Transcriptome *in vivo* analysis (TIVA) of spatially defined single cells in live tissue

Ditte Lovatt[1,6], Brittani K Ruble[2,6], Jaehee Lee[1], Hannah Dueck[3], Tae Kyung Kim[1], Stephen Fisher[3], Chantal Francis[3], Jennifer M Spaethling[1], John A Wolf[4], M Sean Grady[4], Alexandra V Ulyanova[4], Sean B Yeldell[2], Julianne C Griepenburg[2], Peter T Buckley[1], Junhyong Kim[3,5], Jai-Yoon Sul[1], Ivan J Dmochowski[2,7] & James Eberwine[1,5,7]

Transcriptome profiling of single cells resident in their natural microenvironment depends upon RNA capture methods that are both noninvasive and spatially precise. We engineered a transcriptome *in vivo* analysis (TIVA) tag, which upon photoactivation enables mRNA capture from single cells in live tissue. Using the TIVA tag in combination with RNA sequencing (RNA-seq), we analyzed transcriptome variance among single neurons in culture and in mouse and human tissue *in vivo*. Our data showed that the tissue microenvironment shapes the transcriptomic landscape of individual cells. The TIVA methodology is, to our knowledge, the first noninvasive approach for capturing mRNA from live single cells in their natural microenvironment.

Multicellular organisms are composed of a myriad of cells that are categorized into different types based on phenotypic traits. However, even cells of seemingly the same type are not identical at the molecular level[1–3] and demonstrate heterogeneity among their expressed mRNAs and proteins, which can be influenced by cellular stimulation. Most knowledge about variability in gene expression has been extracted from studies using single-cell organisms or cells naturally occurring in suspension[4–7]. It is unknown whether processes that govern variability in gene expression in unicellular organisms can be extrapolated to the cells of multicellular organisms. Notably, the tissue microenvironment can be considerably diverse, and it is expected that with extracellular heterogeneity comes heterogeneity in gene expression. Therefore, tools for investigation of the transcriptome from single cells in tissue provide a unique opportunity to assess heterogeneity in mammalian cells and its biological importance.

RNA-seq is a tool for exploring a single cell's pool of mRNA at unprecedented depth and detail. However, isolating mRNA from single cells is technically challenging, especially for cells in complex tissues such as the brain. Existing methods for isolating mRNA from populations of living cells, including neurons[8–10], rely on sorting cells in suspension from acutely dissociated tissues, in which information about cell morphology and the microenvironment is lost, and where cell variability is masked by the averaging effect[11]. Other methods, such as laser capture microdissection and patch-pipette aspiration[12,13], can be used to isolate single cells in tissue, but both approaches have limitations, including potential RNA contamination from other cells in incidental contact with the patch pipette. Furthermore, the former is performed on fixed tissue, and the latter prompts concern about transcriptional changes associated with mechanical injury during the isolation of RNA[14]. Hence, an mRNA capture methodology with high spatial resolution and that is compatible with live, intact tissue would be a useful tool to explore the transcriptomes of single cells in their natural microenvironment.

Here we describe a method for isolating mRNA from a single cell in complex tissues using a photoactivatable mRNA capture molecule called the TIVA tag. We demonstrate the utility of the TIVA tag in both cell culture and brain tissue for capture of single-cell mRNA for subsequent RNA-seq transcriptome analysis. We apply this approach to study the unique transcriptional landscape of single neurons *in vivo* and how their transcriptomes differ fundamentally from those of cells in culture.

RESULTS

The TIVA tag captures cellular mRNA upon photoactivation

To perform transcriptome analysis of individually selected cells in intact tissue, we engineered a multifunctional photoactivatable mRNA capture molecule that we named the TIVA tag. When added to tissue, the tag penetrates the cell membrane by virtue of a disulfide-linked cell-penetrating peptide (CPP; **Fig. 1**); CPPs have been used to transport a variety of biomolecules into cells *in vitro* and *in vivo*[15–18]. We incorporated a pair of fluorophores that provide a fluorescence resonance energy transfer (FRET) signal to the caged TIVA tag to allow visualization of the uptake of the TIVA tag as well as uncaging in cells.

[1]Department of Pharmacology, University of Pennsylvania, Philadelphia, Pennsylvania, USA. [2]Department of Chemistry, University of Pennsylvania, Philadelphia, Pennsylvania, USA. [3]Department of Biology, University of Pennsylvania, Philadelphia, Pennsylvania, USA. [4]Department of Neurosurgery, University of Pennsylvania Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, USA. [5]PENN Genome Frontiers Institute, University of Pennsylvania, Philadelphia, Pennsylvania, USA. [6]These authors contributed equally to this work. [7]These authors jointly directed this work. Correspondence should be addressed to J.E. (eberwine@mail.med.upenn.edu).

RECEIVED 24 JUNE 2013; ACCEPTED 26 NOVEMBER 2013; PUBLISHED ONLINE 12 JANUARY 2014; DOI:10.1038/NMETH.2804

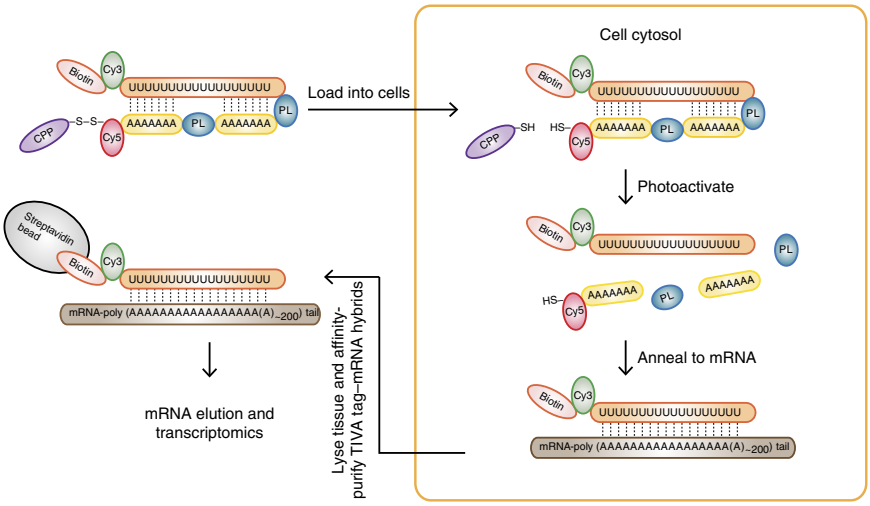

Figure 1 | The TIVA tag is a multifunctional, caged mRNA-capture molecule. The TIVA tag is composed of several functional groups: biotin, Cy3, poly(A) tail binding 2′-F RNA poly(U) oligo (orange), photocleavable linker (PL), 2′-OMe RNA poly(A) oligo (yellow), Cy5, disulfide bond (S-S) and CPP. The CPP is cleaved in the intracellular environment. The Cy3 and Cy5 groups form a FRET pair that allows real-time monitoring of cellular uptake as well as uncaging.

The cytosolic environment cleaves the CPP from the TIVA tag[17], trapping the caged tag inside the cell. Then, by selective laser-mediated photoactivation[19] of the TIVA tag in the desired cell or cells, the mRNA-capturing moiety is revealed and subsequently anneals to a poly(A) tail of cellular mRNA. We additionally engineered an affinity tag (biotin) at the end of the mRNA-capturing moiety to enable affinity purification of the TIVA tag–mRNA hybrids with the captured mRNA. The captured mRNA can be further processed for transcriptome analysis.

To control mRNA capture both spatially and temporally, we engineered the mRNA capture moiety as a light-activated ('caged') hairpin oligonucleotide. Photoactivation of the hairpin uncages and exposes an 18-mer poly(2′-deoxy-2′-fluorouridine) (poly(U)2FRNA) capture oligo, which binds the poly(A) tail of mRNAs[20,21] (**Fig. 2a**). We engineered the lengths of the oligos to achieve (i) a stably hybridized hairpin at physiological temperature, (ii) rapid dissociation after photoactivation and yet (iii) a stably hybridized TIVA tag–poly(U)2FRNA-mRNA hybrid after photoactivation. To this end, we engineered the TIVA tag with two shorter 7-mer poly(2′-O-methyladenosine) oligo-blocking strands separated by photoactivatible linkers to introduce greater instability to the hairpin after photoactivation: melting point $(T_m)_{caged}$ = 59 ± 1 °C (± value indicates the range), $T_{m\ uncaged}$ < 25 °C; Δ T_m > 35 °C (**Fig. 2b**). High mRNA capture efficiency should arise from the thermal stability between the poly(U)2FRNA and a standard 20-mer poly(A) mRNA target ($T_{m\ poly(U)2FRNA:poly(A)}$ = 50 ± 1 °C).

After solid-phase synthesis of the TIVA tag, in a final step we conjugated a (D-Arg)$_9$ CPP, which has been shown to transport cargos into the cytosol and nuclei of cells[22]. We validated chemical synthesis of the tag by high-performance liquid chromatography (HPLC), gel electrophoresis and matrix-assisted laser desorption/ionization–time of flight (MALDI-TOF) mass spectrometry, which recorded a mass-to-charge ratio (*m/z*) of 14,430.31 corresponding to the intact (D-Arg)$_9$ peptide–conjugated TIVA tag with an expected *m/z* of 14,413 (**Supplementary Fig. 1**). To monitor cellular loading and photoactivation inside cells, we engineered the TIVA tag to include a Cy3-Cy5 FRET pair. The Cy3-Cy5 FRET efficiency for the TIVA tag *in vitro* was 83% before photolysis and 9% after activation, with an overall change of 74% (**Fig. 2c**). Evaluation of uncaging conditions for TIVA tag in solution demonstrated that excitation with near-UV or 405-nm wavelengths could cleave the photoactivatible linkers (**Fig. 2d**).

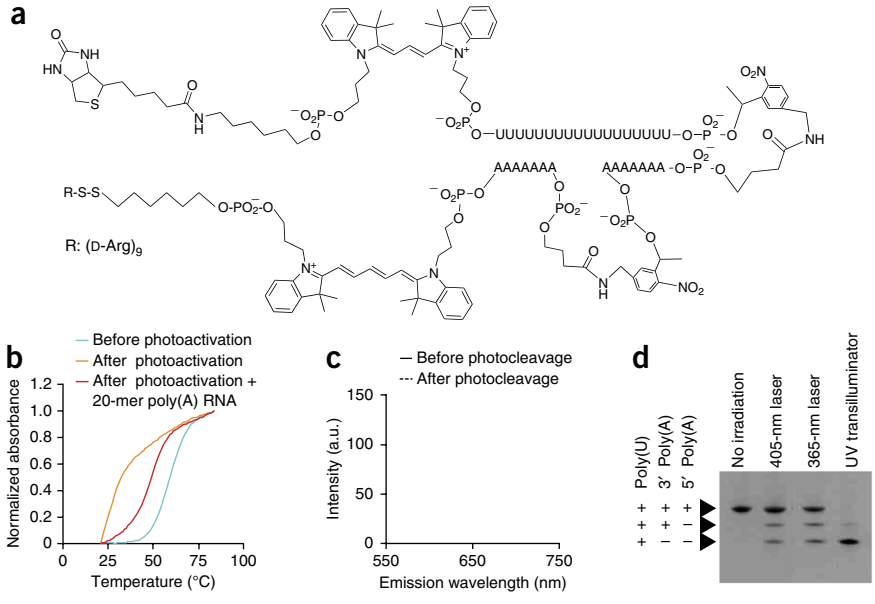

Figure 2 | Validation of the TIVA tag in solution. (**a**) Chemical structure of the TIVA tag. (**b**) Thermal denaturation curves of TIVA tag to determine the T_m before and after photoactivation (365 nm), and after photoactivation in the presence of 20-mer poly(A) RNA to simulate hybridization to mRNA (average of three experiments). (**c**) Efficiency of the Cy3-Cy5 FRET pair in TIVA tag in solution before and after photocleavage given by changes in the Cy3 (blue) and Cy5 (red) emission (average of three experiments). (**d**) Denaturing PAGE of TIVA tag without irradiation or with photoactivation at different wavelengths as indicated. Note that incomplete irradiation with 405-nm and 365-nm laser lanes results from the laser beam intersecting only a fraction of the droplet. 405-nm laser, partial irradiation of droplet for 9 × 101 µs with 405-nm laser at 27 mW; 365-nm laser, partial irradiation of droplet for 10 s with 365-nm laser at 50 mW; UV transilluminator, 15 min irradiation with 365-nm transilluminator (9 mW/cm^2).

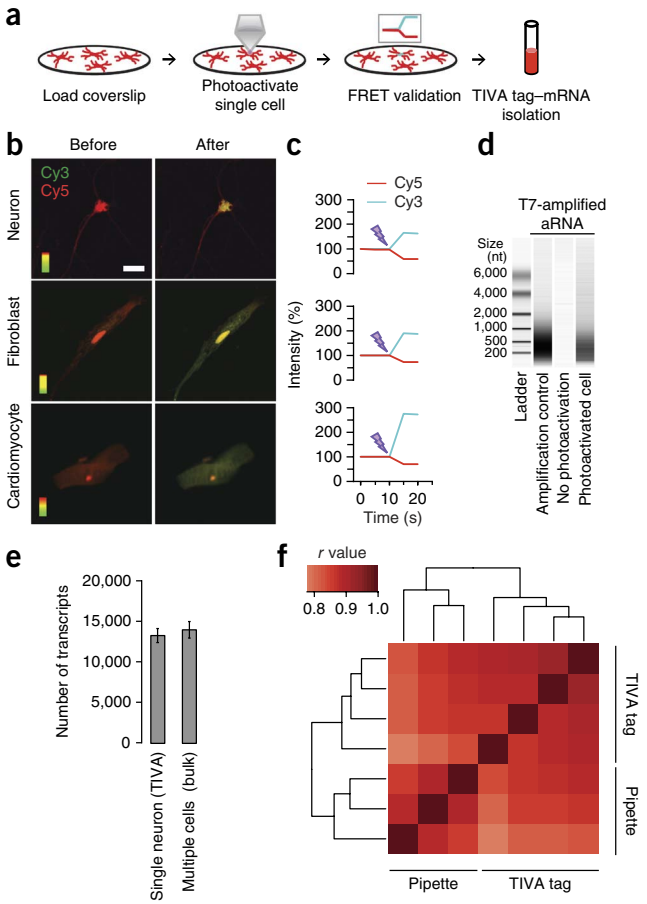

Figure 3 | Validation of uptake and uncaging of TIVA tag in live cells. (**a**) Schematic of the experimental steps. (**b**) Fluorescence photomicrographs showing uptake of $(D\text{-Arg})_9$-labeled TIVA tag by indicated cell types after the entire area of the cell was photoactivated using the 405 nm laser line. Scale bar, 20 µm. (**c**) Fluorescence intensity for the cells in **b**, showing the photoactivation of the TIVA tag by loss of FRET signal with a simultaneous increase in Cy3 fluorescence and decrease in Cy5 fluorescence. (**d**) Bioanalyzer microgel image analysis showing aRNA from coverslips with and without a photoactivated cell. (**e**) Number of expressed transcripts in single neurons or in bulk samples, defined as those with greater than ten unique exon reads per transcript using normalized RNA-seq data ($n = 3$ neurons for pipette isolation and $n = 4$ neurons for TIVA-tag isolation; error bars, s.d.) (**f**) Heatmap of Spearman correlation coefficients among single cells collected with pipette or TIVA tag, performed using normalized $\log_2$ read-count data.

The TIVA tag captures mRNA in dispersed primary cells

We tested isolation of mRNA with the TIVA tag in primary cultures of live, dispersed mouse brain cells (mixed culture; **Fig. 3a**). By exciting only the Cy3 fluorophore and using Cy5 fluorescence as an indicator of loading, we observed the intact TIVA tag to load into several types of cells in addition to neurons (**Fig. 3b**). Without the $D\text{-Arg}_9$ CPP, the TIVA tag did not enter cells. Notably, each type of CPP exhibits cell selectivity[22,23], which may be advantageous. Photoactivation of a single neuron resulted in a loss in FRET signal, showing that the TIVA tag can be uncaged in live cells (**Fig. 3c**)[24]. We next lysed all cells on the coverslip, affinity-isolated the TIVA tag–mRNA hybrids and amplified[25,26] the isolated mRNA. Bioanalyzer analysis of the amplified material showed several hundred nanograms of amplified RNA (aRNA), whereas in the absence of photoactivation there was no amplifiable nucleic acid, indicating that TIVA tag did not spontaneously activate inside cells and that there was little nonspecific mRNA capture during the affinity-purification process (**Fig. 3d**). RNA-seq yielded ~48 million reads per single cell sample, with 90–96% of the reads uniquely aligning, and transcriptome analysis revealed that on average individual neurons expressed about 9,000 different genes with at least ten reads (**Fig. 3e**). The correlation coefficients among single dispersed neuron transcriptomes collected with pipette ($r = 0.87$) or collected with the TIVA tag ($r = 0.89$) were not significantly different ($P = 0.23$, t-test; **Fig. 3f**), suggesting the two approaches capture a similarly complex population of mRNA transcripts.

Spatially controlled photoactivation of TIVA tag in live tissue

To assess the efficacy of TIVA-mediated mRNA capture in live brain tissue, we added TIVA tag to acute hippocampal slices from mice (**Fig. 4a** and **Supplementary Fig. 2**). Pyramidal neurons took up the TIVA tag, and photoactivation of single neurons resulted in a robust loss of FRET signal, validating efficient uncaging (**Fig. 4b**). Neurons adjacent to the photoactivated cell did not demonstrate changes in FRET (**Fig. 4c**), showing cell-specific photoactivation and retention of the photoactivated TIVA tag in the cell(s) of interest and that the photoactivation does not bleed or migrate to neighboring cells. After photoactivation, we aspirated the region containing the field of view, lysed the tissue and isolated TIVA tag–mRNA hybrids by affinity purification. Again, amplification from a single photoactivated cell resulted in generation of aRNA, whereas adjacent cells that were not irradiated upon isolation of the TIVA tag and processing through the RNA-amplification procedure showed no amplification. RNA-seq in single neurons revealed that neuronal markers were greatly enriched, whereas glial and vascular markers were almost entirely absent, validating the single-cell specificity of the TIVA capture procedure (**Fig. 4d**).

We tested whether residual uncaged TIVA tags, which had not annealed to cellular mRNA from a single photoactivated cell, had the ability to capture mRNA liberated from other cells during the affinity-isolation procedure. We carried out the identical uncaging and isolation procedure, and spiked in cardiomyocyte RNA at the affinity-isolation step. By reverse transcription–PCR, we saw no evidence for capture of the abundant cardiomyocyte-specific transcript, troponin, above noise levels (data not shown). Furthermore, loading and uncaging of the TIVA tag did not cause transcriptome changes (**Supplementary Fig. 3** and **Supplementary Table 1**) or cell death (**Supplementary Fig. 4**) and did not trigger cellular pathways that recognize dsRNA (**Supplementary Table 2**), indicating that the tag was not toxic to cells. Finally, we examined the stability of the TIVA tag and determined that we could still visualize and uncage it in single neurons in tissue up to 9 h after loading, which supports the utility of the TIVA methodology in experiments that require extended time periods (**Supplementary Fig. 5**).

Single-cell RNA-seq analysis of the TIVA-isolated mRNA yielded ~52 million reads per sample with 72–90% of the reads uniquely aligning. When evaluating the number of genes expressed, defined as at least ten reads in at least one sample, we found that 12,137 genes were expressed across all

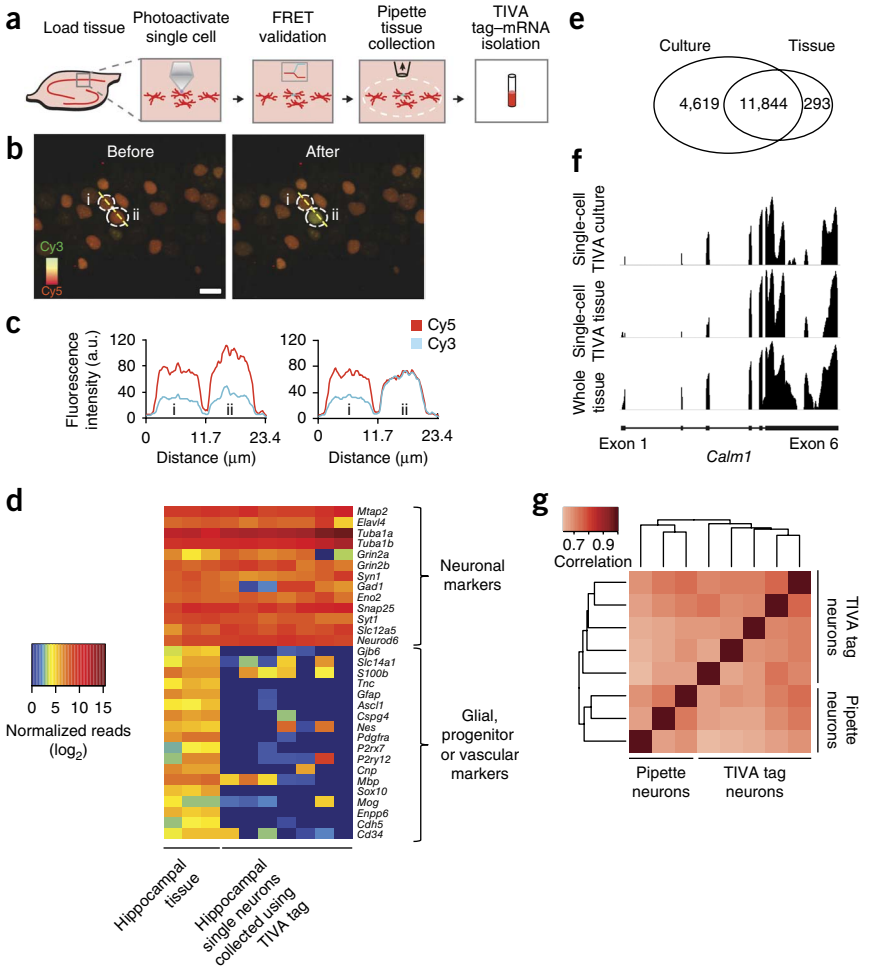

Figure 4 | TIVA tag capture of mRNA from single neurons in mouse hippocampal slices. (**a**) Schematic of the experimental steps. (**b**) Micrograph of TIVA tag–labeled neurons in hippocampus CA1 area before and after uncaging (405 nm). Dotted white lines indicate two neurons, labeled 'i' and 'ii', from which the resting FRET signal was recorded. Cy3 is pseudocolored green and Cy5 red. (**c**) Cy3 and Cy5 fluorescence intensity (a.u., arbitrary units) in neurons 'i' and 'ii', quantified using a line scan (dashed line) intersecting both cells. Images were captured as *Z* stacks (14 sections in 10 μm ranges, uncaging was performed in the middle of image stack) and merged as top view for analysis. Scale bar, 10 μm. (**d**) Heatmap of expressed neuronal, glial, progenitor and vascular markers transcripts using normalized RNA-seq data. (**e**) Overlap of expressed transcripts between single neurons collected using TIVA tag in tissue versus in culture. Expressed transcripts were defined as those with at least ten unique exon reads per transcript in at least one sample within its group using normalized RNA-seq data (3 tissue samples and 7 single cells). (**f**) RNA-seq pileup of unique exon reads aligning to the transcript *Calm1* (4,107 bp) in samples from single-cell TIVA culture (vertical label: 0–307 read counts), single cell TIVA tissue (vertical label: 0–147 read counts), and in whole tissue (vertical label: 0–531 read counts). *Calm1* gene structure shown at bottom depicts exon 1 through exon 6 (blocks, exons; solid lines, introns). (**g**) Heatmap of Spearman correlation coefficients between pipette and TIVA tag collected single hippocampal neurons in tissue.

single neurons in tissue. This is about 30% lower than for single neurons in culture, which expressed 16,463 genes (this is distinct from the average shown in **Fig. 3e**), but is in line with previous reports of freshly dissociated and acutely isolated pools of neurons expressing on average ~10,000 transcripts when assessed by microarray analysis (**Fig. 4e**)[8–10]. Despite these differences, the majority of these genes (~98%) were expressed in both contexts. Among the remaining genes, 293 were specific to the tissue context and 4,619 were only seen in culture (**Supplementary Table 3**). Shared genes were highly expressed in both contexts (1,650 average reads in tissue and 997 average reads in culture), which may indicate that these are genes with more ubiquitous or general function, whereas the average read counts for non-overlapping genes were lower, with 42 reads on average in both tissue and culture contexts.

We evaluated whether the TIVA approach can be used to isolate full-length mRNAs with sufficient 5′-3′ coverage after amplification. To characterize 5′-end transcript coverage in TIVA tag–collected samples, we calculated the fraction of expressed genes with reads aligning to the 5′-most 250 bases. We found that on average 24% of the expressed genes in TIVA tag–collected samples demonstrated coverage at the 5′ end after three rounds of aRNA amplification. Even long transcripts of several thousands of nucleotides had read coverage over the entire length of the mRNA

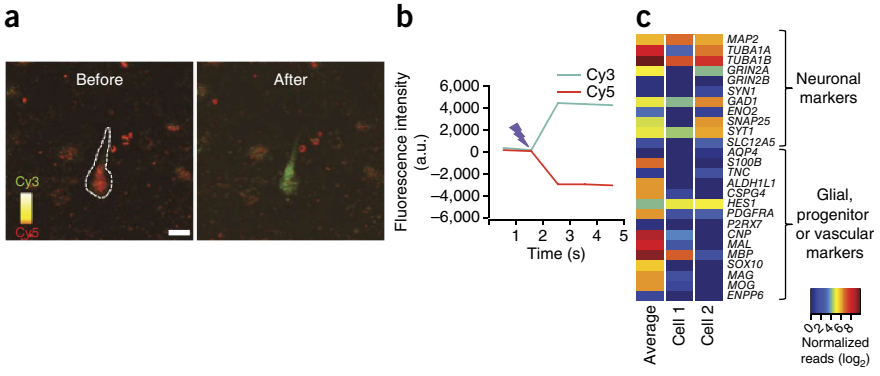

Figure 5 | TIVA tag capture of mRNA from cells in human live brain tissue specimen obtained from biopsy of the right frontal cortex from a subject undergoing surgery for communicating hydrocephalus. (**a**) Micrographs of TIVA tag–loaded cells identified by FRET signal, before and after uncaging, which was performed using the same parameters as in mouse. Scale bar, 10 μm. (**b**) FRET signal upon TIVA-tag activation (lightning bolt). (**c**) Heatmap comparing expression of common cell type–specific markers in an average pool of 13 TIVA tag captured cells and in two TIVA tag captured individual cells.

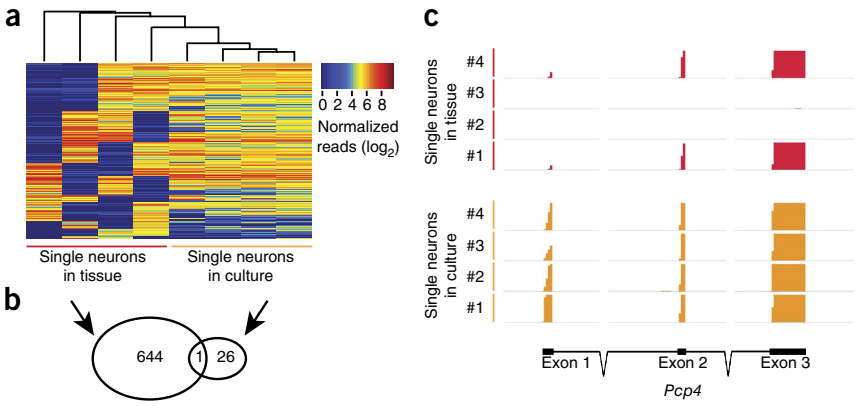

Figure 6 | Bimodal transcripts in single hippocampal neurons in tissue and in culture. (**a**) Heatmaps shows clustering of 645 bimodally expressed genes (horizontal lines) in four single cells from hippocampal tissue compared to four single hippocampal neurons in culture. Bimodally expressed genes were defined as having a gap in expression of at least four log units in two samples. To be especially stringent in this analysis, two samples were required to have expression values on either side of this gap. The cells with low expression were required to have fewer than ten normalized counts for the biomodally expressed transcript. (**b**) Overlap between bimodal genes in single neurons from tissue and from culture (4 cells in each group). (**c**) RNA-seq pileup of unique exon reads aligning to *Pcp4*, a gene showing bimodal expression in four cells in tissue and not in culture (4 cells). *Pcp4* gene structure (bottom) depicts exons 1–3 (blocks, exons; solid lines, introns; dotted lines, part of intron not shown at relative length). Vertical axis read count range for all samples: 0–300.

that was similar to that for bulk-isolated mRNA from whole tissue, suggesting that the isolated mRNA was not degraded during the isolation procedure (**Fig. 4f**). We also collected RNA from single hippocampal neurons using patch-pipette aspiration and found that the correlations among single tissue neuron transcriptomes collected by TIVA tag and patch pipette were $r = 0.71$ and 0.70, respectively, suggesting that both methods captured a similarly complex population of RNA (**Fig. 4g**). We did not find any differences in 5′-end read coverage for TIVA tag–collected samples and pipette-collected samples (**Supplementary Table 4**), nor did we observe differences in the evenness of coverage resulting from collection technique (**Supplementary Fig. 6**). As read depth can be influenced by G+C content and gene length[27–31], we estimated their effect as well as that of the collection technique on abundance estimates using analysis of variance (ANOVA). We observed that the effect sizes are small in all cases, accounting for less than 0.2% of observed variation in mean abundance (**Supplementary Fig. 7** and **Supplementary Table 5**).

These experiments demonstrate that the TIVA tag is an efficient tool for capturing mRNA of varying sizes and abundances from single cells in the live slice preparation.

TIVA on single human cells in live brain tissue

To test the performance of the TIVA tag on cells of the human brain, we obtained live surgically resected human brain tissue from an adult subject undergoing neurosurgery for communicating hydrocephalus. Cells in human tissue loaded efficiently with TIVA tags (**Fig. 5a**), and we observed a loss of FRET signal after uncaging (**Fig. 5b**). Although it is difficult to know a priori the identity of any particular cell in the live slice preparation from visual assessment of the TIVA tag–loaded human tissue, analysis after sequencing suggested that several of the cells were neurons because of the enrichment of multiple neuronal markers and the absence or low expression of multiple markers of other cell types (**Fig. 5c**). From measurements on five single human cells, we observed on average ~5,000 genes expressed per cell and ~12,500 unique expressed genes across all cells. These data show that the TIVA can be used to characterize the transcriptome of human cells in their natural tissue environment, including those from adults.

Tissue neurons demonstrate bimodal expression variability

We next asked whether any of the expressed transcripts in single neurons in tissue demonstrated a bimodal on-or-off expression mode, a subtype of transcriptome variability, between the individual cells. This measure of variability is particularly germane to single-cell analysis as the presence or absence of RNAs is likely reflected in altered cellular physiology. We examined bimodal expression in three whole-brain tissue samples derived from three different mice, finding only six transcripts that were bimodally expressed between these samples. Of these, four genes allocated to the Y chromosome in the two male animals, highlighting that Y chromosome gene-expression bimodality can be quite high for selected genes. Overall, these data demonstrate that we observed little bimodality in transcription when examining a population of cells. In contrast, we identified 27 bimodally expressed transcripts in single TIVA neurons from culture ($n = 8$ cells), and 645 bimodally expressed transcripts in single TIVA neurons from tissue (**Supplementary Table 6**). This is more than expected by chance ($P < 0.001$) and significantly more than found in culture ($P < 0.001$, Online Methods; **Fig. 6**). We found only a single bimodally expressed transcript in both sample types (**Fig. 6b**). We observed bimodality of expression for *Pcp4* (**Fig. 6c**). We then used the Allen Brain Atlas for the developing mouse brain at postnatal day (P)4 and P14 to examine the expression of a subset (87) of the bimodally expressed transcripts that we identified in single cells in tissue (**Supplementary Fig. 8**). We found that 71% of these genes demonstrated a spatially bimodal or speckled expression pattern among individual hippocampal CA1 cells from at least one age, and 22% demonstrated a bimodal expression pattern at both developmental ages. The remaining transcripts showed uniform strong expression during this developmental window.

DISCUSSION

We developed a noninvasive tool enabling capture of mRNA from single spatially defined cells in living, intact tissue for transcriptome analysis. Although pipette isolation permits isolation of RNA from single cells in tissue, the penetration of the pipette through the tissue and process of isolating cytosolic mRNA involves tissue deformation that may alter selective components of the

transcriptional profile even if the overall correlation with TIVA is high. The TIVA methodology is not limited to a certain cell type or species, which contrasts with the case for several existing RNA isolation methods that depend on transgenic rodent models to identify cells of interest. It should also be possible to synthesize RNA specific TIVA tags, which may prove useful for *in vivo* analysis of particular miRNAs and other RNAs that do not possess a poly(A) tail. Finally, the development of new TIVA tags for multiplexing mRNA capture, either in the same cell or in multiple neighboring cells, would allow analysis of gene expression patterns in multiple cells simultaneously or of mRNA expression over time in a single cell after a stimulus. New TIVA tags that incorporate efficient two photon–uncageable linkers (still commercially unavailable), and thus allow uncaging deeper in the tissue and with better three-dimensional resolution[32], should make it possible to use TIVA in conjunction with *in vivo* live-animal functional imaging.

In using TIVA tags, one must be cognizant of various issues, including: (i) if CPPs are used to introduce TIVA tags into cells, different CPPs can promote movement into different cell types and cellular compartments, (ii) upon long-term storage of the described TIVA tag (longer than 3 months), the FRET signal can be weak because of the loss of U-A base-pairing; hence, the TIVA tag should be used within 3 months of synthesis, (iii) CPP-associated TIVA tag should be stored in dried form at −80 °C rather than in solution aliquots to avoid clumping of the compound, (iv) as with all exogenous fluorescent indicators, optimal conditions should be identified for TIVA tag loading and (v) isolation of the TIVA tag-mRNA complex should be optimized for the specific isolation procedure being used (e.g., strepavidin-coated beads, biotin antibodies, phenol separation, etc.).

Tissue measurements mask both low-level mRNA expression in single cells and variation in expression between cells of the same type[11]. In our data on single cells in the context of intact tissue, some neurons expressed one or two traditional glial or vascular markers, although no neurons expressed all of the markers expected to be in a single nonneuronal cell type, and all neurons expressed classical neuronal markers. Similar data have been observed in previous histochemical studies where, for instance, both *Mbp* and *S100b* transcripts expressed primarily in oligodendrocytes and astrocytes, respectively, have been observed in neurons during the early postnatal weeks in human and mouse brain tissue[33–36]. These data suggest that the use of individual markers as surrogate discriminators of cell type is insufficient and that evaluation of an array of markers may provide a more informative evaluation.

Single-cell heterogeneity is a well-accepted phenomenon but has remained understudied in complex tissues owing to technical limitations of mRNA isolation. We showed using the TIVA tag that CA1 hippocampal neurons in live hippocampal tissue expressed fewer genes overall but had more bimodally expressed genes than hippocampal neurons in culture (**Fig. 6**), suggesting an important role for the microenviroment in modulating gene expression in cells. Recently, there have been suggestions that variation in expression is in part stochastic, arising from both intrinsic noise (stochastic nature of biochemical reactions) and extrinsic noise (changes in cellular regulatory proteins)[4]. Our data suggest that the ~30% difference in number of genes expressed may be attributable to extrinsic noise. When a neuron is removed from its natural environment, surface-sensing molecules are no longer subject to the complex regulatory constraints resulting from the thousands of synaptic inputs and cellular interactions converging on individual cells, and which in turn, tightly regulate the activity of transcriptional activators and repressors[37]. The ability of the cell to express 30% more genes when removed from its natural microenvironment further suggests that the epigenetic silencing of genes is either reduced or that the genes that turn on have not been not epigenetically silenced and represent the capacity of that particular cell to respond to different environmental cues. In other words, the microenvironment may serve to place constraints on the functional cellular phenotype, through increased variability and fewer expressed genes, which under different environmental cues is modified to provide an increased flexibility for the cell to transcriptionally respond. The uniqueness of synaptic input present in tissue may shape the transcriptomic landscape of individual neurons in tissue, providing cells with the constraints under which they can function. The power of TIVA in permitting transcriptomic analysis of single cells in intact physiological systems provides a new window into understanding how cells function normally.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. Gene Expression Omnibus: GSE52525 (RNA-seq data).

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

Requests for materials. Requests for TIVA tag should be directed to ivandmo@sas.upenn.edu.

ACKNOWLEDGMENTS

We thank J. Cheung-Lau for assistance with *in vitro* FRET measurements. Funding was provided by the PhRMA foundation to D.L., US National Institutes of Health (NIH) R01 GM083030 to I.J.D., McKnight Foundation Technology Innovations Award to I.J.D. and J.E., U01MH098953 to J.K. and J.E. and NIH DP004117 to J.E. This project is funded, in part, by the Penn Genome Frontiers Institute under a grant with the Pennsylvania Department of Health, which disclaims responsibility for any analyses, interpretations or conclusions.

AUTHOR CONTRIBUTIONS

All authors contributed to the writing of the manuscript. D.L. performed dispersed-cell TIVA experiments and some computational analysis; B.K.R. contributed TIVA-tag characterization and supplied TIVA tag; J.L. perfomed slice experiments and the bulk of TIVA uncaging; H.D. performed the bulk of the computational analysis; T.K.K. performed TIVA-mediated RNA amplifications; S.F. contributed computational analysis; C.F. contributed some control samples; J.M.S. contributed TIVA-mediated RNA amplifications; J.A.W., M.S.G. and A.V.U. organized human tissue use; S.B.Y. and J.C.G. contributed TIVA tag; P.T.B. contributed TIVA-mediated RNA amplifications; J.K. directed the computational analysis; J.Y.S. contributed the TIVA uncaging parameters and oversaw the biophotonics; I.J.D. designed experiments and contributed oversight of TIVA-tag synthesis; J.E. designed experiments and contributed oversight of the biological experiments.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Kaern, M., Elston, T.C., Blake, W.J. & Collins, J.J. Stochasticity in gene expression: from theories to phenotypes. *Nat. Rev. Genet.* **6**, 451–464 (2005).
2. Raj, A. & van Oudenaarden, A. Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell* **135**, 216–226 (2008).

3. Eldar, A. & Elowitz, M.B. Functional roles for noise in genetic circuits. *Nature* **467**, 167–173 (2010).
4. Elowitz, M.B., Levine, A.J., Siggia, E.D. & Swain, P.S. Stochastic gene expression in a single cell. *Science* **297**, 1183–1186 (2002).
5. Flatz, L. *et al.* Single-cell gene-expression profiling reveals qualitatively distinct CD8 T cells elicited by different gene-based vaccines. *Proc. Natl. Acad. Sci. USA* **108**, 5724–5729 (2011).
6. Taniguchi, Y. *et al.* Quantifying *E. coli* proteome and transcriptome with single-molecule sensitivity in single cells. *Science* **329**, 533–538 (2010).
7. Pedraza, J.M. & van Oudenaarden, A. Noise propagation in gene networks. *Science* **307**, 1965–1969 (2005).
8. Cahoy, J.D. *et al.* A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function. *J. Neurosci.* **28**, 264–278 (2008).
9. Lovatt, D. *et al.* The transcriptome and metabolic gene signature of protoplasmic astrocytes in the adult murine cortex. *J. Neurosci.* **27**, 12255–12266 (2007).
10. Sugino, K. *et al.* Molecular taxonomy of major neuronal classes in the adult mouse forebrain. *Nat. Neurosci.* **9**, 99–107 (2006).
11. Eberwine, J. *et al.* Quantitative biology of single neurons. *J. R. Soc. Interface* **9**, 3165–3183 (2012).
12. Espina, V. *et al.* Laser-capture microdissection. *Nat. Protoc.* **1**, 586–603 (2006).
13. Tang, F. *et al.* mRNA-seq whole-transcriptome analysis of a single cell. *Nat. Methods* **6**, 377–382 (2009).
14. Okaty, B.W., Sugino, K. & Nelson, S.B. A quantitative comparison of cell-type-specific microarray gene expression profiling methods in the mouse brain. *PLoS ONE* **6**, e16493 (2011).
15. Joliot, A. & Prochiantz, A. Transduction peptides: from technology to physiology. *Nat. Cell Biol.* **6**, 189–196 (2004).
16. Kumar, P. *et al.* Transvascular delivery of small interfering RNA to the central nervous system. *Nature* **448**, 39–43 (2007).
17. Zeng, F. *et al.* A protocol for PAIR: PNA-assisted identification of RNA binding proteins in living cells. *Nat. Protoc.* **1**, 920–927 (2006).
18. Zielinski, J. *et al. In vivo* identification of ribonucleoprotein-RNA interactions. *Proc. Natl. Acad. Sci. USA* **103**, 1557–1562 (2006).
19. Adams, S.R. & Tsien, R.Y. Controlling cell chemistry with caged compounds. *Annu. Rev. Physiol.* **55**, 755–784 (1993).
20. Tang, X. & Dmochowski, I.J. Synthesis of light-activated antisense oligodeoxynucleotide. *Nat. Protoc.* **1**, 3041–3048 (2006).
21. Dmochowski, I.J. & Tang, X. Taking control of gene expression with light-activated oligonucleotides. *Biotechniques* **43**, 161–165 (2007).
22. Madani, F., Lindberg, S., Langel, U., Futaki, S. & Graslund, A. Mechanisms of cellular uptake of cell-penetrating peptides. *J. Biophys.* **2011**, 414729 (2011).
23. Svensen, N., Walton, J.G. & Bradley, M. Peptides for cell-selective drug delivery. *Trends Pharmacol. Sci.* **33**, 186–192 (2012).
24. Roy, R., Hohng, S. & Ha, T. A practical guide to single-molecule FRET. *Nat. Methods* **5**, 507–516 (2008).
25. Eberwine, J. *et al.* Analysis of gene expression in single live neurons. *Proc. Natl. Acad. Sci. USA* **89**, 3010–3014 (1992).
26. Morris, J., Singh, J.M. & Eberwine, J.H. Transcriptome analysis of single cells. *J. Vis. Exp.* **2011**, 2634 (2011).
27. Ramskold, D. *et al.* Full-length mRNA-Seq from single-cell levels of RNA and individual circulating tumor cells. *Nat. Biotechnol.* **30**, 777–782 (2012).
28. Griffith, M. *et al.* Alternative expression analysis by RNA sequencing. *Nat. Methods* **7**, 843–847 (2010).
29. Zheng, W., Chung, L.M. & Zhao, H. Bias detection and correction in RNA-Sequencing data. *BMC Bioinformatics* **12**, 290 (2011).
30. Adiconis, X. *et al.* Comparative analysis of RNA sequencing methods for degraded or low-input samples. *Nat. Methods* **10**, 623–629 (2013).
31. Gertz, J. *et al.* Transposase mediated construction of RNA-seq libraries. *Genome Res.* **22**, 134–141 (2012).
32. Ellis-Davies, G.C. Caged compounds: photorelease technology for control of cellular chemistry and physiology. *Nat. Methods* **4**, 619–628 (2007).
33. Zhang, S.C. Defining glial cells during CNS development. *Nat. Rev. Neurosci.* **2**, 840–843 (2001).
34. Pribyl, T.M. *et al.* Expression of the myelin basic protein gene locus in neurons and oligodendrocytes in the human fetal central nervous system. *J. Comp. Neurol.* **374**, 342–353 (1996).
35. Landry, C.F. *et al.* Myelin basic protein gene expression in neurons: developmental and regional changes in protein targeting within neuronal nuclei, cell bodies, and processes. *The J. Neurosci.* **16**, 2452–2462 (1996).
36. Vives, V., Alonso, G., Solal, A.C., Joubert, D. & Legraverend, C. Visualization of S100B-positive neurons and glia in the central nervous system of EGFP transgenic mice. *J. Comp. Neurol.* **457**, 404–419 (2003).
37. West, A.E., Griffith, E.C. & Greenberg, M.E. Regulation of transcription factors by neuronal activity. *Nat. Rev. Neurosci.* **3**, 921–931 (2002).

ONLINE METHODS

Synthesis and purification of TIVA tags. Custom chemical synthesis of the TIVA tag without the CPP was carried out using standard phosphoramidite chemistry on an ABI DNA/RNA synthesizer (model 394). TIVA-tag synthesis was done according to the structure outlined in **Figure 1** on either 1 µmol or 10 µmol scales with the following reagents (Glen Research): 2′-deoxy-2′-fluoro-uridine phosphoramidite (10-3430), 2′-O-methyladenosine phosphoramidite (10-3100), Cy3 phosphoramidite (10-5913), Cy5 phosphoramidite (10-5915), photocleavable spacer phosphoramidite (10-4913), thiol modifier C6 S-S phosphoramidite (10-1936) and 3′-biotin TEG CPG (20-2955). Coupling times were adjusted to manufacturer's recommendations, and 0.02 M iodine was used for oxidation steps. After cleavage and deprotection using ammonium hydroxide at room temperature for 24 h, TIVA tag was purified on a C18 column using reverse-phase HPLC (Agilent 1100S) with eluents of 0.05 M triethylammonium acetate (A) and acetonitrile (B) gradient: 10–60% B the first 0–40 min; then 60–80% B the remaining 40–50 min in A+B; flow rate was 1 ml/min; temperature was 40 °C. The retention time of the purified TIVA tag with 5′ thiol modification was ~50 min. The TIVA tag was desalted on a NAP-5 column (GE Healthcare) and dried under vacuum, before conjugation to the CPP.

Conjugation of TIVA tag to CPP. The method of conjugating the TIVA tag to CPP through a disulfide bond was modified from ref. 38. Briefly, about 5 nmol of oligonucleotide with 5′ thiol modification was deprotected using 50 mM TCEP for 2 h. The TCEP was removed by desalting on a NAP-5 column, and the TIVA tag was dried under vacuum. After drying, the TIVA tag was redissolved in 50 µl of 0.33 M TEAA, 150 µl of formamide was added and the sample was vortexed. The CPP, $(D\text{-}Arg)_9$, with an activated Cys(Npys) residue (Anaspec) was dissolved at a concentration of 1 mM in water. A fourfold excess of CPP was added to the dissolved TIVA tag, and the reaction was allowed to proceed overnight. The conjugation product was purified by anion exchange on 1 ml Resource Q column (Agilent 1100S HPLC) using a flow rate of 1 ml/min and a gradient of 0–100% buffer B in 30 min (buffer A: 20 mM Tris-HCl (pH 6.8) and 50% formamide; buffer B: 20 mM Tris-HCl (pH 6.8), 50% formamide and 400 mM $NaClO_4$). Finally, the product was desalted on a NAP-5 column, concentrated and characterized by MALDI-TOF mass spectrometry (**Supplementary Fig. 1**) using a MALDI-TOF mass spectrometer (Applied Biosystems Voyager System 6030) operated in negative ion mode with 3-hydroxypicolinic acid matrix.

Determination of melting point. Melting point studies were conducted on a Beckman Coulter DU800 UV-Vis spectrophotometer equipped with a programmable Peltier temperature controller. Samples were monitored at 260 nm while heating or cooling at a rate of 1.0 °C/min, with a 1 min hold per degree Celsius. Melting temperatures were determined from the peak of the first derivative plot of aborbance at 260 nm versus temperature. TIVA tag was prepared at 1 µM concentration in standard buffer (10 mM Tris pH 7.5, 300 mM NaCl and 10 mM $MgCl_2$). The melting temperature of the poly(U):poly(A) duplex was determined by preparing a solution of 1 µM TIVA tag and 1 µM 20-mer poly(A) RNA and irradiating the solution with 365 nm light for 15 min. To ensure proper annealing of the TIVA tag, samples were heated to 90 °C in a water bath for 5 min, and then slowly cooled to room temperature over ~3 h. For samples that were photolysed, irradiation was carried out using a UV transilluminator (Spectronics Corporation TL-365R) at wavelengths centered on 365 nm (9 mW/cm^2 at peak intensity) for 15 min with the samples in open 200-µl microcentrifuge tubes. Notably, thermal denaturation analysis for the uncaged TIVA tag indicates the melting temperature is <25 °C and is too low to be accurately determined experimentally by UV-vis spectrophotometry. Previous RNA bandages of short strand lengths have been shown to have melting temperatures that are <20 °C, which is consistent with the data presented here[39].

FRET analysis of TIVA tag in solution. TIVA tag was prepared at 1 µM concentration in standard buffer (see above), and hybridization and photolysis were carried out in the same way as above. Measurements were made in a sub-microcuvette incubated at 37 °C during collection of emission. Fluorescence emission from Cy3 at 565 nm and Cy5 at 667 nm, upon excitation at 552 nm, was monitored by a Varian Eclipse fluorimeter (scanning rate of 120 nm/min and averaging time of 0.50 s). The FRET efficiency was defined as: $I_a / ((\gamma \times I_d) + I_a)$, where I_a is the intensity of the acceptor (Cy5) fluorescence, I_d is the intensity of the donor (Cy3) fluorescence, and γ is the correction factor for the difference in donor and acceptor quantum yields.

PAGE of TIVA tag. TIVA tag in a water droplet was photolysed at 365 nm or 405 nm (177.32 µs per pixel over 303 µm × 303 µm area for 9 times repetition at 90% laser power) and analyzed by PAGE before and after photolysis. Samples were loaded on a 7 M urea, 20% polyacrylamide gel, and the gel was electrophoresed at 300 V for 40 min. After staining with ethidium bromide, the gel was imaged on a Bio-Rad Gel Doc 2000 system.

Loading TIVA tag into cell cultures and brain slices. Cultured hippocampal neurons from C57BL/6 mice were grown on coverslips as described elsewhere[40]. The majority of cells in these mixed cultures were neuronal with only with a small fraction of glia. After 7–9 d in culture, the coverslips were rinsed in prewarmed cell saline buffer (140 mM NaCl, 5.4 mM KCl, 1 mM $MgCl_2$, 2 mM $CaCl_2$, 16 mM glucose and 10 mM HEPES) and then placed in an empty Petri dish in a humidified chamber. Immediately after, 50 µl of 10 µM TIVA tag in cell saline buffer was added to the coverslip. After 15 min incubation and two rinses of cell saline buffer, the coverslip was transferred to an imaging chamber with cell saline buffer. Imaging and photolysis were performed immediately on cells using a confocal microscope (Zeiss 710 Meta, 40× water-immersion objective, numerical aperture (NA) 1.0). FRET was recorded under the same configuration as the tissue imaging experiments described below. Photolysis was performed using the 405-nm laser at 30% power and 50 µs per pixel. Coronal brain slices (220–270 µm) were prepared from 8–10-day-old C57BL/6 mice (Charles River) with a Leica VT2000 vibratome. Acutely isolated tissues were transferred to artificial cerebrospinal fluid (aCSF) (in mM: NaCl 122, $NaHCO_3$ 28, glucose 5.5, HEPES 10, KCl 3.5, $MgCl_2$ 1, $CaCl_2$ 2, pH 7.4 with 5% CO_2 and 95% O_2 gas mixture) for 90 min, and subsequently loaded with TIVA tag (30 µM in aCSF) for 90 min at room temperature. Imaging and photoactivation were performed in the CA1 area of the hippocampus under the same configuration as the cell culture

experiments with the exception of using the 405-nm laser for uncaging at 80% power and 100.85 µs per pixel. Loading was confirmed by detecting Cy5 signal in the emission range excited by 561 nm, and uncaging was performed using the 405-nm laser while recording FRET excited by 514 nm and simultaneously capturing in Cy3 (538–599 nm) and Cy5 (637–704 nm) emission ranges. In some experiments, a line-scan analysis was performed using a 4-pixel-width line and averaging pixel values. To confirm that photodamage was not caused to the cell by uncaging, cells were loaded with Fluo-4 AM at 5 mg/ml in aCSF for 60 min at room temperature. Cells were imaged with the confocal microscope and excited at 488 nm, and emission was recorded at 530 nm ± 20 nm. In calcium-imaging experiments, photolysis laser power and duration used in TIVA tag experiment was confirmed not to photodamage the targeted cell (**Supplementary Fig. 4**). We verified that the power and duration used to phototactivate TIVA tag in cultures and slices did not cause cytosolic Ca^{2+} increases. After uncaging a single cell in the tissue, the imaged field and surrounding area, including the photolysed cell, were isolated by aspiration using a wide-bore glass pipette and further processed for mRNA analysis. A biopsy from a 71-year-old male subject undergoing neurosurgery for hydrocephalus was collected according to University of Pennsylvania Perelman School of Medicine Internal Review Board–approved guidelines. The subject, who provided informed consent, had been administered the following drugs: aspirin, finasteride, glipizide, metformin, metoporolol, multivitamin, omega-3 fatty acids, rosuvastatin calcium and tamsulosin HCl. The tissue specimen was placed into a 250-ml bottle with 4 °C sucrose cutting solution (248 mM sucrose, 1 mM KCl, 26 mM $NaHCO_3$, 10 mM glucose, 1 mM $CaCl_2$ and 10 mM $MgCl_2$, bubbled with 95% O_2 and 5% CO_2) immediately after resection and was quickly transported to the laboratory for slicing. The tissue was sliced with a vibratome into 200-µm sections. The sections were immediately transferred to the incubation chamber with normal aCSF (124 mM NaCl, 4 mM KCl, 26 mM $NaHCO_3$, 10 mM glucose, 2 mM $CaCl_2$ and 2 mM $MgCl_2$, bubbled with 95% O_2 and 5% CO_2) at room temperature to rest for 1 h and then the slice was loaded with 6 µM TIVA–tag in normal aCSF (bubbled with 95% O_2 and 5% CO_2) at room temperature for 2 h. Loaded cells were uncaged and mRNA was affinity-purified according to the procedure used for mouse tissue.

Affinity purification of TIVA tag. Cells or tissue were lysed in a final volume of 200 µl lysis buffer (10 mM Tris-HCl pH 7.5, 300 mM NaCl, 10 mM $MgCl_2$, 0.1–0.5% NP40 and 1 U/µl SUPERase-In RNase inhibitor) for 5 min on ice and in the dark. Then, 10 µl prewashed MyOne T1 or C1 streptavidin magnetic Dynabead was mixed in, and incubated for 15 min on ice in the dark. After three washes in wash buffer (20 mM Tris-HCl pH 7.5 and 50 mM NaCl), TIVA tag–mRNA hybrids were eluted at 60–70 °C in 9 µl nuclease-free water. The eluate was either immediately frozen or amplified.

Spiking assay. Cardiomyocyte total RNA was isolated from whole mouse cardiac tissue using Trizol reagent following the manufacturer's recommendation. TIVA tag was loaded into mouse brains slices as described above, and then two groups (TIVA tag–loaded uncaged and TIVA tag–loaded but not uncaged) of neurons were harvested into tubes containing 5 µl of the binding and washing (B&W) solution (Invitrogen). The volume of samples was adjusted to 200 µl using B&W solution and then 500 pg, 100 pg or 10 pg cardiomyocyte total RNA was added to the sample tubes, followed by the TIVA tag affinity purification as described above. TIVA tag–isolated RNA was then reverse-transcribed as described above. As positive controls, 500 pg, 100 pg and 10 pg cardiomyocyte total RNA were transcribed into cDNA. The cardiac cDNA was assayed for Troponin T using PCR with a 3′ end directed mouse cTnT specific primer set. The relative amounts of PCR amplicons were measured by agarose gel electrophoresis and band intensity was quantified using MetaMorph software.

Amplification, library construction and Illumina RNA-seq. mRNA was amplified using three rounds of linear *in vitro* transcription–based amplification as described elsewhere[26]. The size and amount of the aRNA was evaluated on a Bioanalyzer RNA Nanochip (Agilent). Between 100 ng and 1,000 ng of aRNA was used for Tru-Seq library construction according to the manufacturer's instructions and then submitted for Illumina 1000 sequencing.

RNA-seq data analysis and bioinformatics. After trimming for sequencing adaptor and poly(A) contamination using in-house software, raw sequencing data were aligned to the mouse genome and transcriptome using the RNA-seq unified mapper (RUM) and mouse genome build mm9 (ref. 41). We used HTSeq (analyzing high-throughput sequencing data with Python, developed by S. Anders at EMBL Heidelberg (Genome Biology Unit)) to match unique reads with University of California Santa Cruz (UCSC) genome browser known gene annotations to generate gene counts. To mitigate differences in read depth across samples, counts were normalized using the method as described[42]. Heatmaps were generated in R, using the $\log_2$ of normalized counts and the heatmap.2 method of the gplots library (http://www.r-project.org/, G.R. Warnes, gplots: various R programming tools for plotting data; 2008). Spearman correlations were used to generate correlation heatmaps. Genes hypothesized to have bimodal expression distributions were identified as follows: log expression values for a given set of samples were required to demonstrate a gap in expression of at least four log units. In addition, at least two samples were required to have expression values on either side of this gap (except in the case of bulk tissue where, because there were only three samples, at least two samples were required to have low expression values). Finally, samples with low expression were required to have fewer than ten normalized counts. If all these criteria were met, a gene was hypothesized to be bimodal. The R library 'VennDiagram' was used to visualize the sets of bimodal genes observed in tissue and in culture (Hanbo Chen Venn diagram: generate high-resolution Venn and Eurler plots. R Package v1.6.0 (2013), http://cran.r-project.org/web/packages/VennDiagram/index.html). For cases where values are compared across single cells in culture and tissue, only four tissue samples were used in order to match group sample sizes. Tissue samples with highest percentage alignment were used for these comparisons. Pile-up diagrams were prepared using Integrative Genomics Viewer version 2.1 using mm9 as the reference genome.

G+C content and transcript length. To determine the relationship between read counts generated from TIVA-collected samples

doi:10.1038/nmeth.2804

and transcript sequence characteristics such as G+C content or length, we examined the correlation between these gene traits and normalized gene counts. We tabulated gene length and G+C content for UCSC genome browser known genes and calculated the correlations of these traits with average read depth across four different experimental groups: pipette-collected culture samples ($n = 7$), TIVA-collected culture samples ($n = 8$), pipette-collected tissue samples ($n = 3$) and TIVA-collected tissue samples ($n = 7$). Only genes with an average read depth of 10 or more for a particular experimental group were included. Evenness of coverage: to assess the effect of TIVA collection on evenness of coverage, we employed two methods used in ref. 30. First, we assessed evenness of coverage across the length of a transcript, from 5′ to 3′. To do this we selected highly expressed genes, retaining those with greater than 500 uniquely aligned reads. We divided each gene into 100 equally sized bins by length and then calculated the relative read depth observed for each bin. Finally, for each sample, we examined the average relative coverage across all transcripts. Second, as a global measure of evenness of coverage we calculated the average coefficient of variation in per-nucleotide coverage across the 1,000 most highly expressed transcripts for each sample. For all coverage analysis, we considered only reads where both mates aligned uniquely to genes on autosomal chromosomes. Transcripts shorter than 100 nucleotides in length were excluded from analysis.

To measure the extent of coverage at the 5′ end of expressed genes, we calculated the fraction of genes demonstrating read coverage within the 250 most 5′ bases. We limited our attention to expressed genes, defined here as those with at least ten read counts.

38. Turner, J.J. *et al.* Cell-penetrating peptide conjugates of peptide nucleic acids (PNA) as inhibitors of HIV-1 Tat-dependent trans-activation in cells. *Nucleic Acids Res.* **33**, 6837–6849 (2005).
39. Richards, J.L., Tang, X., Turetsky, A. & Dmochowski, I.J. RNA bandages for photoregulating *in vitro* protein synthesis. *Bioorg. Med. Chem. Lett.* **18**, 6255–6258 (2008).
40. Cummings, D.D., Wilcox, K.S. & Dichter, M.A. Calcium-dependent paired-pulse facilitation of miniature EPSC frequency accompanies depression of EPSCs at hippocampal synapses in culture. *J. Neurosci.* **16**, 5312–5323 (1996).
41. Grant, G.R. *et al.* Comparative analysis of RNA-Seq alignment algorithms and the RNA-Seq unified mapper (RUM). *Bioinformatics* **27**, 2518–2528 (2011).
42. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).