RNA dynamics revealed by metabolic RNA labeling and biochemical nucleoside conversions

Marisa A P Baptista & Lars Dölken

The combination of metabolic RNA labeling with biochemical nucleoside conversion now adds a broadly applicable temporal dimension to RNA sequencing.

The life cycle of RNA is governed by tightly regulated molecular events. Massively parallel RNA sequencing (RNA-seq) provides insight into the general regulation of gene expression but does not differentiate its molecular basis, namely, whether the observed differences result from alterations in RNA transcription, processing or decay. Two studies published in *Nature Methods*[1,2] and a third study published in *Angewandte Chemie International Edition*[3] now provide three equivalent, rapid and sensitive approaches for studying RNA dynamics in mammalian cells (**Fig. 1**). All three are based on metabolic RNA labeling and simple biochemical nucleoside conversions followed by RNA-seq.

In recent years, powerful biochemical approaches have been developed to determine real-time changes in RNA transcription, processing and decay. One of the most widely applied approaches, termed 4sU-tagging, involves metabolic labeling of newly transcribed RNA in living cells with thiol-labeled nucleoside analogs such as 4-thiouridine (4sU). After the isolation of total cellular RNA and thiol-specific biotinylation, total cellular RNA can be efficiently separated into newly transcribed and pre-existing (unlabeled) RNA[3,4]. The approach is applicable to virtually all model organisms, including vertebrates, insects[5] and yeast (after transporter-enhanced 4sU uptake)[6], and dramatically increases the temporal resolution of short-term changes in gene expression occurring within the frame of minutes to hours. It provides noninvasive access to precise RNA half-life measurements[4]. When combined with ultra-short (~5 min) RNA labeling, it reveals the kinetics of RNA processing[7] and depicts the transcription of even the most unstable RNAs[8]. A modified version uses 4-thiouracil (4tU) and a special uracil salvage pathway that requires uracil-phosphoribosyl transferase (UPRT) to convert 4tU to 4sU-monophosphate and thereby initiate metabolic labeling only in cells expressing UPRT[3]. When this approach is used with transgenic systems, cell-type-specific newly transcribed RNA profiles can be obtained from complex tissues in mice and *Drosophila*[5,9].

As with any biochemical separation method, the underlying protocols are laborious and require a high level of specific expertise, as well as ample starting material. All three of the recently published approaches now provide a simple solution to this problem. They are based on the identification of single point mutations in RNA-seq data introduced by chemical conversion of 4sU residues into cytosine analogs. Whereas Herzog *et al.*[1] used the thiol-reactive compound iodoacetamide to efficiently alkylate 4sU residues (>95% in 15 min) in their SLAM-seq (thiol(SH)-linked alkylation for the metabolic sequencing of RNA) approach, Schofield *et al.*[2] used the combination of 2,2,2-trifluoroethylamine and sodium periodate (~80% in 1 h) for their TimeLapse-seq method. Both approaches rapidly convert 4sU into cytosine analogs, which are efficiently transcribed by common reverse transcriptases. In contrast, Riml *et al.*[3] use osmium-mediated conversion of 4sU to cytidine by aqueous ammonia (>98% in 4 h) in their TUC-seq method. Thus labeling and conversion rates of about 1 in 40 to 1 in 150 uridines can now be achieved, allowing bioinformatic identification of reads originating from newly transcribed RNA based on U-to-C conversion. This became possible only with recent advancements in the fidelity of Illumina sequencing technology, with error rates dropping below 1 in 1,000 nucleotides, as observed with, for example, the HiSeq 2500 sequencing system. Thus even sequencing reads with a single U-to-C conversion may be attributed to newly transcribed RNA. However, it is important to note that the latest and most powerful sequencing systems (MiniSeq, NextSeq and NovaSeq), which are based on two-color rather than four-color chemistry, have higher error rates and are thus currently prohibitive for nucleotide-conversion experiments.

Results from all three studies demonstrate that 4sU-labeling and subsequent nucleotide conversion increase the temporal resolution for the detection of short-term changes in gene expression to a similar extent as reported for 4sU-tagging after as little as 45 min of 4sU-labeling. However, it is important to note that after a 60-min period, only about 1–5% of the total cellular RNA pool in mammalian cells will consist of newly transcribed RNA. Therefore, a maximum of 1–5% of all sequencing reads will contain U-to-C conversions. Thus, short 4sU pulses require a large number of sequencing reads to also achieve sufficient coverage for less abundantly transcribed genes.

Metabolic labeling of newly transcribed RNA enables measurement of RNA decay rates in two different ways. In steady-state conditions, RNA transcription compensates for RNA decay. Precise RNA half-lives can thus be determined on the basis of the ratio of newly transcribed to total or unlabeled, pre-existing RNA[4]. Alternatively, prolonged metabolic labeling followed by washout and chase can be used to analyze changes in RNA decay rates after changes in conditions. In particular, SLAM-seq gave RNA decay rates for >8,400 different genes in mouse embryonic

Marisa A.P. Baptista and Lars Dölken are at the Institute of Virology and Immunobiology, Julius Maximilian University of Würzburg, Würzburg, Germany.
e-mail: lars.doelken@uni-wuerzburg.de

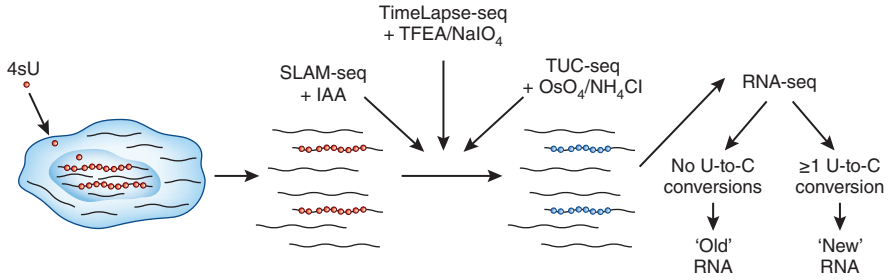

Figure 1 | The principle of the three new nucleotide-conversion approaches. Exposure of cells to a thiol-labeled nucleoside (4sU) results in rapid uptake and incorporation into newly transcribed RNAs, which can be isolated and sequenced. IAA, iodoacetamide; TFEA, 2,2,2-trifluoroethylamine.

stem cells after comprehensive 4sU-labeling of cellular mRNAs using a 24-h 4sU pulse at non-toxic 4sU concentrations followed by washout and chase[1]. The method reliably depicted the regulatory effects of cellular microRNAs and post-translational RNA modifications (N^6-methyladenosine) on RNA stability[1]. Given the simplicity of nucleotide-conversion approaches and high 4sU-labeling densities, pulse-chase experiments thus provide a very interesting new means to analyze changes in RNA decay rates under non-steady-state conditions.

The TT-seq approach uses ultra-short metabolic labeling (5 min of 4sU-labeling) and partial fragmentation of the isolated total cellular RNA before purification of newly transcribed RNAs, and can be used to visualize even the most transient RNA species, including enhancer- and promoter-derived RNAs[8]. Although the contaminating background of unlabeled RNA carried over in 4sU-tagging experiments is usually extremely low (owing to the stringent wash conditions enabled by the streptavidin affinity purification), it can still result in an amount of contaminating reads exceeding 30% when only the minute amounts of nascent RNA transcribed during very short 4sU pulses are purified. Until now, this RNA could not be differentiated from nascent RNA reads and could be detrimental for TT-seq experiments. With the combination of TT-seq and TimeLapse-Seq, the absence of U-to-C conversions now allows the identification and removal of such contaminating reads[2]. The same holds true for 4sU-tagging carried out *in vivo*, for example, using 4tU and mice with cell-type-specific UPRT expression[5,9]. Here, computational removal of contaminating background RNA will substantially boost the applicability of these approaches.

In summary, nucleotide-conversion approaches offer fast, easy and broadly applicable techniques that can provide highly valuable information on the molecular mechanisms underlying the observed changes in total RNA levels for a broad range of applications from *in vivo* experiments to single-cell RNA-seq.

COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

1. Herzog, V.A. *et al. Nat. Methods* **14**, 1198–1204 (2017).
2. Schofield, J.A., Duffy, E.E., Kiefer, L., Sullivan, M.C. & Simon, M.D. *Nat. Methods* **15**, 221–225 (2018).
3. Riml, C. *et al. Angew. Chem. Int. Edn. Engl.* **56**, 13479–13483 (2017).
4. Dölken, L. *et al. RNA* **14**, 1959–1972 (2008).
5. Miller, M.R., Robinson, K.J., Cleary, M.D. & Doe, C.Q. *Nat. Methods* **6**, 439–441 (2009).
6. Miller, C. *et al. Mol. Syst. Biol.* **7**, 458 (2011).
7. Windhager, L. *et al. Genome Res.* **22**, 2031–2042 (2012).
8. Schwalb, B. *et al. Science* **352**, 1225–1228 (2016).
9. Gay, L. *et al. Genes Dev.* **27**, 98–115 (2013).