A convenient single-cell newly synthesized transcriptome assay reveals gene expression dynamics during early-stage T-cell activation

Jun Lyu[1], Xiaoyan Xu[2], Chongyi Chen[1]*

1. Laboratory of Biochemistry and Molecular Biology, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA.
2. Immunoregulation Section, National Eye Institute, National Institutes of Health, Bethesda, MD 20892, USA.

* Correspondence should be addressed to chongyi.chen@nih.gov

Abstract

Sequencing newly synthesized transcriptome alongside regular transcriptome in single cells enables the study of gene expression temporal dynamics during rapid chromatin and gene regulation processes. However, existing assays to profile single-cell newly synthesized transcriptome require in-house technical expertise to achieve high cellular throughput, limiting their widespread application. Here, we developed NOTE-seq, a method that simultaneously profiles regular and newly synthesized transcriptomes in single cells. NOTE-seq integrates 4-thiouridine labeling of newly synthesized RNA, thiol-alkylation-based chemical conversion, and a streamlined workflow on the 10X Genomics platform, offering high cellular throughput that is accessible and convenient for regular biology laboratories without specialized single-cell expertise. Using NOTE-seq, we characterized the temporal dynamics of gene expression during early-stage T-cell activation in Jurkat and naïve T cells, identified transcription factors and regulons, and discovered *Fli-1* as a master transcription factor for gene regulation upon T-cell stimulation. Interestingly, chemotherapeutic topoisomerase inhibitor affects *Fli-1* level in T cells, indicating potential complications for the immune system.

Introduction

Gene expression is a dynamic process involving the synthesis of new RNA transcripts and the degradation of old ones. Regular RNA sequencing methods cannot distinguish newly synthesized RNA from pre-existing RNA, thus unable to dissect transcription kinetics during biological processes involving rapid temporal changes in chromatin regulation and gene expression. Computational approaches, such as RNA velocity analysis, estimate RNA kinetics from existing RNA sequencing datasets by distinguishing unspliced and spliced mRNA molecules[1, 2]. However, these methods were based on certain assumptions that may lead to potential pitfalls[3]. Experimental approaches, which involve the metabolic labeling of newly synthesized RNA using nucleoside analogs followed by biochemical isolation or chemical conversion based on the incorporated nucleoside analogs, offer the capability to distinguish newly synthesized RNA from pre-existing RNA. This distinction enables accurate quantification of RNA synthesis and degradation rates[4]. Building on this strategy, several methods to detect and sequence newly synthesized RNA have been developed and applied to cultured cells in different settings, advancing our understanding of the regulation of transcription and RNA kinetics[5-9].

More recently, the strategy of metabolic labeling of newly synthesized RNA has been integrated into single-cell RNA sequencing[10], enabling the characterization of RNA kinetics with higher sensitivity, allowing for the detection of subtle temporal alterations in the transcriptome of single cells, and providing higher resolution in cell clustering and cell-fate trajectory analysis[11-18]. However, the application of single-cell newly synthesized transcriptome sequencing to address scientific questions is not yet widespread, largely due to the technical limitations of existing methods. On one hand, plate-based assays, including scSLAM-seq[11], NACS-seq[12], and scEU-seq[13], have limited cellular throughput and potential batch-to-batch variations between individual plates. On the other hand, among methods offering high cellular throughput, sci-fate suffers from a low cell recovery rate of less than 5% due to the split-pool procedures[14], while scNT-seq, SLAM-Drop-seq, and Well-TEMP-seq require a specialized microfluidic system or customized micro-well chips [15-18], significantly limiting their accessibility to regular biology laboratories without in-house single-cell technology expertise. In contrast, the 10X Genomics platform is widely available in core facilities that are often accessible to regular biology laboratories. This platform offers high cellular throughput and excellent performance in single-cell and single-molecule RNA capture for transcriptome analysis, and is convenient for regular laboratories to use due to its streamlined protocol and automation. However, it is currently unable to probe newly synthesized RNA molecules in single cells.

In this study, we developed the **N**ewly synthesized transcriptome **o**n **10**X **E**xpression sequencing (NOTE-seq) assay, which combines newly synthesized RNA detection with the 10X Genomics platform for the simultaneous profiling of single-cell transcriptome and newly synthesized transcriptome. NOTE-seq retains the streamlined workflow and data analysis pipeline of the 10X Genomics platform, making it a convenient assay accessible to regular biology laboratories without extensive experience and specialized expertise in single-cell technology. Using NOTE-seq, we characterized the temporal dynamics of gene expression during early-stage T-cell activation, revealing patterns of transcription activity and gene regulation by transcription factors at different time points following T-cell stimulation. We identified *Fli-1* as a master transcription factor with many downstream targeted genes in its regulon, playing an

important role in gene expression dynamics upon T-cell stimulation. Additionally, we found that *Fli-1* levels in naïve T cells were affected by treatment with a human topoisomerase inhibitor used in cancer chemotherapy, indicating potential complications for the immune system caused by topoisomerase inhibition.

Result

NOTE-seq enables simultaneous profiling of transcriptome and newly synthesized transcriptome in single cells

The workflow of the NOTE-seq assay included metabolic labeling of the newly synthesized RNA, specific alkylation of the labeled nucleoside, and a modified protocol to profile single-cell transcriptome on the 10X Genomics platform (Fig. 1a). Specifically, after incubation with the thymidine analog 4-thiouridine (4sU) to label newly synthesized RNA, cells were formaldehyde crosslinked to immobilize RNA molecules in situ and iodoacetamide (IAA) treated to achieve thiol-specific alkylation of the incorporated 4sU, leading to the conversion from the thymidine analog to a cytosine analog, thus an effective T-to-C substitution in downstream sequencing analysis. Single-cell transcriptome profiling was then performed on the 10X Genomics platform using a modified protocol, including reverse transcription (RT) with cellular barcoding, reverse crosslinking, and a second round of RT to enhance template-switching efficiency before cDNA amplification, followed by library preparation and sequencing. Generally speaking, sequencing reads containing T-to-C substitution represented the newly synthesized RNA, while those without T-to-C substitution represented the pre-existing RNA. However, considering potential confounders arising from the relatively infrequent events of 4sU incorporation and sequencing errors, the exact fractions of newly synthesized RNA and pre-existing RNA were estimated based on a binomial mixture statistical model[19].

To characterize the performance of NOTE-seq, we carried out both NOTE-seq and regular single-cell transcriptome sequencing on the same 10X Genomics platform using the same human cell line eHAP1, and compared their parameters of single-cell transcriptome profiling. Comparable library size distributions (Fig. S1a) and similar fractions of mitochondrial and ribosomal genes (Fig. S1b) were observed, suggesting negligible RNA leakage in the NOTE-seq workflow involving additional steps of 4sU labeling and subsequent thiol-alkylation-based chemical conversion. Moreover, single-cell transcriptomes characterized by both methods were highly correlated (Fig. 1b), suggesting the additional steps and modified workflow in NOTE-seq did not significantly affect cell physiology, alter gene expression pattern, or skew the RNA composition.

We further determined NOTE-seq's specificity in labeling and identifying newly synthesized RNA. First, in the NOTE-seq data, T-to-C conversion rate was more than 10-fold higher than other types of conversion events (Fig. 1c), a strong indicator of successful 4sU incorporation and specific chemical conversion. Next, T-to-C events were mostly detected in intronic regions, with a 10-fold higher frequency than exonic T-to-C events (Fig. 1d), suggesting specific labeling of newly synthesized nascent RNA molecules instead of

pre-existing RNA. Moreover, for the housekeeping gene *GAPDH* with a low transcript turnover rate, indeed only a small fraction of its RNA was newly synthesized, that was opposite to the *MYC* gene with a high transcript turnover rate and a dominant fraction of newly synthesized RNA (Fig. 1e), further confirming the specificity of NOTE-seq in detecting newly synthesized RNA. Finally, we analyzed single cells in different cell-cycle stages in the population. After confirming that NOTE-seq does not significantly alter the cell-cycle distribution (Fig. S2a), we focused on cell cycle markers in single cells within each stage. Consistent with the long mRNA half-life in human cells[20], we found that the detection of newly synthesized RNA often preceded the emergence of regular RNA during cell cycling (Fig. S2b), again demonstrating NOTE-seq's capability and specificity to probe newly synthesized RNA in single cells.

As the cost of detecting newly synthesized transcriptome in addition to the regular transcriptome in single cells, NOTE-seq had some drawbacks compared to regular single-cell transcriptome sequencing on the 10X Genomics platform without distinguishing between newly synthesized and pre-existing RNA. First, cDNA length and the mapping ratio of sequencing reads were often compromised (Fig. S1c-d), likely due to the reduced processivity of the reverse transcriptase due to formaldehyde crosslinking and the chemical conversion of 4sU-incorporated RNA bases[15]. Consistently, NOTE-seq also showed a slight reduction in the detection sensitivity estimated by the number of genes and UMIs detected in single cells (Fig. 1g-h). In addition, NOTE-seq reported an elevated collision rate between single cells (Fig. 1f), yet still within the reasonable range, that was commonly observed in single-cell transcriptome assays involving formaldehyde crosslinking[21].

As a conclusion, we developed the NOTE-seq assay that enables simultaneous single-cell profiling of transcriptome and newly synthesized transcriptome. The slightly compromised gene detection sensitivity and elevated collision rate of the NOTE-seq assay compared to regular single-cell transcriptome analysis were largely tolerable in most scientific applications. More importantly, relying on the widely available 10X Genomics platform to offer a high cellular throughput, NOTE-seq was a feasible and convenient method with an easy access for regular biology labs. As a result, NOTE-seq was particularly suitable for labs without much experience in single-cell technologies and lacking expertise in microfluidic automation, enabling them to conveniently characterize single-cell transcriptome and newly synthesized transcriptome, which is crucial to understand the temporal dynamics of gene expression in single cells during processes involving rapid chromatin and gene regulation.

NOTE-seq reveals gene expression temporal dynamics during early-stage T-cell activation

Using NOTE-seq, we studied T lymphocytes, a vital component of the immune system exerting central functions against infectious diseases and tumors. In particular, we investigated the rapid dynamics of gene expression during early-stage T-cell activation, a prerequisite for T-cell functions including T-cell development, proliferation, and differentiation[22]. In vivo, T cells can be activated via rapid stimulation of the canonical T-cell activation signaling pathway upon cognate antigen recognition. The human Jurkat T cells provide a convenient and robust in vitro system to study the temporal dynamics of gene expression during T-cell activation, after its readily stimulation by phorbol 12-myristate 13-acetate (PMA) and ionomycin (ION) to directly activate the downstream of the T-cell activation signaling pathway (Fig. 2a).

We applied NOTE-seq to characterize the temporal dynamics of gene expression during early-stage Jurkat T-cell activation (Fig. 2a). We found that the T-to-C conversion rate, the detection rate of newly synthesized RNA, the fraction of newly synthesized RNA per gene, and the sequencing depth per cell all showed a good similarity between conditions of stimulated T-cell activation and unstimulated negative control (Fig. S3a-d), suggesting the successful implementation of NOTE-seq to study the gene expression dynamics during T-cell activation.

First, we performed cell clustering analysis. Principal component analysis (PCA) based on the single-cell transcriptome grouped individual cells into two partially overlapped clusters, indicating a lack of clear separation between stimulated and unstimulated cells (Fig. 2b, left panel). Next, we performed PCA based on the newly synthesized transcriptome revealed by NOTE-seq and successfully identified two clearly separated clusters, representing the stimulated and unstimulated T cells, respectively (Fig. 2b, middle panel). The better clustering performance could be attributed to the possibility that newly synthesized transcriptome in single cells, compared to regular transcriptome, more accurately reflects the difference between stimulated and unstimulated cells. We further reasoned that, considering the rapid gene expression dynamics upon T-cell activation and the relatively long RNA lifetime, the dominant fraction of the single-cell RNA could be pre-existing RNA transcribed before T-cell stimulation. Indeed, we found that PCA based on the pre-existing transcriptome generated two clusters that were mostly overlapped and indistinguishable (Fig. 2b, right panel). These results highlighted the importance of detecting the newly synthesized transcriptome by NOTE-seq, that was often missing in regular single-cell transcriptome assays, especially when studying dynamic processes involving rapid gene expression changes that could determine cell type and cell fate.

Next, we performed gene regulatory network analysis based on the temporal fluctuations of the transcriptome in single cells[23]. Specifically, after quantifying the transcription level of each gene that fluctuates in individual cells within the cell population, gene regulatory networks (a.k.a. regulons) were inferred by correlating the expression level of transcription factors (TFs) and the transcription level of their corresponding downstream targeted genes in single cells (Fig. 2c). Based on the NOTE-seq data, we were able to identify multiple regulons representing the major T-cell activation signaling pathway in stimulated T cells (Fig. 2d-e), as expected from the previous knowledge[22].

Finally, we investigated differentially expressed genes (DEGs) during early-stage T-cell activation (Fig. 2f). Not surprisingly, many canonical T-cell activation makers such as *CD69*[24] were upregulated at the RNA level. Moreover, we often noted a more significant increase in the level of newly synthesized RNA compared to the level of regular RNA (Fig. 2f), suggesting a superior DEG detection sensitivity by probing the newly synthesized RNA in NOTE-seq. Indeed, based on the level of newly synthesized RNA, we identified additional DEGs that otherwise would go unnoticed on the regular RNA level (Fig. 2f). Focusing on differentially expressed TFs during T-cell activation (Fig. 2f, Table S1), we found that most TFs were upregulated with reported relevance in T-cell function (Table S1), including families of *NR4A*, *EGR*, and *NFκB*[25, 26], consistent with the expectation. However, we also identified a downregulated TF, *FLI-1*, whose exact role during early-stage T-cell activation remains unclear.

NOTE-seq identifies transcription factors for gene regulation during early-stage T-cell activation

The circumstances during T-cell activation in vivo may not be fully recapitulated by Jurkat T-cell activation in vitro. For example, naïve T cells in vivo are often in quiescence, characterized by a small cell size, a low metabolism rate, and a low transcription activity[27], that is very different from the proliferating Jurkat T cells (Fig. 3a). Consequently, quiescence exit is an essential feature of naïve T-cell activation, together with other potential differences between naïve and Jurkat T cells, that could lead to a different gene expression program regulated by a different set of TFs.

We applied NOTE-seq to characterize the gene expression dynamics during early-stage naïve T-cell activation ex vivo. Specifically, we isolated mouse naïve T cells for stimulation using agonistic antibodies and collected cells in a two-hour interval with one hour labeling of newly synthesized RNA prior to each harvest (Fig. 3b). The expected naïve T-cell activation and quiescence exit were observed upon stimulation, as evidenced by the upregulation of multiple relevant makers (Fig. S4). The expected dose-dependence of T-to-C conversion frequency was also observed (Fig. S5a), suggesting successful 4sU labeling of newly synthesized RNA without triggering detrimental stresses or altering the gene expression program in T cells (Fig. S5b-c).

First, we performed cell clustering analysis to study the gene expression dynamics during naïve T-cell activation. Based on single-cell transcriptome, we found that cells without stimulation, two-hour post-stimulation, and four-hour or longer post-stimulation were already well separated into distinct clusters (Fig. 3c), suggesting significant temporal alterations of the gene expression program during T-cell activation within the first four hours. Additional clustering based on both the transcriptome and the newly synthesized transcriptome offered a better resolution, and further separated cells stimulated for four, six, and eight hours into distinct clusters (Fig. 3c), suggesting minor yet distinguishable gene expression changes after four hours post-stimulation. Furthermore, the first principal component of the single-cell joint transcriptome demonstrated a temporal trend following the initial stimulation (Fig. 3d), again demonstrating a dynamic program of gene expression progressing over time during T-cell activation.

Next, we analyzed differentially expressed TFs over time during naïve T-cell activation, and found different TFs at different time points (Fig. 3e, Table S2), revealing the temporal dynamics of the gene expression program after the initial stimulation. For many TFs, the change in regular RNA level was often preceded by the change in newly synthesized RNA level (Fig. 3f), that reflects the actual timing of the transcription response to T-cell stimulation. In addition, some TFs showed similar trends between the levels of newly synthesized RNA and regular RNA at all time points (Fig. 3f), indicating a high transcript turnover rate for these TFs, a key feature for important gene regulators capable of fast modulation and acute response to the environmental stimuli.

Most upregulated TFs identified during naïve T-cell activation have reported roles relevant to T-cell function (Table S2). Many of them also coincided with TFs observed during Jurkat T-cell activation, including families of *Nr4a*, *Egr*, and *Nfκb*, that have known involvement in T-cell activation. The differences of a few TFs were likely due to different stimulation conditions, sampling time points, and different T cells. Furthermore, we identified various downregulated TFs specifically during the activation

of naïve T cells but not Jurkat T cells (Fig. 3e). Many of them were involved in the maintenance of the quiescence state (Table S2), consistent with the additional essential step of quiescence exit during naïve T-cell activation.

Interestingly, *Fli-1* was also downregulated during naïve T-cell activation (Fig. 3e), the same as our previous observation during Jurkat T-cell activation, suggesting *Fli-1* as a newly identified TF with potential importance. In addition, *Fli-1* was downregulated rapidly at the start of T-cell activation, with a fast turnover rate since both the levels of its RNA and newly synthesized RNA rapidly respond to T-cell stimulation with the same trend (Fig. 3f), further indicating its role in gene regulation upon T-cell stimulation.

Fli-1 **is a master transcription factor during early-stage T-cell activation modulated by topoisomerase inhibition**

To investigate the role of *Fli-1* as a down-regulated TF during early-stage T-cell activation, we performed gene regulatory network analysis to identify master TFs that regulate a large number of downstream targeted genes and cause significant changes in their transcription activity, thus acting as driving factors to modulate the gene expression program. Among the master TFs (Fig. 4a), the upregulated *Nfκb* is associated with the T-cell activation signaling pathway to promote cell proliferation (Fig. S6)[22, 28], and the downregulated *Klf3* and *Stat1* have reported roles in the maintenance of T cell naivety (Fig. S6)[29, 30], with their downregulation expected for the quiescence exit during naïve T-cell activation. More importantly, *Fli-1* was also identified as a master TF, regulating 59 downstream targeted genes in the Fli-1 regulon (Fig. 4a). Most genes in the *Fli-1* regulon were significantly downregulated over time (Fig. 4b), making *Fli-1* the most downregulated regulon among all the identified regulons during T-cell activation (Fig. 4c). Furthermore, gene ontology analysis showed that many genes in the *Fli-1* regulon were categorized in functional groups related to the regulation of T-cell differentiation and proliferation (Fig. 4d), highlighting the importance of *Fli-1* as a master TF to regulate the temporal dynamics of gene expression during T-cell activation.

On the other hand, *Fli-1* is a proto-oncogene[31] overexpressed in various types of human malignancies[32], most notably in Ewing sarcoma, which is induced by *Fli-1* translocation leading to the expression of the oncogenic fusion protein EWS/FLI-1[32, 33]. To treat malignancies associated with FLI-1 overexpression, small-molecule FLI-1 inhibitors, including camptothecin (CPT) and etoposide[32, 34], have been identified as effective treatments for leukemia and Ewing sarcoma[34]. Interestingly, both CPT and etoposide are topoisomerase inhibitors widely used in cancer chemotherapy. Specifically, CPT and its derivatives targeting human topoisomerase I are used for the treatment of colon cancer, small-cell lung cancer, and ovarian cancer[35], while etoposide targeting human topoisomerase II remains the first-line treatment for small-cell lung cancer[36, 37].

We reasoned that the chemotherapeutic topoisomerase inhibitors may also inhibit FLI-1 in naïve T cells. Indeed, we found that CPT treatment induced a dose-dependent reduction of the FLI-1 level in mouse naïve T cells (Fig. 4e), that cannot be completely explained by the reduced T-cell viability (Fig. S7). These results indicated that topoisomerase inhibitors used in cancer chemotherapy, in addition to the expected

inhibition of topoisomerase-dependent DNA replication and cell division in tumor cells that proliferate much faster than normal cells, may have other complicated effects on the immune system. In particular, topoisomerase inhibition may affect the regulation of T-cell activation by triggering the downregulation of FLI-1, a master TF involved in the temporal dynamics of gene expression during early-stage T-cell activation.

Discussion

Single-cell newly synthesized RNA sequencing, in addition to regular single-cell RNA sequencing, enables the simultaneous profiling of newly synthesized and regular transcriptomes in single cells, facilitating the study of the temporal dynamics of gene expression and gene regulation. Based on the 10X Genomics platform, we developed the NOTE-seq assay, which offers simultaneous quantification of both newly synthesized and pre-existing transcriptomes in single cells. In addition, NOTE-seq provides high cellular throughput, is easily accessible to many biology laboratories, and can be conveniently performed without requiring in-house expertise in single-cell technology. The NOTE-seq workflow is essentially a modified version of the streamlined and optimized workflow of the 10X platform, enhancing its capability to profile the newly synthesized single-cell transcriptome in addition to the regular single-cell transcriptome already offered by the standard 10X workflow.

Using NOTE-seq to characterize the gene expression temporal dynamics upon T-cell stimulation, we identified TFs and regulons, particularly discovering a new master TF during early-stage T-cell activation. However, the application of NOTE-seq extends far beyond the process studied in this work. In principle, NOTE-seq is applicable to investigating any rapid dynamic process involving temporal changes in chromatin and gene regulation at the single-cell level, as newly synthesized RNA accurately reflects transcription kinetics, which are often delayed and hidden in regular RNA analysis. Additionally, NOTE-seq quantifies the abundance of both pre-existing and newly synthesized RNA in single cells with high accuracy, that can be used to estimate important parameters of RNA synthesis and degradation rates, enabling further study of transcription and RNA kinetics.

It is intriguing that chemotherapeutic topoisomerase inhibitors, including CPT and etoposide, which specifically inhibit human topoisomerase I and II respectively, also reduce the level of FLI-1 in naïve T cells. One possibility is that CPT and etoposide directly target FLI-1 via a topoisomerase-independent mechanism. However, considering that two different types of small-molecule inhibitors targeting different topoisomerases have the same effect on FLI-1, it is more likely that FLI-1 downregulation upon CPT or etoposide treatment is due to the inhibition of topoisomerase activities, that modulate DNA supercoiling tension in the chromatin to further affect TF binding, alter RNA polymerase activity, and ultimately regulate gene expression. This is consistent with the observation that both FLI-1 protein and *Fli-1* mRNA levels are reduced upon topoisomerase inhibition[32], suggesting the downregulation occurs at the transcription level. Moreover, a recent report showed the involvement of R-loop alterations during CPT-induced Fli-1 downregulation[38], further indicating its relevance to topoisomerase and DNA supercoiling, which may alter the chromatin environment. Future studies will be needed to elucidate the mechanism of

Fli-1 downregulation upon topoisomerase inhibition and its connection to chromatin dynamics and DNA supercoiling.

In addition to the role in early-stage T-cell activation, *Fli-1* is implicated in regulating T-cell development[32] and the response of effector T cells[39], underscoring the potential broad impact of *Fli-1* downregulation on the immune system following topoisomerase inhibition. Indeed, various studies across different contexts suggest that topoisomerase activities have a role in immune diseases[40-42], and that topoisomerase inhibition could potentially modulate the immune response during inflammation[43, 44]. Further research is essential to deepen our understanding of the roles of Fli-1 and topoisomerase inhibition in T-cell function and the overall immune response, including exploring different T-cell subsets, various stages of T-cell activation, and other components of the immune system.

Method

Cell culture

The eHAP human cells (Horizon discovery Ltd, Cat # C669) were cultured in IMDM medium (Gibco) containing 10% FBS and 1% PS, and supplemented with 1 mM 4sU one hour for RNA labeling before harvest. Jurkat T cells (E6.1 clone) were cultured in RPMI1640 medium (Gibco) containing 10% FBS and 1% PS. For RNA labeling and cell stimulation, one million Jurkat T cells were resuspended in RPMI1640 containing 250 μM 4sU, 81 nM phorbol 12- myristate 13-acetate (PMA) and 1.34 μM ionomycin (eBioscience™ Cell Stimulation Cocktail) for one-hour incubation. Mouse naïve T cells were purified from splenocytes and lymphocytes of C57BL/6J WT mice using EasyStem™ Mouse Pan-Naïve T cell isolation kit (STEMCELL), cultured in RPMI1640 medium containing L-glutamine, 50 μM 2-mercaptoethanol, 10% FBS, and 1% PS. For RNA labeling and cell stimulation, 0.5-1 million cells were placed in a 24-well plate coated with 5 ng/μl plate-coated anti-CD3 (Biolegend, clone 2C11) and anti-CD28 (Biolegend, clone 37.51) for incubation for the time indicated, and followed by 500 μM 4sU addition one hour before harvest.

The NOTE-seq assay

One million 4sU-labeled cells were harvested, resuspended in 500 μl ice-cold PBS, fixed by adding 500 μl 8% formaldehyde, and incubated on ice for 15 minutes. After washing with PSD (0.5% Superase•In and 10 mM DTT in PBS), cells were resuspended with PSD containing 0.2% Triton X-100, incubated on ice for 5 minutes, washed with PS (0.5% Superase•In in PBS), and resuspended in 100 μl PS. Then, 40 μl fresh-made iodoacetamide (IAA), 40 μl 500 mM Sodium phosphate buffer (pH 8.0), and 20 μl nuclease-free water were sequentially added to the 100 μl cell suspension, before mixing with 200 μl DMSO. After 15-minute incubation at 50°C on a mixer, cells were put on ice with immediate addition of 8 μl 1 M DTT to quench the chemical conversion. Cells were then washed with 5 ml ice-cold PBS containing 0.2% Triton X-100 and resuspended in 200 μl PS. Cell suspension was loaded on the 10X Genomics platform, following the workflow according to the Chromium Next GEM Single Cell 3′ Kit v3.1 user guide (Rev E) with the following modifications: (1) After the GEM break (step 2.1a), the aqueous phase was transferred

to a new PCR tube, with the addition of 2 µl 20 mg/ml protease K (Invitrogen) and one-hour incubation at 53°C on a mixer. (2) After Dynabeads Cleanup (step2.1n), beads were resuspended with 10.5 µl MixtureA consisting of 50 mM Tris-HCl (pH 8.0), 60 mM NaCl, 5 mM $MgCl_2$, and 1 mM dNTP. 10 µl elution was mixed with 10 µl MixtureB consisting of 2 mM GTP, 10% PEG8000, 16 mM DTT, 2U/µl Superase•In, 4 µM TSO, and 200 U Maxima H-minus, and incubated at 53°C for 45 minutes (second round of reverse transcription). (3) cDNA was amplified by mixing 50 µl Amp Mix, 15 µl cDNA primers, and 15 µl nuclease-free water, following the PCR program and cycles in the user guide (step2.2d). (4) The PCR products were purified twice, following step2.3. The downstream library preparation and sequencing was performed the same as the user guide.

The barnyard experiment

Human Jurkart T cells and NIH3T3 cells were mixed at a 1:1 ratio, followed by the NOTE-seq workflow without T-to-C conversion. Reads were aligned to GRCh38_and_mm10-2020-A reference by cellranger7.1 with cellular barcodes assigned to human, mouse, or multiplets. Briefly, cellular barcodes are initially classified as human or mouse by which genome has more total UMI counts for that barcode. Barcodes with total UMI counts that exceed the 10th percentile of the distributions for both human and mouse are called observed multiplets. Because cellranger only detects the observable multiplets (human, mouse), it computes an inferred multiplet rate by estimating the total number of multiplets, including (human, human) and (mouse, mouse). This is done by estimating via maximum likelihood the total number of multiplet GEMs from the observed multiplets and the inferred ratio of human to mouse cells.

The scRNA-seq data analysis

Raw sequencing data was subjected to fastq calling, reads alignment, and quantification using cellranger7.1 with default parameters and standard pipeline. The gene expression matrix from the filtered_feature_bc_matrix was used to create a Seurat object using Seurat v4.3.0 with default parameters. Cellular barcodes (CBs) were retrieved from the Seurat object, and were used to subset the BAM file. Subsequently, confidently unique-aligned reads representing UMIs (corresponding to BAM tag: xf:i:25) were extracted from reads associated with the selected CBs. A modified pipeline of scNT-seq was adopted to quantify metabolically labeled and unlabeled transcripts. Briefly, Sites of T-to-C substitution with the base Phred Quality Score > 27 (A to G substitution in the reverse complementary strand) in each transcript were extracted and then filtered by background control without 4sU labeling and chemical conversion. The resultant transcript containing T-to-C substitution was counted as newly synthesized transcripts and then corrected by a statistical model described below.

Estimation of the fraction of newly synthesized transcripts

The metabolic labeling of transcripts was incomplete due to insufficient 4sU substitution, resulting in underestimation of labeled transcripts. To remedy such quantification bias, a binomial mixture model was adopted to approximate the real distribution of T-to-C substitution and the real mutation rate. Briefly, the likelihood function for each gene transcript i follows:

$$f(\theta, p, q) = \theta Binom(y_i; p, n_i) + (1 - \theta) Binom(y_i; q, n_i)$$

where θ is the fraction of new transcripts in each experiment, p and q are the probabilities of a T-to-C substitution at each nucleotide for new and old transcripts, respectively, and n_i is the number of uridine nucleotides observed in the transcript i. As the p and q are global parameters (not gene-specific) for all transcripts, for each sample, 10,000 UMIs were randomly sampled to estimate the global p and q. The model was fit by maximizing the likelihood function using the Nelder–Mead algorithm. The optimization was repeated 100 times with random initialization values for θ, p, and q in the range [0,1]. The p and q were kept if θ is in the range [0,1]. The p and q corresponding to the maximal function value were used for calculating θ of individual genes with >= 100 UMIs, using the Brent algorithm with the constraint $\theta \in [10^{-8},1]$. The labeled transcript levels of these genes (N_{gene}) were re-estimated as follows:

$$N_{gene} = \theta(L_{gene} + U_{gene})$$

where θ is the fraction of newly transcribed transcripts for a gene, L_{gene} is the number of labeled transcripts of a gene and U_{gene} is the number of unlabeled transcripts of a gene.

Since the detection rates of labeled transcript (α) are highly similar in one cell, the average α was calculated for all genes in one cell as follows:

$$\alpha_{cell} = \frac{\sum L_{gene}}{\sum N_{gene}}$$

where L_{gene} is the number of observed labeled transcripts of a gene with estimated N_{gene} in that cell.

Then, the new transcript level for each gene in each cell can be re-calculated as follows:

$$N_{gene} = \min\left(\frac{L_{gene}}{\alpha_{cell}}, L_{gene} + U_{gene}\right)$$

The statistic correction pipeline was adapted from Well-TEMP-seq[16].

Cell cycle analysis

The normalized Seurat objects containing single-cell expression matrices were subjected to CellCycleScoring() function of Seurat (v4.3.0), where individual cells were scored, based on updated gene sets of cell cycle markers (cc.genes.updated.2019), and then classified into distinct cell cycle phases. The average expression levels of canonical cell cycle markers were visualized using heatmap[45].

SLAM-seq on naïve T cells

One million naïve T cells were incubated with varying concentrations of 4sU and stimulated by 5 ng/μl plate-coated anti-CD3 and anti-CD28 for 2 hours before being harvested for RNA extraction following the manual of RNeasy Mini Kit (Qiagen). Next, RNA was subject to SLAM-seq protocol with minor modifications. Briefly, 2 μg RNA were combined with 10 mM iodoacetamide (Sigma-Aldrich, St. Louis, MO, USA), 50 mM sodium phosphate buffer (pH 8.0) and 50% DMSO, followed by 15 min of incubation at 50°C. The reaction was stopped by adding 20 mM DTT, followed by RNA purification using 1.8X RNA beads (Aline Bioscience) and RNA elution with 11 μl nuclease-free water. Then, 500 ng RNA was subject to library construction using Quant-seq 3′ mRNA-Seq Library Prep Kit (Lexogen). Data were first

trimmed using cutadpt (v1.15) and then analyzed using SLAM-DUNK (v0.4.3)[46] to map, filter and quantify reads containing T-to-C mutation.

PCA analysis, DEG identification and GO enrichment

Seurat objects were generated with new, old, and total transcript matrices. Functions in the Seurat package including RunPCA and FindMarkers were used to perform principal component analysis (PCA) and identify differentially expressed genes (DEGs) using Wilcoxon test, respectively. Genes with absolute fold changes in expression > 1.5 and adjusted p-value < 0.05 were selected as DEGs. The list of transcription factors to start with was according to previous studies[47, 48]. Gene Oncology analysis and KEGG pathway analysis were performed by clusterProfiler[49].

Gene regulatory network analysis

Regulon activity analysis was performed using single-cell regulatory network inference and clustering (SCENIC). Briefly, the newly synthesized transcript matrix and total transcript matrix were filtered with default parameters of geneFiltering, and the correlation matrix was computed based on gene expression level by runCorrelation. The correlation matrix of TFs and their target gene were computed and weighed by GENIE3, and then correlation modules were validated (with RcisTarget), selected (with weight > 0.005), and scored using runSCENIC. The resultant modules (regulons) with highly confident annotation for motifs of target genes were used for downstream analysis.

Flow cytometry

Flow cytometry was performed in Beckman CytoFlex LX flow Cytometer, and data were analyzed with Flowjo v10 software (TreeStar). The following antibodies, conjugated with designated fluorochromes were used: anti-mouse CD69 (H1.2F3), CD62L (MEL-14), CD4 (RM4-5), CD8 (53-5.8), CD90.2 (53-2.1), CD44 (1M7), anti-phospho AKT (S473), anti-phospho-S6 (Ser235/236). Viakrome 808 Fixable Viability dye (Beckman Coulter) was used for viability staining for 20 min at room temperature. For cell surface staining, cells were stained with different antibody cocktails in a staining buffer (PBS with 5% FCS, 2 mM EDTA, 1ug/ml Fc receptor blocking antibody) for 30 min on ice, followed by viability staining and fixation with PFA. For intracellular staining, cells were fixed and permeabilized with Cytofix/CytopermTM solution (BD), followed by staining against intracellular antigens. For Phosflow staining, cells were fixed with BD Phosflow Fix Buffer for 10 min and permeabilized in BD Phosflow Perm buffer for 30 min on ice. Then cells were stained with indicated antibodies in staining buffers for 1 hours on ice.

Western blot

Naïve T cells were treated with different concentrations of CPT for 24 hours. Cells were lysed in RIPA buffer supplemented with protease inhibitors for protein extraction. The equal volume of protein samples was loaded to and separated by 10% SDS-PAGE and then transferred onto an NC membrane (Millipore). The membrane was blocked in 5% non-fat milk dissolved in Tris-buffered saline with Tween-20 (TBST) for 1 hour and then washed with TBST three times. Subsequently, the membrane was incubated with rabbit

anti-mouse primary antibodies against FLI1 (Abcam, ab133485, 1:500 dilution) or beta-actin (Abcam, ab272085, 1:1000 dilution) overnight at 4°C and horseradish peroxidase-labeled secondary antibody (Abcam, catalog no. ab97080, 1:5000 dilution) at room temperature for 1 hour. The protein signals were detected by ECL Kit (Thermo Scientific, 34580) on Cytiva800.

Cell viability measurement

The naïve T cells in triplicate wells were treated with different concentrations of CPT for 24 hours. Then, 10 µl single cell suspension was mixed with 10 µl trypan blue in a PCR tube. The mixture was loaded into a chamber slide and subjected to Countess II for the number and viability quantification.

Sequencing

Libraries were sequenced on an Illumina NextSeq 2000 instrument. For NOTE-seq libraries, sequencing parameters follow the Chromium Next GEM Single Cell 3′ Kit v3.1 user guide. For SLAM-seq libraries, 100 cycles-single-end sequencing was performed for read1.

Data availability

All count tables generated by NOTE-seq pipeline are available at https://github.com/lyuj2022/NOTE-seq. The raw sequencing data were deposited at the National Center for Biotechnology Information Gene Expression Omnibus (GEO), with the accession number GSE270292.

Code availability

The main scripts used to analyze data and generate main figures and conclusions are available at https://github.com/lyuj2022/NOTE-seq.

Acknowledgements

We thank Madeline Wong and Steve Shema from the CCR Genomics Core for sequencing; Charlie Seibert from Single Cell Analysis Facility for 10X Genomics training; and other members of the Chen lab for discussions. The Jurkat T cell line is a gift from Dr. Peter D. Aplan. The mouse primary T cells are provided by Dr. Rachel Caspi. This study used the Biowulf Linux cluster at the National Institutes of Health. This work was supported by the Intramural Research Program of the National Institutes of Health, National Cancer Institute.

Author contributions

J.L. and C.C. conceived the project and wrote the manuscript. J.L. developed the assay protocol, modified the data analysis pipeline, performed most of the experiments, analyzed most of the data, and prepared the figures. X.X. performed some experiments, analyzed some data, and provided helpful suggestions. C.C. supervised this project.

Competing interests

The authors declare no competing interests.

Reference:

1. La Manno, G., et al., *RNA velocity of single cells.* Nature, 2018. **560**(7719): p. 494-498.
2. Bergen, V., et al., *Generalizing RNA velocity to transient cell states through dynamical modeling.* Nat Biotechnol, 2020. **38**(12): p. 1408-1414.
3. Bergen, V., et al., *RNA velocity-current challenges and future perspectives.* Mol Syst Biol, 2021. **17**(8): p. e10282.
4. Rabani, M., et al., *Metabolic labeling of RNA uncovers principles of RNA production and degradation dynamics in mammalian cells.* Nat Biotechnol, 2011. **29**(5): p. 436-42.
5. Schwalb, B., et al., *TT-seq maps the human transient transcriptome.* Science, 2016. **352**(6290): p. 1225-8.
6. Herzog, V.A., et al., *Thiol-linked alkylation of RNA to assess expression dynamics.* Nat Methods, 2017. **14**(12): p. 1198-1204.
7. Riml, C., et al., *Osmium-Mediated Transformation of 4-Thiouridine to Cytidine as Key To Study RNA Dynamics by Sequencing.* Angew Chem Int Ed Engl, 2017. **56**(43): p. 13479-13483.
8. Schofield, J.A., et al., *TimeLapse-seq: adding a temporal dimension to RNA sequencing through nucleoside recoding.* Nat Methods, 2018. **15**(3): p. 221-225.
9. Muhar, M., et al., *SLAM-seq defines direct gene-regulatory functions of the BRD4-MYC axis.* Science, 2018. **360**(6390): p. 800-805.
10. Erhard, F., et al., *Time-resolved single-cell RNA-seq using metabolic RNA labelling.* Nature Reviews Methods Primers, 2022. **2**(1).
11. Erhard, F., et al., *scSLAM-seq reveals core features of transcription dynamics in single cells.* Nature, 2019. **571**(7765): p. 419-423.
12. Hendriks, G.J., et al., *NASC-seq monitors RNA synthesis in single cells.* Nat Commun, 2019. **10**(1): p. 3138.
13. Battich, N., et al., *Sequencing metabolically labeled transcripts in single cells reveals mRNA turnover strategies.* Science, 2020. **367**(6482): p. 1151-1156.
14. Cao, J., et al., *Sci-fate characterizes the dynamics of gene expression in single cells.* Nat Biotechnol, 2020. **38**(8): p. 980-988.

15. Qiu, Q., et al., *Massively parallel and time-resolved RNA sequencing in single cells with scNT-seq.* Nat Methods, 2020. **17**(10): p. 991-1001.
16. Lin, S., et al., *Well-TEMP-seq as a microwell-based strategy for massively parallel profiling of single-cell temporal RNA dynamics.* Nature Communications, 2023. **14**(1): p. 1272.
17. Liu, H., et al., *SLAM-Drop-seq reveals mRNA kinetic rates throughout the cell cycle.* Molecular Systems Biology, 2023.
18. Fishman, L., et al., *Cell-type-specific mRNA transcription and degradation kinetics in zebrafish embryogenesis from metabolically labeled single-cell RNA-seq.* Nature Communications, 2024. **15**(1): p. 3104.
19. Jurges, C., L. Dolken, and F. Erhard, *Dissecting newly transcribed and old RNA using GRAND-SLAM.* Bioinformatics, 2018. **34**(13): p. i218-i226.
20. Shamir, M., et al., *SnapShot: Timescales in Cell Biology.* Cell, 2016. **164**(6): p. 1302-1302 e1.
21. Chen, A.F., et al., *NEAT-seq: simultaneous profiling of intra-nuclear proteins, chromatin accessibility and gene expression in single cells.* Nat Methods, 2022. **19**(5): p. 547-553.
22. Hwang, J.-R., et al., *Recent insights of T cell receptor-mediated signaling pathways for T cell activation and development.* Experimental & Molecular Medicine, 2020. **52**(5): p. 750-761.
23. Aibar, S., et al., *SCENIC: single-cell regulatory network inference and clustering.* Nature Methods, 2017. **14**(11): p. 1083-1086.
24. Cibrián, D. and F. Sánchez-Madrid, *CD69: from activation marker to metabolic gatekeeper.* Eur J Immunol, 2017. **47**(6): p. 946-953.
25. Odagiu, L., et al., *Role of the Orphan Nuclear Receptor NR4A Family in T-Cell Biology.* Front Endocrinol (Lausanne), 2020. **11**: p. 624122.
26. Shao, H., et al., *Induction of the early growth response (Egr) family of transcription factors during thymic selection.* J Exp Med, 1997. **185**(4): p. 731-44.
27. Chapman, N.M. and H. Chi, *Hallmarks of T-cell Exit from Quiescence.* Cancer Immunology Research, 2018. **6**(5): p. 502-508.
28. Wang, R., et al., *The transcription factor Myc controls metabolic reprogramming upon T lymphocyte activation.* Immunity, 2011. **35**(6): p. 871-82.
29. Fang, F., et al., *Human Transcription Factor KLF3 Maintains T Lymphocyte Quiescent Phenotype Via Inhibiting SHP-1 Expression.* Blood, 2015. **126**(23): p. 3426-3426.
30. Kye, Y.C., et al., *STAT1 maintains naive CD8(+) T cell quiescence by suppressing the type I IFN-STAT4-mTORC1 signaling axis.* Sci Adv, 2021. **7**(36): p. eabg8764.
31. Ben-David, Y., E.B. Giddens, and A. Bernstein, *Identification and mapping of a common proviral integration site Fli-1 in erythroleukemia cells induced by Friend murine leukemia virus.* Proc Natl Acad Sci U S A, 1990. **87**(4): p. 1332-6.
32. Ben-David, Y., et al., *Current insights into the role of Fli-1 in hematopoiesis and malignant transformation.* Cell Mol Life Sci, 2022. **79**(3): p. 163.
33. Corradini, N. and F. Redini, *New therapeutic targets in Ewing sarcoma: from pre-clinical proof-of-concept to clinical trials.* Bone Cancer: Primary Bone Cancer and Bone Metastases, 2nd Edition, 2015: p. 393-405.

34. Li, Y.J., et al., *Drug-mediated inhibition of Fli-1 for the treatment of leukemia.* Blood Cancer Journal, 2012. **2**(1): p. e54-e54.
35. Liew, S.T. and L.X. Yang, *Design, synthesis and development of novel camptothecin drugs.* Curr Pharm Des, 2008. **14**(11): p. 1078-97.
36. Wang, S., et al., *Current Diagnosis and Management of Small-Cell Lung Cancer.* Mayo Clin Proc, 2019. **94**(8): p. 1599-1622.
37. Uuskula-Reimand, L. and M.D. Wilson, *Untangling the roles of TOP2A and TOP2B in transcription and cancer.* Sci Adv, 2022. **8**(44): p. eadd4920.
38. Joaquin Olmedo-Pelayo, E.G.-C., Daniel Delgado-Bellido, Laura Lobo-Selma, Carmen Jordán-Pérez, Ana T. Monteiro-Amaral, Anna C. Ehlers, Shunya Ohmura, Angel M. Carcaboso, Javier Alonso, Isidro Machado, Antonio Llombart-Bosch, Thomas G.P. Grünewald, Fernando Gómez-Herreros, Enrique de Álava, *EWS::FLI1-DHX9 interaction promotes Ewing sarcoma sensitivity to DNA topoisomerase 1 poisons by altering R-loop metabolism.* BioRxiv, 2023.
39. Chen, Z., et al., *In vivo CD8(+) T cell CRISPR screening reveals control by Fli1 in infection and cancer.* Cell, 2021. **184**(5): p. 1262-1280 e22.
40. Li, S., C. Vemuri, and C. Chen, *DNA topology: A central dynamic coordinator in chromatin regulation.* Curr Opin Struct Biol, 2024. **87**: p. 102868.
41. Broderick, L., et al., *Mutations in topoisomerase IIbeta result in a B cell immunodeficiency.* Nat Commun, 2019. **10**(1): p. 3644.
42. Amatullah, H., et al., *Epigenetic reader SP140 loss of function drives Crohn's disease due to uncontrolled macrophage topoisomerases.* Cell, 2022. **185**(17): p. 3232-3247 e18.
43. Rialdi, A., et al., *Topoisomerase 1 inhibition suppresses inflammatory genes and protects from death by inflammation.* Science, 2016. **352**(6289): p. aad7993.
44. Ho, J.S.Y., et al., *TOP1 inhibition therapy protects against SARS-CoV-2-induced lethal inflammation.* Cell, 2021. **184**(10): p. 2618-2632 e17.
45. *whitfield-et-al-2002-identification-of-genes-periodically-expressed-in-the-human-cell-cycle-and-their-expression-in.*
46. Neumann, T., et al., *Quantification of experimentally induced nucleotide conversions in high-throughput sequencing datasets.* BMC Bioinformatics, 2019. **20**(1): p. 258.
47. Lambert, S.A., et al., *The Human Transcription Factors.* Cell, 2018. **172**(4): p. 650-665.
48. Hammelman, J., et al., *Ranking Reprogramming Factors for Directed Differentiation.* bioRxiv, 2021: p. 2021.05.14.444080.
49. Yu, G., et al., *clusterProfiler: an R package for comparing biological themes among gene clusters.* Omics, 2012. **16**(5): p. 284-7.

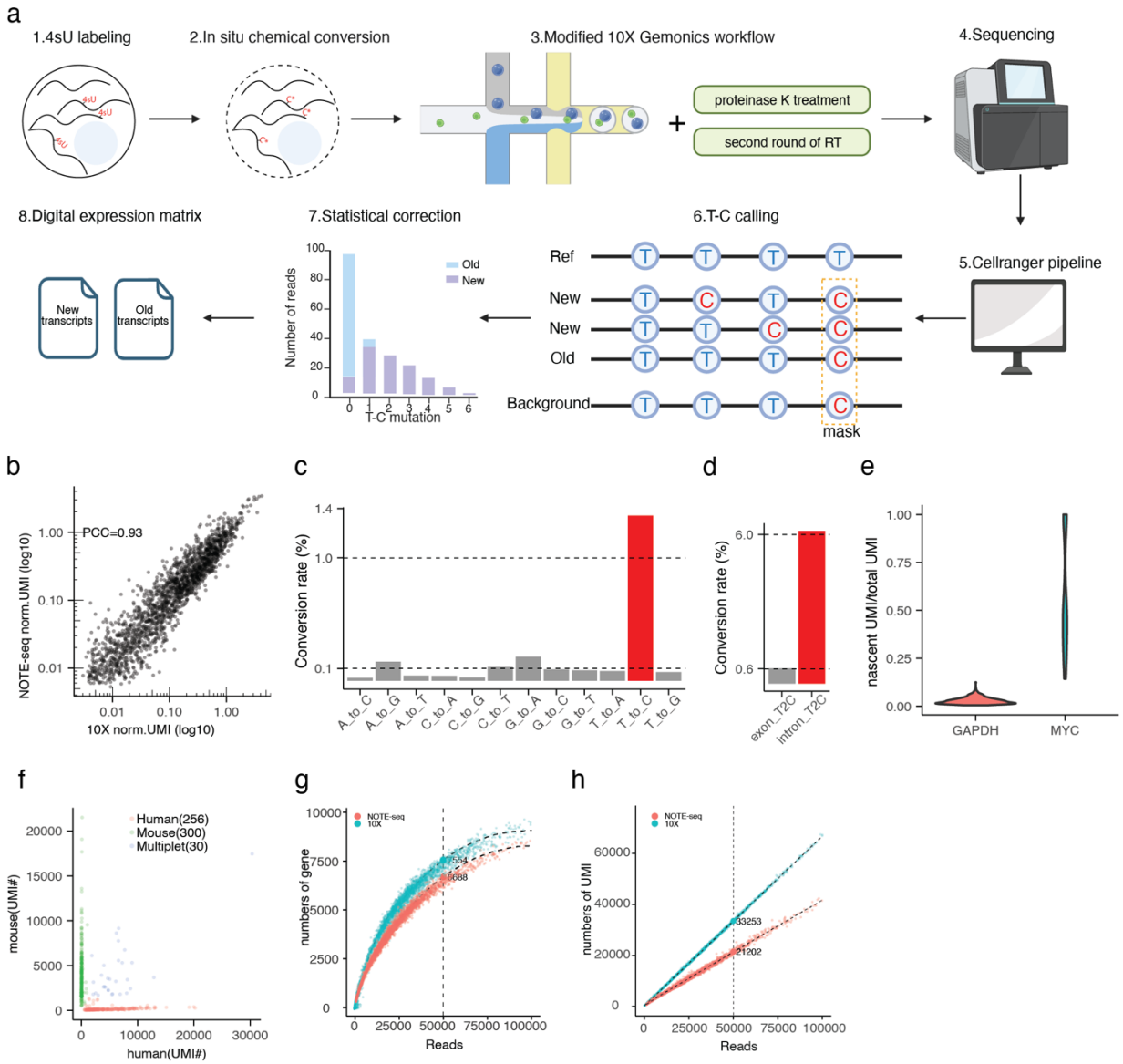

Figure 1. NOTE-seq workflow and performance. a. Workflow of NOTE-seq. **b**. Correlation of average normalized UMI counts of eHAP between NOTE-seq (n=3046 cells) and the standard 10X workflow (n=2774 cells). The Pearson's correlation coefficient (PCC) was calculated based on 14771 genes. **c**. Overall base conversion rates derived from the NOTE-seq data, with the T-to-C conversion rate highlighted in red. **d**. T-to-C conversion rate in exonic reads and intronic reads. **e**. Fraction of newly synthesized RNA in a low-turnover gene (GAPDH, n=674 cells) and a high-turnover gene (MYC, n=128 cells). **f**. Transcript counts of individual cells mapped to the merged human-mouse genome, with cells colored based on their assigned categories of human, mouse, or multiplet. **g**. Comparison of detected gene numbers per cell between NOTE-seq and the standard 10X workflow. **h**. Comparison of detected UMI numbers per cell between NOTE-seq and the standard 10X workflow.

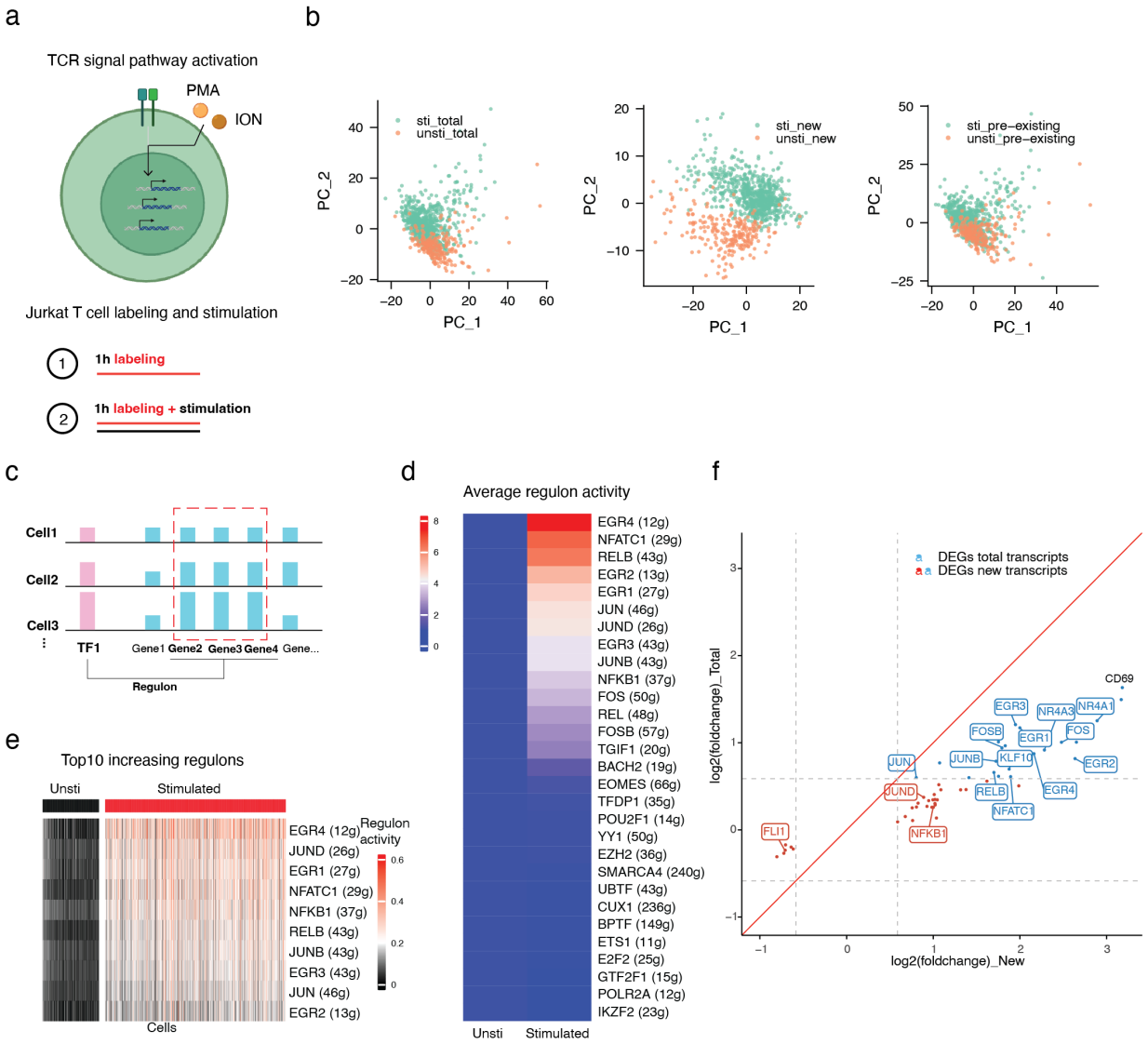

Figure 2. NOTE-seq reveals gene expression temporal dynamics during early-stage T-cell activation. **a**. Schematic of PMA/ION-mediated T-cell activation and 4sU labeling upon T-cell stimulation. **b**. Cell clustering analysis of stimulated (n=748 cells) and unstimulated (n=233 cells) Jurkat T cells by PCA, using total transcripts, newly synthesized transcripts (new), and pre-existing transcripts. **c**. Schematic of regulon analysis by correlating expression levels of transcriptional factors and their downstream target genes. **d**. Comparison of regulon activities between unstimulated and stimulated Jurkat T cells. The stimulated regulon activity was normalized to the corresponding unstimulated regulon activity. **e**. Top 10 significantly activated regulons upon T-cell stimulation plotted for each single cell, with data derived from single-cell transcriptome matrix and scaled for each row. **f**. Differentially expressed genes identified using total transcripts or new transcripts (blue) and only using new transcripts (red) upon T-cell stimulation. Among the genes, names of transcription factors were highlighted in the boxes.

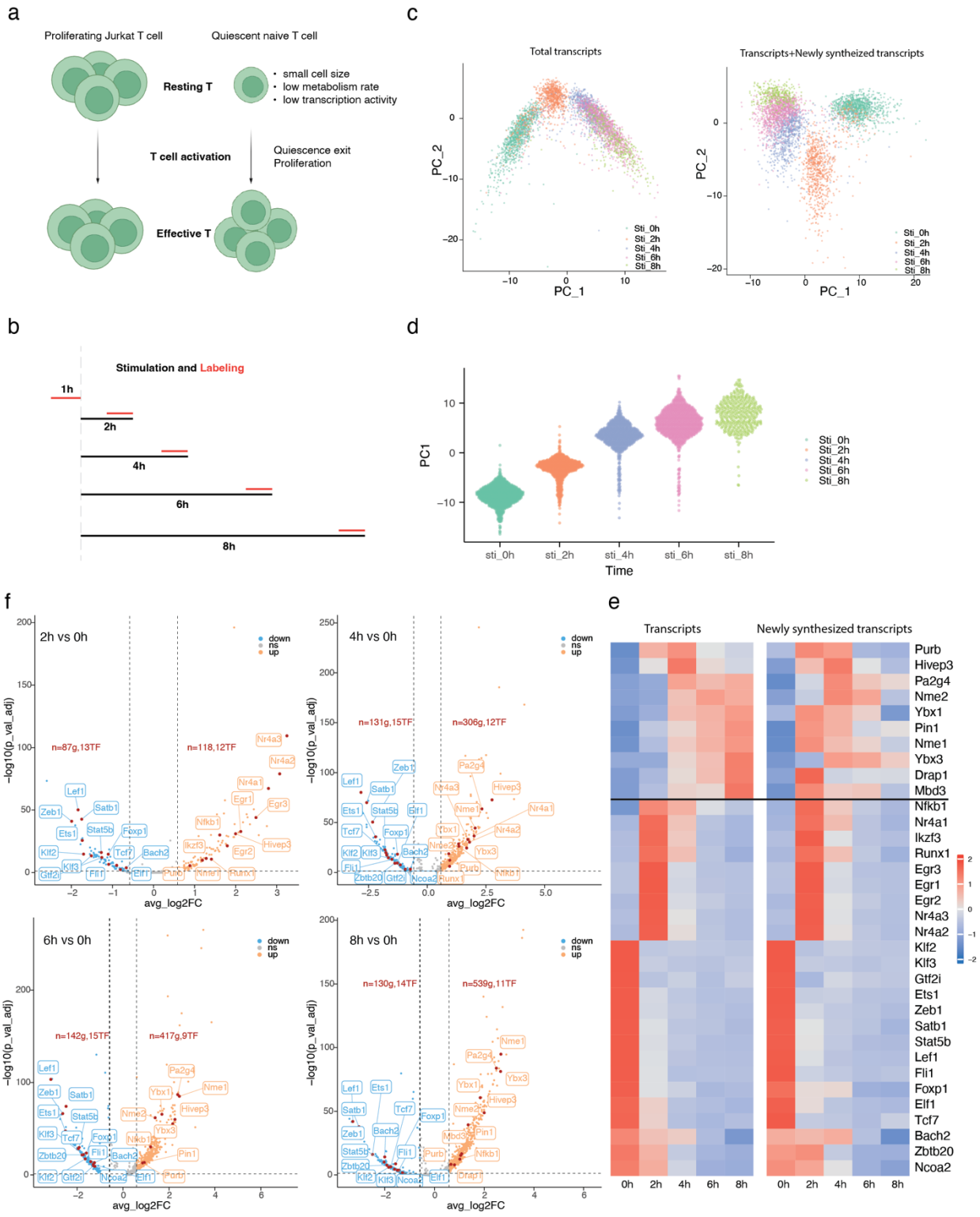
a
Proliferating Jurkat T cell
Quiescent naive T cell
Resting T
• small cell size
• low metabolism rate
• low transcription activity
T cell activation
Quiescence exit
Proliferation
Effective T
b
Stimulation and Labeling
1h
2h
4h
6h
8h
c
Total transcripts
Transcripts+Newly syntheized transcripts
PC_1
PC_2
Sti_0h
Sti_2h
Sti_4h
Sti_6h
Sti_8h
d
PC1
Time
sti_0h
sti_2h
sti_4h
sti_6h
sti_8h
f
2h vs 0h
4h vs 0h
6h vs 0h
8h vs 0h
down
ns
up
n=87g,13TF
n=118,12TF
n=131g,15TF
n=306g,12TF
n=142g,15TF
n=417g,9TF
n=130g,14TF
n=539g,11TF
−log10(p_val_adj)
avg_log2FC
e
Transcripts
Newly synthesized transcripts
Purb
Hivep3
Pa2g4
Nme2
Ybx1
Pin1
Nme1
Ybx3
Drap1
Mbd3
Nfkb1
Nr4a1
Ikzf3
Runx1
Egr3
Egr1
Egr2
Nr4a3
Nr4a2
Klf2
Klf3
Gtf2i
Ets1
Zeb1
Satb1
Stat5b
Lef1
Fli1
Foxp1
Elf1
Tcf7
Bach2
Zbtb20
Ncoa2
0h
2h
4h
6h
8h

Figure 3. NOTE-seq identifies transcription factors for gene regulation during early-stage T-cell activation. a. Schematic of distinct processes during early-stage T-cell activation for Jurkat T cells and naïve T cells. **b**. Strategy of T-cell stimulation and 4sU labeling. Black line represents time for stimulation and the red line represents time for 4sU labeling. **c**. Cell clustering analysis of naïve T cells stimulated for the indicated time (1014 cells for 0 hour, 809 cells for 2 hours, 831 cells for 4 hours, 1016 cells for 6 hours, and 359 cells for 8 hours). PCA analyses and plots were based on total transcripts (left) or total and newly synthesized transcripts jointly (joint transcriptome, right). **d**. Principal components of the joint transcriptome at different time points after T-cell stimulation. **e**. Differentially expressed genes identified upon T-cell stimulation at different time points. Among the genes, names of transcriptional factors were highlighted in the boxes. **f**. Heatmap of the level of total transcripts and newly synthesized transcripts for each differentially expressed transcription factor, with the data matrix scaled for each row.

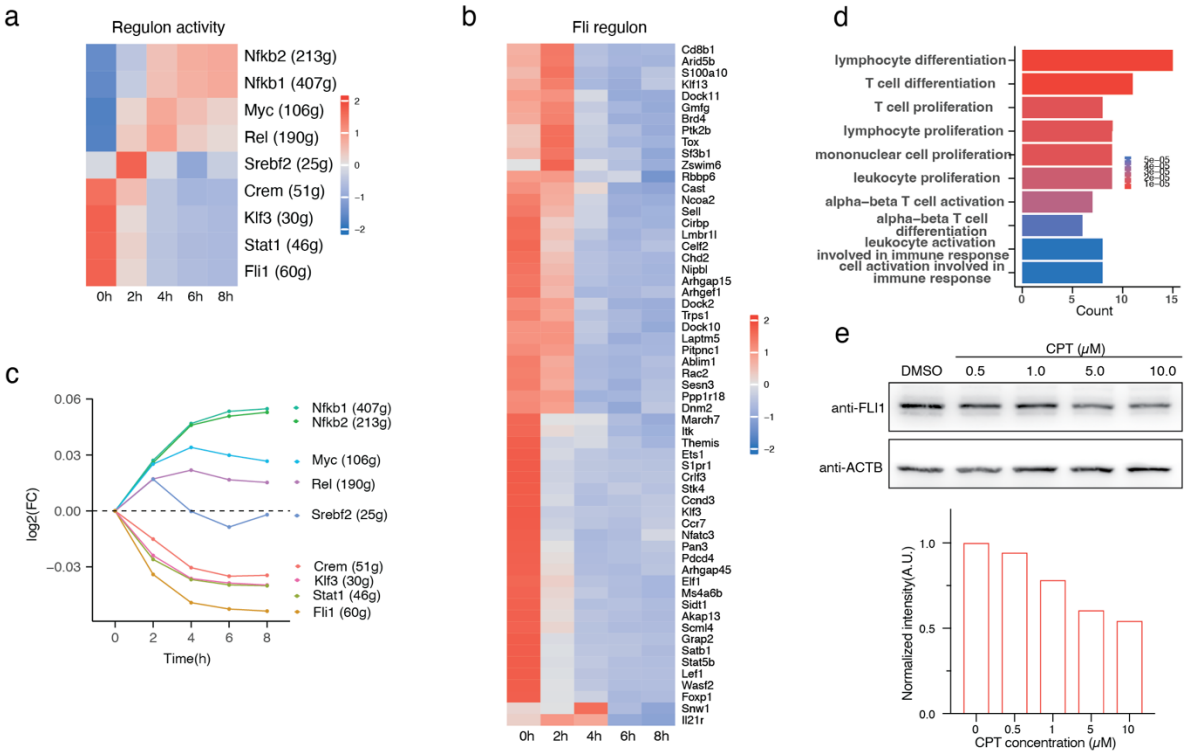

Figure 4. *Fli-1* is a master transcription factor during early-stage naïve T-cell activation modulated by topoisomerase inhibition. a. Activity of regulons at different time points during early-stage naïve T-cell activation, with the data matrix scaled for each row. **b**. Transcription level of downstream target genes in the *Fli-1* regulon at different time points during early-stage naïve T-cell activation, with the data matrix scaled for each row. **c**. Fold changes of regulon activity at different time points during early-stage naïve T-cell activation **d**. GO enrichment analysis of genes in the *Fli-1* regulon. **e**. Western blot of FLI-1 after CPT treatment, with FLI-1 level normalized to the ACTB level, and quantified relative to the negative control without CPT treatment. The data is representative of two experiments.