC7A1 protein with a murinized binding site (238-244(KEGKPGV > NVKYGE)). The resulting polyclonal cell line was monoclonalized by limited dilution, and clones were tested for ecotropic mGL delivery via integrase-deficient lentivirus (IDLV). Green fluorescent clones were expanded. The pegRNA sequence can be found in Supplementary Table 1.

Generation of stable HEK293T sender cell lines for delivery of mRNA or prime editor RNP

The polyclonal sender cell lines were designed by combining doxycycline-inducible expression of NTVE components (gag-PCP, glycoprotein) with constitutive expression of TetON-3G transactivator, a selection and fluorescence marker of choice, and either the PP7-tagged mRNA or gene-editor-RNP cargo onto a PiggyBac transposon (see **Supplementary Figure 14** for a detailed schematic). Sender cell lines were then generated by transfecting and selecting HEK293T cells with 1 µg/ml puromycin.

Co-cultivation of sender and receiver cells

Sender and receiver cells were seeded in a 1:1 ratio onto poly-D-lysine-coated 2-well slides (Ibidi) in 1.2 ml medium at a combined density of 300,000 cells/ml and then co-cultivated for 72 hours either with or without doxycycline before being analyzed with confocal microscopy and FACS.

Perturbation experiments

ASCL1 transactivation

For stable upregulation of ASCL1, 600,000 HEK293T:NTVE$_{PABP}$ cells were seeded on a 6-well plate and transfected with 1.2 µg plasmid-DNA encoding an ITR-flanked, EF1a-driven dCas9(D10A;H840A):VPR[24] and four U6-driven sgRNAs targeting the transcription start site[25] (800 ng) and the hyperactive PiggyBac transposase (400 ng).

-886 nt GTTGTGAGCCGTCCTGTAGG
-572 nt GCCAATTTCTAGGGTCACCG
-451 nt GAGAACTTGAAGCAAAGCGC
-196 nt GGGGAGAAAGGAACGGGAGG

IFN-γ perturbation

HEK293T:NTVE$_{PABP}$ cells were seeded on pre-coated (poly-D-lysine) 6-well plates at a density of 750,000 cells in 3 ml advanced DMEM supplemented with 50 ng/ml interferon gamma. One day post seeding, the NTVE system was induced with 500 ng/ml doxycycline and cultured for three days before the RNA was prepped from the supernatant and cell lysates.

Trilineage differentiation of hiPSC

Differentiation of hiPSC:NTVE$_{PABP}$ into ecto-, meso-, and endoderm was performed on 48-well plates (500 µl/well feed volume) using the STEMdiff Trilineage Differentiation Kit (StemCell Technologies) according to the manufacturer's instructions. For ectoderm and endoderm differentiation, 200,000 cells were seeded, while 50,000 cells were seeded for differentiation into mesoderm.

Cells were seeded as single cells in E8 flex media supplemented with 10 µM Y27632. One day post seeding, cells were washed using DPBS, and subsequently, pre-warmed differentiation media was added to the cells. This procedure was repeated each day; the supernatants were collected, and the RNA from the NTVE fraction was purified for analysis via RT-qPCR.

The differentiation media was supplemented with 500 ng/ml doxycycline throughout the entire course of differentiation. As a control condition, 200,000 cells were cultured in E8 flex media to maintain the iPSC state. Cells were sacrificed for RNA isolation at the end of the differentiation.

All supernatants were analyzed for NTVE export strength via HiBiT luminescence assay.

Viability and toxicity assays

To assess viability and cytotoxicity, NTVE hiPSC were seeded as single cells (25,000 cells) on a Geltrex-coated 96 well plate in 100 µl E8 flex media supplemented with 10 µM Y27632 (Enzo Life Sciences, ALX-270333-M005). One day post-seeding,

cells were washed with DPBS, and the NTVE system was induced with different doxycycline concentrations. The CellTiter-Glo 2.0 Cell Viability Assay (Promega, G9241), as well as LDH-Glo Cytotoxicity Assay (Promega, J2380), was performed in triplicates according to the manufacturer's protocol 48 hours post-induction.

Affinity purification

40 µL Anti-FLAG® M2 magnetic beads (Merck, M8823) were equilibrated with 200 µL of supplemented DMEM for pull-down of NTVE vesicles obtained from a 6-well of HEK293T cells. After one washing step with 400 µL PBS, the sterile filtered supernatants were added to the beads. Samples were incubated at room temperature with slow rotation for 2 hours. After three washing steps (500 µL PBS), NTVE vesicles were eluted using 100 µL of a 4x Flag peptide solution (Merck, F3290).

For HA-mediated purification, 35 µL Pierce Anti-HA Magnetic Beads (Thermo Fisher Scientific, 88836) were used. The purification of NTVE vesicles via HA-mediated pull-down was similar to the anti-FLAG purification. Finally, the vesicles were eluted from the beads using 100 µL HA peptide (1 mg/mL).

If cells were co-cultivated expressing either dVSV-G(W72A):HA or dVSV-G(W72A):FLAG, the supernatants were first incubated with anti-FLAG beads, followed by anti-HA beads.

The flow-through of the anti-FLAG purification was applied on anti-HA beads.

All elution fractions were analyzed for secretion strength via HiBiT luminescence assay.

Characterization of NTVE particles

Dynamic light scattering (DLS) analysis

Dynamic Light Scattering (DLS) measurements were performed using a DynaPro NanoStar DLS system (Wyatt Technology). The supernatant from $NTVE_{PABP}$ cells was collected 48 hours post dox induction or addition of DMSO vehicle control. Samples were sterile filtered (0.45 µm) to remove any large particulates that could interfere with the measurement and concentrated via ultracentrifugation (120,000× g, 2 hours). Pellets were resuspended in PBS (1/100 of the original volume). For each measurement, 2 µL of the sample was loaded into a clean, dust-free quartz cuvette (WTC JC-325). The cuvette was carefully inspected for any scratches or imperfections before use. The sample was allowed to equilibrate to 25°C prior to data acquisition. Between measurements, pure PBS samples were run to verify that there was no remaining signal after washing the cuvette.

DLS measurements were performed at a scattering angle of 90° using a 662 nm laser. The laser power was automatically adjusted by the instrument to maintain an optimal count rate between 1 x 10^5 and 1 x 10^6 counts per second. For each sample, 10 acquisitions were recorded with 5 seconds per acquisition, resulting in a total measurement time of approximately 1 minute per sample.

The autocorrelation functions were analyzed using the DYNAMICS software (version 8.2.0.297, Wyatt Technology) to obtain the size distribution. A regularization algorithm was employed to fit the autocorrelation data, assuming a spherical model for the particles. The intensity-weighted size distribution was converted to a volume-weighted distribution.

Cryo-electron microscopy

Freshly prepared suspensions (3 µL) of the identical samples also subjected to DLS analysis (see above) were pipetted onto glow-discharged, perforated, carbon-coated 200 mesh copper grids (Quantifoil Micro Tools), incubated for 4 sec, and blotted. Immediately thereafter, the grids were snap-frozen in liquid ethane using a Vitrobot Mark 4 (Thermo Scientific) and loaded into the autoloader of a 200 kV Glacios cryo-TEM (Thermo Scientific). Low-magnification overview images were collected manually using a Falcon4i direct electron detector camera (Thermo Scientific). To increase the signal-to-noise ratio, a median filter (despeckle operation) was applied twice to each overview image before further processing. Ten images from the dox-induced sample were randomly selected for manual vesicle segmentation in Fiji (ImageJ)[26]. Segmented areas were used to estimate the radii of equivalent circular areas. For display, image areas containing the selected vesicles were cropped and contrast-enhanced using the Contrast Limited Adaptive Histogram Equalization (CLAHE) function in Fiji image software (block size 127, histogram bins 256, maximum slope 1.5)[27] and assembled into a montage.

HiBiT bioluminescence quantification

For quantification of NTVE particles in culture supernatants, samples were centrifuged at 200× g for 5 minutes and then analyzed with the Nano-Glo HiBiT Lytic Detection System (Promega) according to the manufacturer's protocol. Luminescence was measured on the Centro LB 960 plate reader (Berthold Technologies) with 0.5 s acquisition time.

Fluorescence microscopy

Epifluorescence microscopy images were taken on an EVOS FL Auto Imaging System (Invitrogen, Thermo Fisher Scientific) with identical settings across all

samples. Confocal microscopy was conducted on a Leica SP5 system (Leica Microsystems) at 37 °C.

Flow cytometry

Cells were gently detached in a suitable volume of Accutase, pelleted (200× g, 5 minutes), and resuspended in ice-cold 0.4% formalin for 10 minutes. Fixed cells were pelleted again (200× g, 5 minutes) and resuspended in ice-cold 200 µl DPBS. Flow cytometry was performed on the BD FACSaria II system (controlled with the BD FACSDiva Software (v.6.1.3, BD Biosciences)). Briefly, the main population of the cells was gated according to their FSC-A and SSC-A. Secondly, single cells were gated using FSC-A and FSC-W. The final gate (green fluorescence) was used to determine the proportion of edited reporter cells after successful mRNA/PE RNP delivery in co-culture.

Cell fractionation

For the separation of the cytoplasmic and nuclear RNA, cells were fractionated (Fig. 1M, reference). After removal of the cultivation media, cells were rinsed with DPBS (Gibco; Thermo Fisher Scientific) and incubated in Accutase solution (Gibco; Thermo Fisher Scientific) for 10 minutes. The detached cells were collected and centrifuged for 3 minutes at 300× g and 4 °C. Cells were resuspended in 750 µl DPBS (Gibco; Thermo Fisher Scientific) and centrifuged again (300× g at 4 °C). Cell pellets were resuspended in 750 µl ice-cold mild lysis buffer (50 mM Tris, 100 mM NaCl, 5 mM MgCl2 and 0.5% Nonidet P-40, pH 8) and incubated for 5 minutes to allow lysis of the cell membrane only. The lysate was centrifuged for 2 minutes (18,000× g at 4 °C), and the supernatant was carefully transferred to the fresh Eppendorf reaction tube. The pelleted nuclei were resuspended in 750 µl mild lysis buffer. Subsequently, 1.5 µl proteinase K (20 mg ml−1; Thermo Fisher Scientific) and 4 µl 40% sodium dodecyl sulfate solution were added to both fractions, which were vortexed vigorously before incubation for 15 minutes at 37 °C. Finally, the fractions were centrifuged at 18,000× g for 2 minutes to clear impurities from the samples.

RNA isolation from cell lysates

The Monarch Total RNA Miniprep Kit (NEB) was used for RNA purification from cell lysates. For experiments with high sample numbers, the Quick-RNA 96 Kit (Zymo Research) was used according to protocol. RNA was stored at −80 °C.

RNA isolation from NTVE vesicles

To isolate RNA from NTVE vesicles, supernatants were sterile-filtered to remove the remaining dead cells. Subsequently, the supernatant was concentrated 20x prior to lysis and purification using the Quick-RNA Viral 96 Kit according to protocol. RNA was stored at −80 °C.

RT-qPCR

RNA samples were subjected to DNAse digestion with DNAse I (NEB) with the following protocol changes: the DNAse concentration was doubled, and the incubation time was extended to 15 minutes, followed by inactivation at 75 °C for 10 minutes. The Luna Universal One-Step RT-qPCR Kit (NEB) was used according to the manufacturer's protocol. Reactions were run in 384-well plates (11 µl per well) in technical duplicates for 45 cycles. Control amplification reactions without reverse transcriptase were carried out as a reference for each RT-qPCR run. The reactions were performed and monitored in an Applied Biosystems QuantStudio 12K Flex Real-Time PCR system. Primer sequences are provided in **Supplementary Table 1.**

Library preparation and Illumina sequencing

Library construction was performed as described in the high-throughput protocol of the TruSeq stranded mRNA Sample Prep Guide (Illumina) in an automated manner using the Bravo Automated Liquid Handling Platform (Agilent). Samples were either selected for mRNA sequencing (poly(A)-enriched via poly(T) beads), or total RNA was used in combination with rRNA depletion (**Supplementary Fig. 10**).

RNA libraries were reverse transcribed, and the resulting cDNA was PCR amplified after tagmentation. Fragments were multiplexed and sequenced as 100 bp paired-end run on an Illumina HiSeq4000 or HiSeq X platform with a depth of ≥ 30 Million reads per sample.

Quantification of NTVE-exported RNA

A poly-adenylated transcript encoding mGL was *in-vitro* transcribed using the HiScribe T7 mRNA Kit with CleanCap Reagent AG (NEB, E2080S). The transcript contains a genetically encoded 67 nt poly(A) tail. After quantifying the mRNA amount via Qubit RNA High Sensitivity, 1 pg of mGL transcript was spiked into the sample after the total RNA extraction and before the NGS library preparation. To obtain the absolute RNA amount of a sample, the proportion of fragments mapping to mGL was used for a known amount of mGL spike-in (1pg).

The amount of exported transcripts per day was calculated from a spike-in experiment three days post NTVE induction. We assumed that cells divided two times during cultivation and that the average length of an mRNA is ~1500 nt[28,29].

RNA-seq analysis pipeline

Given the low amount of RNA input in our experiments, duplicate reads were removed using czid-dedup (v0.1.0) in the paired-end mode to correct for a possible excess of PCR duplicates, followed by different data analysis workflows schematized in the flow chart in **Supplementary Figure 5**.

STAR-based analysis

To establish correspondence between NTVE and lysate, we obtained counts at the level of genes and aligned transcripts to a reference (STAR, v2.7.10b_alpha_220111)[30]. RNA-seq paired-end reads were aligned using a combined human-murine index built with assembly of GRCh38 with annotations from Ensembl release 108 for *Homo sapiens* and GRCm39 with annotations from Ensembl release 109 for *Mus musculus*. To account for reads mapping to the mRNA encoding $NTVE_{PABP}$ and the mGL-encoding mRNA used as a reference of a known amount, the corresponding sequences were added to the index[31]. STAR was run with the --outSAMtype BAM SortedByCoordinate parameter on the paired-end reads. Uniquely mapped fragments were used as inputs to calculate counts with featureCounts (v2.0.3) [32] on the human-murine index, using the following parameters to account for paired-end, reverse stranded data: "featureCounts -a [combined_index.gtf] -o [output_file] -g gene_id -t exon -s 2 -p". The resulting counts were normalized to counts per million (CPM).

Salmon-based analysis

To assess which transcript lengths were exported by $NTVE_{PABP}$, we obtained transcript-level RNA-seq counts by Salmon (v1.10.1)[16]. The index was constructed from *Homo sapiens* cDNA and ncRNA sequences (GRCh38, Ensembl release v108), *Mus musculus* (GRCm39, Ensembl release v109)*, and Bos taurus* cDNA and ncRNA sequences (ARS-UCDI1.2, Ensembl release v108), to identify possible contaminations from bovine serum contained in the cell culture media. The sequences for the mRNA encoding $NTVE_{PABP}$ and the mGL-encoding mRNA were added to the index. The resulting counts were normalized to transcripts per million (TPM) to account for differences in sequencing depth across samples. Human transcripts were identified and isolated by filtering for ENST (Ensembl transcript) identifiers. For each gene, we calculated the relative abundance of all expressed isoforms for each sample. We defined a dominant isoform of a gene in a sample as a transcript that accounted for ≥90% of the total gene expression (sum of all isoform TPM values) of the gene in the specific sample. For each gene with a dominant isoform, we recorded the transcript identifier, the proportion of total gene expression represented by the dominant isoform, and its normalized abundance in TPM. To plot the mean TPM for bins of transcript lengths, we stratified transcripts by the length of the corresponding annotation and calculated the mean of the TPM of all transcripts per bin per sample. We then calculated the mean and the standard deviation across all samples of one condition. For analyses where only coding genes selected, we renormalized the TPMs obtained from Salmon again to 1 million after subselection.

Differential expression analysis

Differential expression estimation was performed using the DESeq2 package (v1.34.0,[33]) in R (v4.1.2). P-values were adjusted for multiple testing using the Benjamini-Hochberg procedure to control the false discovery rate (FDR). Log2 fold changes were shrunk using the 'apeglm' method to reduce the effect size estimates of low-count genes. Analyses were performed separately for lysate and supernatant samples, comparing dox-treated samples against DMSO (non-induced), CRISPRa *ASCL1* transactivated, and IFNγ stimulated samples. For the IFNγ stimulated condition, six dox-treated replicates were compared against three replicates after IFNγ stimulation for both lysate and NTVE. For CRISPRa *ASCL1* transactivated samples, the DE calculation was conducted in triplicates versus reference triplicates for both NTVE and lysate. Effects of DMSO were assessed in the lysate.

Read-coverage along transcripts

We filtered for dominant isoforms as outlined above to then compute the read coverage with a pipeline implemented in Python (v3.12.3) utilizing the pandas data analysis library for efficient data manipulation and processing. For the calculation of the coverage depth, gene annotations were obtained from the Ensembl annotation GTF file for *Homo sapiens* (v108) containing genomic coordinates and feature information[34]. The pipeline filtered the GTF annotations to retain only exon features corresponding to the identified dominant isoforms.

We extracted and compiled the genomic coordinates, including chromosome, start position, end position, and strand information. The coordinates were processed to create a BED format file containing all exonic regions of the dominant transcripts.

Samtools (v1.19.2) was used with the depth command to calculate the read depth for the corresponding exonic features using "samtools depth -b dominant_isoform_genes_coordinates.bed -@ 8 -o

[output_file] [input_bam]". Coverage depths were normalized to account for differences in sequencing depth and transcript abundance. The normalization process involved scaling each transcript's coverage by its total read count to obtain position-wise proportions, multiplying by the transcript length to maintain relative depth information. For visualization, transcripts were aligned by the 3' end of the annotation, and coverage was analyzed across the last 5,000 nt.

Binary classifier performance evaluation

We evaluated the performance of $NTVE_{PABP}$ as a binary classifier by establishing ground truth labels from lysate samples and assessing prediction accuracy using supernatant samples from the same cells 96 hours after doxycycline induction and 72 hours after IFNγ stimulation (or vehicle control). Ground truth labels were determined using DESeq2 hypothesis testing framework on lysate samples, comparing six replicates without stimulation and three with IFN-γ stimulation. Genes were classified into two sets: a positive set comprising genes with significant absolute log2 fold changes greater than 1 (FDR < 0.05, using the 'greaterAbs' alternative hypothesis option) and a negative set containing genes with significant absolute log2 fold changes less than 1 (FDR < 0.05, with the 'lessAbs' alternative hypothesis option). This classification yielded 39 positive and 16,095 negative gene IDs. Log2 fold changes were shrunk using the 'ashr' method to mitigate effect size overestimation in low-count genes, and p-values were adjusted for multiple testing using the Benjamini-Hochberg procedure. For prediction, we analyzed matched supernatant samples using DESeq2's greaterAbs test to identify genes with absolute log2 fold changes greater than 1 between the same treatment conditions (six doxycycline replicates *vs.* three IFNγ replicates). The adjusted p-values from this analysis served as prediction scores. Classifier performance was evaluated using Receiver Operating Characteristic (ROC) and Precision-Recall (PR) curves, with corresponding Area Under the Curve (AUC) values calculated using the sklearn.metrics package in Python.

Data availability

All sequences of the genetic constructs will be available via addgene. Source data will be provided and NGS data will be available at Gene Expression Omnibus (GEO).

Code availability

Scripts for the processing workflows schematized in Supplementary Figure 5 will be available at https://github.com/ggwlab/NTVE.

Acknowledgments

D-J.J.T, N.A., J.Ge, and G.G.W. gratefully acknowledge the financial support provided by the EIC Pathfinder project under grant number 101115574, and the ERC PoC under grant number 101138939. We thank Dr. Inti de la Rosa Velazques and Dr. Gertrud Eckstein from the Core Facility Genomics of Helmholtz Munich for library preparation and NGS and Dr. Carsten Peters from the TUM Electron Microscopy Facility for the help with cryo-EM data acquisition. We also thank the Groll Laboratory at TUM for access to their DLS instrument.

Contributions

N.A. developed methodology, designed genetic constructs, designed and conducted the cell and biochemical experiments, generated stable cell lines, generated schematics and figures, and co-wrote the manuscript;

J.Ge developed methodology, designed genetic constructs, co-conducted experiments on barcoded RNA export, generated stable cell lines, designed and conducted the cell-cell-communication experiments, and drafted some of the respective accompanying figures;

M.G. analyzed and visualized the NGS data, co-designed the perturbation experiments, conducted the IFN-gamma and Doxycycline vs DMSO perturbation experiments,

L.S. worked with N.A. on the biochemical experiments;

S.Sch. worked with N.A. on the differentiation of hIPSC;

T. P. worked with N.A. on the RT-qPCR analysis of $NTVE_{PABP}$ and the ASCL1 transactivation;

E.S., F.W., N.W., M.S., A.S., and T.S. assisted with cell and biochemical experiments;

O.B. obtained and analyzed the cryo-EM data;

F.T. advised population-specific experiments and data analysis;

J.Ga. advised on perturbation experiments and methodology, reviewed the text and NGS analyses;

D-J.J.T developed methodology, designed genetic constructs, designed experiments, supervised the research project, reviewed the data analysis and presentation;

G.G.W. conceptualized, initiated, and secured funding for the research program, developed methodology, designed experiments, reviewed data analysis and presentation, supervised and administered research, and wrote the manuscript.

Competing interests

All authors declare no competing interest.

References

1. Chen, W. *et al.* Live-seq enables temporal transcriptomic recording of single cells. *Nature* **608**, 733–740 (2022).
2. Cao, Y. *et al.* Nondestructive nanostraw intracellular sampling for longitudinal cell monitoring. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E1866–E1874 (2017).
3. Shueng, P.-W., Shih, K.-C., Gambhir, S. S., Kuo, D.-Y. & Chuang, H.-Y. Cancer detection using an artificial secretable MicroRNA found in blood and urine. *Biomedicines* **10**, 621 (2022).
4. Cherbonneau, F. *et al.* TRACE-seq: A transgenic system for unbiased and non-invasive transcriptome profiling of living cells. *iScience* **25**, 103806 (2022).
5. Gheysen, D. *et al.* Assembly and release of HIV-1 precursor Pr55gag virus-like particles from recombinant baculovirus-infected insect cells. *Cell* **59**, 103–112 (1989).
6. Galla, M., Will, E., Kraunus, J., Chen, L. & Baum, C. Retroviral pseudotransduction for targeted cell manipulation. *Mol. Cell* **16**, 309–315 (2004).
7. Prel, A. *et al.* Highly efficient in vitro and in vivo delivery of functional RNAs using new versatile MS2-chimeric retrovirus-like particles. *Mol. Ther. Methods Clin. Dev.* **2**, 15039 (2015).
8. Truong, D.-J. J. *et al.* Intron-encoded cistronic transcripts for minimally invasive monitoring of coding and non-coding RNAs. *Nat. Cell Biol.* **24**, 1666–1676 (2022).
9. Horns, F. *et al.* Engineering RNA export for measurement and manipulation of living cells. *Cell* **186**, 3642–3658.e32 (2023).
10. Litke, J. L. & Jaffrey, S. R. Highly efficient expression of circular RNA aptamers in cells using autocatalytic transcripts. *Nat. Biotechnol.* **37**, 667–675 (2019).
11. Stanifer, M. L., Cureton, D. K. & Whelan, S. P. J. A recombinant vesicular stomatitis virus bearing a lethal mutation in the glycoprotein gene uncovers a second site suppressor that restores fusion. *J. Virol.* **85**, 8105–8115 (2011).
12. Machida, K. *et al.* Dynamic interaction of poly(A)-binding protein with the ribosome. *Sci. Rep.* **8**, 17435 (2018).
13. Groft, C. M. & Burley, S. K. Recognition of eIF4G by rotavirus NSP3 reveals a basis for mRNA circularization. *Mol. Cell* **9**, 1273–1283 (2002).
14. Kahvejian, A., Svitkin, Y. V., Sukarieh, R., M'Boutchou, M.-N. & Sonenberg, N. Mammalian poly(A)-binding protein is a eukaryotic translation initiation factor, which acts via multiple mechanisms. *Genes Dev.* **19**, 104–113 (2005).
15. Ahler, E. *et al.* Doxycycline alters metabolism and proliferation of human cell lines. *PLoS One* **8**, e64561 (2013).
16. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).
17. Albritton, L. M., Kim, J. W., Tseng, L. & Cunningham, J. M. Envelope-binding domain in the cationic amino acid transporter determines the host range of ecotropic murine retroviruses. *J. Virol.* **67**, 2091–2096 (1993).
18. Durrant, M. G. *et al.* Systematic discovery of recombinases for efficient integration of large DNA sequences into the human genome. *Nat. Biotechnol.* **41**, 488–499 (2023).
19. Truong, D.-J. J. *et al.* Exonuclease-enhanced prime editors. *Nat. Methods* **21**, 455–464 (2024).
20. Sekine, K. *et al.* Robust detection of undifferentiated iPSC among differentiated cells. *Sci. Rep.* **10**, 10293 (2020).
21. Lopes, I., Altab, G., Raina, P. & de Magalhães, J. P. Gene size matters: An analysis of gene length in the human genome. *Front. Genet.* **12**, 559998 (2021).
22. Wen, W. *et al.* An efficient screening system in yeast to select a hyperactive piggyBac transposase for mammalian applications. *Int. J. Mol. Sci.* **21**, 3064 (2020).
23. Yusa, K., Zhou, L., Li, M. A., Bradley, A. & Craig, N. L. A hyperactive piggyBac transposase for mammalian applications. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 1531–1536 (2011).
24. Vora, J. *et al.* Alejandro Chavez.
25. Giehrl-Schwab, J. *et al.* Parkinson's disease motor symptoms rescue by CRISPRa-reprogramming astrocytes into GABAergic neurons. *EMBO Mol. Med.* **14**, e14797 (2022).
26. Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
27. Zuiderveld, K. Contrast limited adaptive histogram equalization. Graphics gems. 474–485 (1994).
28. Gao, Y. *et al.* ESPRESSO: Robust discovery and quantification of transcript isoforms from error-prone long-read RNA-seq data. *Sci. Adv.* **9**, eabq5072 (2023).
29. Harrow, J. *et al.* GENCODE: The reference human genome annotation for The ENCODE Project. *Genome Res.* **22**, 1760–1774 (2012).

30. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
31. Harrison, P. W. *et al.* Ensembl 2024. *Nucleic Acids Res.* **52**, D891–D899 (2024).
32. 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).
33. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
34. Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *Gigascience* **10**, (2021).