


# Dynamics of transcriptional and post-transcriptional regulation

Mattia Furlan, Stefano de Pretis and Mattia Pelizzola 

Corresponding authors: Stefano de Pretis. Center for Genomic Science of IIT@SEMM, Fondazione Istituto Italiano di Tecnologia (IIT), 20139 Milan, Italy; E-mail: Stefano.depretis@iit.it; Mattia Pelizzola. Center for Genomic Science of IIT@SEMM, Fondazione Istituto Italiano di Tecnologia (IIT), 20139 Milan, Italy; Tel.: +39-0294375019; E-mail: mattia.pelizzola@iit.it

## Abstract

Despite gene expression programs being notoriously complex, RNA abundance is usually assumed as a proxy for transcriptional activity. Recently developed approaches, able to disentangle transcriptional and post-transcriptional regulatory processes, have revealed a more complex scenario. It is now possible to work out how synthesis, processing and degradation kinetic rates collectively determine the abundance of each gene's RNA. It has become clear that the same transcriptional output can correspond to different combinations of the kinetic rates. This underscores the fact that markedly different modes of gene expression regulation exist, each with profound effects on a gene's ability to modulate its own expression. This review describes the development of the experimental and computational approaches, including RNA metabolic labeling and mathematical modeling, that have been disclosing the mechanisms underlying complex transcriptional programs. Current limitations and future perspectives in the field are also discussed.

**Key words:** RNA metabolic labeling; nascent RNA; RNA synthesis; RNA processing; RNA degradation; mathematical modeling

## Introduction

The genetic information encoded in transcriptional units is transferred to end products, non-coding transcripts and proteins, via the complex series of events that constitute the cell's gene expression programs, controlled by transcriptional, co-transcriptional, and post-transcriptional regulatory cues [1]. Profiling of transcriptional maps by high-throughput sequencing is currently considered routine, and public repositories include thousands of RNA-seq experiments that enable both absolute and comparative gene expression quantification [2]. However, RNAs being transient species, the mere quantification of a transcript's copy number is poorly informative of the underlying

dynamics and could actually lead to misleading conclusions [3]. Taking a transcript's abundance as a direct measurement of its corresponding gene's transcriptional activity is a widespread oversimplification. Rather, a poorly transcribed gene could see many of its RNA molecules accumulate just because they are highly stable (i.e. long half-life) and therefore can persist long after the transcriptional unit has been turned off. Conversely, if a gene is very actively transcribed but its transcripts are highly unstable species, only a few copies of its RNAs are found in the cell [4].

In general, gene expression could be better deciphered if the transcriptional output were broken down into its two key components, premature and mature RNA, each with its own

**Mattia Furlan** is a postdoc at the Center for Genomic Science of IIT@SEMM (Fondazione Istituto Italiano di Tecnologia). He was trained as a theoretical physicist and his research interests include the development of mathematical modeling approaches for the study of RNA dynamics, and their characterization through single molecule and single cell approaches.

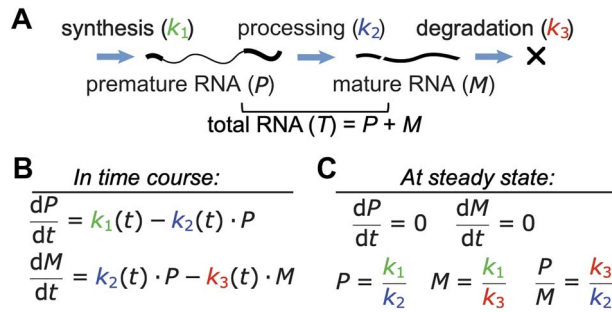
**Stefano de Pretis** is a senior postdoc at the Center for Genomic Science of IIT@SEMM (Fondazione Istituto Italiano di Tecnologia). He was trained as a bioinformatician and his research interests include the development of mathematical modeling approaches for the study of the dynamics of RNA and RNA polymerase II.

**Mattia Pelizzola** is the Head of the Epigenomics and Transcriptional Regulation Unit at the Center for Genomic Science of IIT@SEMM (Fondazione Istituto Italiano di Tecnologia), which relies on interdisciplinary approaches to study the dynamics of RNA metabolism in physiological and disease conditions.

Submitted: 12 October 2020; Received (in revised form): 12 November 2020

© The Author(s) 2020. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact [journals.permissions@oup.com](mailto:journals.permissions@oup.com)



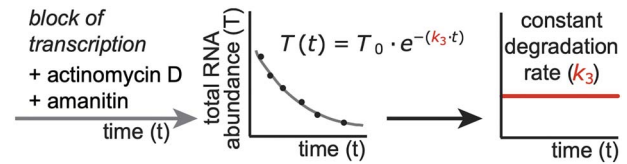
**Figure 1.** Mathematical modeling of the RNA life cycle. (A) Schema of the RNA life cycle, including the steps of premature RNA synthesis, its processing into mature RNA, and the degradation of the latter; the corresponding kinetic rates are indicated in parenthesis. (B) Mathematical representation of the RNA life cycle through ODEs. The first equation describes the modulation of premature RNA over time ( $dP/dt$ ) as the balance between the amount of RNA molecules synthesized ( $k_1(t)$ ) and those processed to become mature transcripts ( $k_2(t) \cdot P$ ). The second equation describes the modulation of mature RNA ( $dM/dt$ ) as the balance between the amount of premature RNA molecules processed ( $k_2(t) \cdot P$ ) and the amount of mature RNA transcripts degraded ( $k_3(t) \cdot M$ ). (C) Steady state solution of the ODEs system.

rates of production and decay [5]. This set of information would allow a fuller understanding of the expression status of a given gene in a given biological condition. First, this information would reveal the close link between a gene's transcriptional activity and its actual transcriptional output, and that, at steady state, the population of its RNAs is only apparently static [3]. In fact, the tight balance between RNA production and decay results into a constant flowing of genetic information, comparable to the constant flowing of water in and out of a sink where both the faucet and the drain are open. Second, this information would unveil a cell's ability to modulate gene expression levels at a given speed [3]. For example, if the faucet were suddenly closed, the time the sink would take to become empty would be entirely dictated by the size of the drain, i.e. the rate of degradation of a given RNA species. Indeed, transcripts with either short or extended half-lives are, respectively, highly and poorly responsive to perturbations [6].

### The kinetic rates of the RNA life cycle

Altogether, the RNA life cycle can be depicted with a simple model where premature transcripts are produced at a given 'synthesis rate' and transformed into mature transcripts at a given 'processing rate' (Figure 1A). Eventually, mature transcripts are lost at a given decay or 'degradation rate'. Therefore, processing and degradation rates determine the half-life of premature and mature species, respectively. This model is commonly represented as a system of ordinary differential equations (ODEs, Figure 1B) [5, 7], which reveals that, at steady state, the expression of premature RNAs equals the ratio between synthesis and processing rates, while the expression of mature RNAs equals the ratio between synthesis and degradation rates (Figure 1C). As an obvious consequence of this model, processing and degradation rates have no impact on the steady state abundance of mature and premature species, respectively [3].

For years, RNA half-life has been the only other piece of information that could complement the quantification of total RNA abundance. Nowadays, multiple approaches are available to quantify the whole set of RNA kinetic rates, disclose the behind-the-scenes details of gene expression programs, and shed light on their implications. This review will cover the experimental



**Figure 2.** Quantification of a transcript's half-life through block-of-transcription experiments. Key steps in the estimation of RNA degradation rates through block-of-transcription experiments. First, a drug blocking transcription is provided. Then, the progressive reduction in gene expression is measured at several time points. Finally, the exponential decay coefficient is used to quantify the transcript half-life.

and computational key approaches that are currently available for the genome-wide quantification of the kinetic rates of RNA synthesis, processing, and degradation. First, the approaches that can be applied to total RNA-seq experiments will be covered. Then, the approaches that rely on nascent RNA profiling will be discussed. Finally, recent approaches enabling the study of single cells will be summarized, and key assumptions, limitations, and future perspectives will be discussed.

### Quantification of RNA half-lives through block-of-transcription experiments

The first approach to be established in the field was RNA half-lives quantification through block-of-transcription experiments [8]. Cells are treated with drugs blocking transcription, such as actinomycin D or amanitin. Then, a simple first-order kinetic model is used to describe the decrease in transcripts' levels and their half-lives are quantified by estimating their exponential decays (Figure 2). This approach has been (and still is) often adopted for the quantification of the half-lives of individual RNAs. This can be done by PCR, designing primers that match exons or exon-intron junctions and quantifying the progressive reduction in the abundance of mature or premature RNA at different times following drug treatment. In some cases, this approach has also been adopted in the genome-wide quantification of RNA half-lives, by quantifying the transcripts' reduction with microarrays [9, 10] and high-throughput sequencing [11–13]. However, this method is extremely invasive for the cell, and among other effects, it can also affect RNA molecules stability [14]. Moreover, fitting the exponential decay of transcripts expression levels requires multiple data points and the optimal spacing between time points can vary between genes with extremely short or extended half-lives. Finally, the extrapolation of the half-life of mature RNAs from their progressive reduction neglects the fact that, despite the block of transcription, additional mature molecules could have originated following the processing of residual premature molecules. This is potentially very important for those genes that are characterized by markedly slow processing rates. Altogether, this technique is cumbersome and not suitable to the study of the dynamic regulation of RNA half-life.

### The importance of the quantification of premature and mature RNA expression

As depicted in Figure 1, the determination of the RNA kinetic rates requires the quantification of premature and mature RNA abundance [15]. Commonly, these are respectively obtained from exonic and intronic RNA-seq signals in high-throughput sequencing short-read experiments. However, complicating the interpretation of intronic signals is the potential occurrence of

alternative splicing events, including the possibility that introns are retained in the mature form [16].

Premature transcripts have shorter half-lives than their mature counterparts. For example, in mouse fibroblast cells, RNA processing rates are more than 10 times faster than degradation rates [17]. In other words, RNA processing has a faster kinetic rate than RNA degradation, and premature RNAs are less abundant than mature ones. Moreover, both species are largely outnumbered by transcripts of ribosomal genes. For these reasons and in order to maintain the information on premature RNA abundance, RNA-seq experiments based on the depletion of ribosomal RNAs are preferable to the most commonly used preparations that include the selection of the poly-adenylated transcripts [15].

Nonetheless, standard sequencing kits profiling poly-adenylated transcripts have a substantial (~10% of the reads) residual intronic expression [18], which can be used to measure the relative levels of premature RNA expression with good approximation [3]. The drawback of this approach is that the inferred expression levels of premature RNAs are significantly underestimated.

### Approaches to quantify RNA kinetic rates based on total RNA-seq

The quantification of premature and mature RNA expression levels is necessary but not sufficient for the quantification of RNA kinetic rates. Indeed, the model presented in Figure 1 is hampered by the indetermination of the system, where three unknowns are governed by two equations. A few approaches based on steady-state observations were developed that consider intronic expression as a proxy for synthesis rates. While this greatly simplifies the problem from a mathematical point of view, it neglects the fact that intronic signals, i.e. premature RNA expression, result from the joint action of two processes: premature RNA synthesis and processing into mature form. Therefore, these approaches neglect the contribution of RNA processing or rely on the strong assumption that the rate of processing is constant.

#### Available methods

One of these approaches is EISA, which allows the quantification of changes in synthesis and degradation. By adopting the assumptions discussed above, this method uses changes in intronic signals as a proxy for changes in transcriptional activity [19]. A variation of this approach was recently employed in Rembrandts, a new tool that, by including a term modeling the coupling between synthesis and processing, allows a more refined estimation of the variations in degradation rates between conditions [20]. Noteworthy, all these methods only provide relative changes in the kinetic rates between different steady state conditions. Conversely, Snapshot-Seq aims at determining the absolute rates of synthesis and degradation, using intronic signals for the former, and the ratio between exonic and intronic signals for the latter [21]. Finally, an approach for modeling synthesis and degradation rates on time course data was proposed by Zeisel *et al.* [6], but it requires a priori information on RNA processing rates and assumes them to be constant.

In a complementary approach, histone marks closely related to transcription (H3-K4me3, -K27me3, and -K36me3) were used in place of RNA-seq intronic signals as proxy for transcriptional activity, and a regression approach was adopted to compare

them with the RNA-seq transcriptional output. Here, the rationale is that the difference between expected (histone marks) and observed (RNA-seq) transcriptional outputs can be used to infer decay rates [22].

Finally, we recently proposed a novel approach, INSPECT-, which, for the first time, enabled us to model the whole set of RNA kinetic rates based on time course RNA-seq datasets [3, 23]. INSPECT- considers the temporal delay between the responses of premature and mature forms proportional to mature RNA half-life, which can thus be estimated. Following this, the rates of RNA synthesis and processing can be determined by solving the system at steady state:  $k_1 = M \cdot k_3$  and  $k_2 = k_3 \cdot \frac{M}{P}$ , respectively (Figure 1). Further modeling is used to refine first-guess values of all kinetic rates and to quantify the significance of their variation over time. If time course data are not available, INSPECT- determines the ratio between premature and mature RNA expression to quantify the ratio between degradation and processing rates (Figure 1). While this does not permit the deconvolution of the contribution of each rate, changes in this ratio for a given gene across conditions indicate that one of those rates was altered, suggesting the occurrence of post-transcriptional regulation. This approach was formerly introduced by Gaidatzis *et al.* [19] and implemented in EISA. However, EISA assumes that RNA processing rates are constant across conditions, and changes in the ratio are entirely attributed to RNA degradation, which could be an issue in those conditions where both RNA processing and degradation are modulated.

Among these approaches, EISA, Rembrandts, and INSPECT- are the only ones for which executable codes have been released. EISA and INSPECT- are available as part of the eisaR and INSPECT Bioconductor packages, while Rembrandts is available in the form of scripts.

#### Advantages of approaches that only require total RNA-seq data

Despite their assumptions and limitations, these approaches have two key advantages: they are straightforward to apply, since they only require total RNA-seq data, and they permit to retrospectively analyze the thousands of sequencing datasets archived in public repositories.

When applied to analyze the transcriptomes of various tissues, EISA revealed the post-transcriptional regulation of gene sets targeted by specific miRNAs [19].

When used to reanalyze a number of publicly available time-course datasets in human, mouse, and plants, INSPECT- permitted to distinguish between the contributions of transcriptional and post-transcriptional regulation by complementing and expanding already published gene expression analyses [3]. INSPECT- was also used to discern between the transcriptional and post-transcriptional impact of the MYC transcription factor during B Cell activation [24]. Finally, exploiting its ability to map changes in post-transcriptional regulation between steady state conditions, INSPECT- was used to chart landscapes of post-transcriptional regulation in hundreds of RNA-seq experiments, covering dozens of tissues and disease conditions. Those data revealed a signature of brain genes, some of which involved in amyotrophic lateral sclerosis, which are potentially post-transcriptionally regulated by m6A RNA modifications [3].

These approaches demonstrated that it is possible to extract precious information from total RNA-seq datasets, decoupling transcriptional and post-transcriptional regulation through the quantification of RNA kinetic rates. However, this comes at the cost of important assumptions, such as neglecting or assuming

constant rates of RNA processing, or at the cost of specific requirements on the experimental design, such as time course transcriptional profiling.

## Approaches that quantify RNA kinetic rates based on nascent RNA profiling

As discussed in the previous section, quantifying all three RNA kinetic rates based on premature and mature RNA expression is not trivial, especially at steady state. Indeed, available approaches either require priori for the RNA processing rates [6] or require time course data [3].

To overcome this issue and find a unique solution to the ODEs system introduced in Figure 1, additional information must be gathered. An obvious choice would be the independent quantification of one kinetic rate, for instance the estimation of transcripts' half-lives. However, this solution would carry with it all the limitations previously mentioned for the block-of-transcription technique. A turning point in the field came when approaches for direct genome-wide measurements of nascent transcription became available. Indeed, when profiled over relatively short-time intervals, nascent transcription provides a direct measurement of the rate of premature RNA synthesis [25]. Quantification of RNA synthesis rates through nascent RNA profiling greatly facilitates the resolution of the equations described in Figure 1 and enables the quantification of processing and degradation rates.

### Profiling nascent transcription

Multiple approaches are available that profile nascent transcription [26] by relying on either RNA metabolic labeling or the isolation of chromatin associated RNAs or of transcripts associated to transcriptionally active polymerase complexes (Figure 3A). Among these approaches, those employing RNA metabolic labeling have proved to be particularly successful [27]. Indeed, the enrichment of chromatin-associated RNAs also captures stably associated pre-existing transcripts. The isolation of transcripts associated to transcriptionally engaged polymerases, as in GRO-seq [28] and PRO-seq [29], circumvents this issue but cannot disentangle the speed of the RNA polymerase II (RNAPII) complex from the transcriptional output. Hence, two genes (X and Y) of the same length, associated with the same number of productive RNAPII complexes, would return the same number of nascent transcripts, suggesting that their synthesis rates are identical. Rather, in metabolic labeling experiments, modified nucleotides are provided to the cells for a fixed length of time and are incorporated only into the transcripts that are synthesized over this time interval. Hence, if RNAPII complexes reading gene X elongate their RNA transcripts faster than those reading gene Y, they will generate more nascent transcripts and gene X's synthesis rate will result to be higher than gene Y's.

When profiling nascent RNA via metabolic labeling, two main approaches are available. One is based on physically separating nascent transcripts from pre-existing species. The other is based on the chemical derivatization of nascent transcripts. Specific computational tools have been developed to quantify the RNA kinetic rates depending on the experimental approach adopted.

### Quantification of RNA kinetic rates through purification of labeled RNA

The approaches that are based on nascent RNA purification typically rely on the incorporation of 4-thiouridine (4sU)

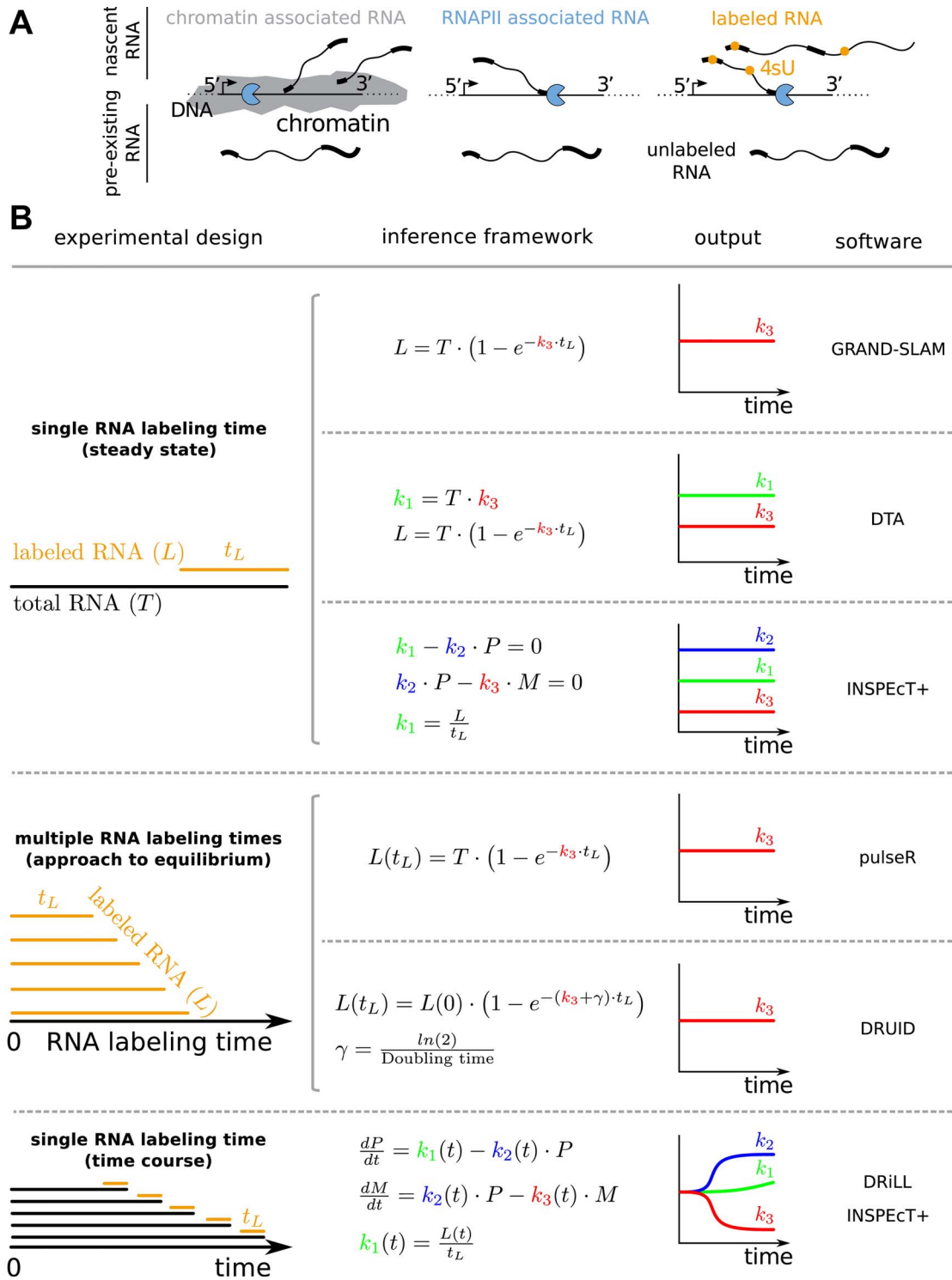
modified nucleotides into nascent transcripts. Following total RNA extraction, the 4sU nucleotides in the labeled transcripts are tagged with biotin, and nascent RNAs are physically separated from pre-existing ones and eventually sequenced [25]. Recently, TT-seq implemented a variation of this approach by adding a fragmentation step. In this way, data resolution was increased, and ultra-short labeling times could be adopted. This technique was applied for the characterization of nascent transcription at enhancers [30].

Different methods were developed that rely on nascent RNA profiling and involve the adoption of specific experimental designs and tailored computational methods. In terms of output, these methods differ for their ability to assess either the magnitude only or also the changes in RNA kinetic rates across different conditions (Figure 3B).

In the more straightforward approach, the sequencing of total RNA is complemented by profiling the nascent RNA with single 4sU pulses, from a few minutes to a number of hours long. In theory, the optimal labeling time depends on several experimental parameters and on the distribution of the degradation rates of interest. In general, short labeling times provide good estimates for fast decay kinetics, while long half-lives should be studied using longer 4sU pulses. Moreover, quickly degraded transcripts are affected by long labeling times more than stable RNAs that are affected by brief 4sU pulses; for this reason, the second of these two experimental settings should be generally preferred [31]. A number of computational tools for the analysis of these data are available. The first tool developed for this purpose was DTA, which quantifies the rates of synthesis and degradation under a single condition. For each condition, it requires three samples to be profiled: labeled, unlabeled, and total transcripts. In case the unlabeled fraction is not available, the ratio between labeled and total transcripts must be provided [32]. This approach was subsequently extended to allow the comparison among different conditions (cDTA) by using spike-ins [33]. More recently, pulseR was introduced to also provide the quantification of degradation rates [34]. pulseR is based on negative binomial models, which are particularly appropriate for RNA-seq counting data. It takes advantage of spike-ins, if available, otherwise it requires sequencing of the total RNA population in addition to the profiling of labeled and unlabeled conditions. Available from the INSPECT Bioconductor library together with INSPECT-, INSPECT+ is the only tool that can quantify all RNA kinetic rates, RNA processing rate included, in steady state conditions [7]. Moreover, it provides a statistical assessment of the variation of the kinetic rates across different conditions. INSPECT+ requires only the profiling of nascent and total RNAs, which are normalized with a computational procedure.

The approach-to-equilibrium methods were developed as an alternative to those based on single pulse. These methods are optimized to only quantify the RNA degradation rate and can be broadly divided into two categories: one where a series of experiments with increasing labeling times is conducted (pulse), and the other where RNA is quantified at different time points following complete labeling (chase). Currently, two different tools can be used for the quantification of degradation rates based on pulse approach-to-equilibrium data, DRUID [12] and pulseR [34]. While approach-to-equilibrium methods can be considered robust, they are restricted to the analysis of steady states and are less commonly adopted in the field due to the large number of required samples and high costs.

The temporal profiling of nascent transcription through metabolic labeling at different time points following a perturbation was adopted for the first time to characterize the response



**Figure 3.** Experimental and computational approaches to RNA kinetic rates quantification through nascent RNA profiling. (A) Alternative experimental approaches to the separation of nascent RNA from its pre-existing form. (B) Alternative approaches to RNA kinetic rates quantification. The corresponding experimental design, computational framework, output, and software are reported.

of dendritic cells to LPS [15]. In this study, all RNA kinetic rates were modeled, even though the rates of processing were assumed to be constant. Further development of the analysis methodology allowed to avoid this assumption and disclosed the full complexity of the RNA life cycle. Two different tools implemented this approach: DRILL [5] and INSPECT+ [7].

### Quantification of RNA kinetic rates through chemical derivatization of labeled RNA

Improving on the issues related to the purification of nascent RNA, various approaches were recently proposed in which alternative chemical reactions specifically converted modified

nucleotides incorporated in nascent transcripts into different bases [35]. These methods include SLAM-seq [36], TimeLapse [37], and TUC-seq [38]. A variation of the latter was recently proposed in which two modified nucleotides are used in subsequent pulses. The rationale is that differences in detecting the corresponding conversion signatures give information on both RNA synthesis and decay [39]. A couple of tools, GRAND-SLAM [40] and SLAM-DUNK [41], were developed that to be tailored to these data. While GRAND-SLAM only returns the rates of degradation, the rates of synthesis can be easily determined, in the context of the tool assumptions and equations reported in Figure 3B. SLAM-DUNK is particularly suited to the quantification of conversion rates and to the ability to simulate SLAM-seq datasets.

Noteworthy, once the quantification of nascent and pre-existing transcripts is obtained with these approaches, many of the computational tools developed for handling single pulse metabolic labeling data obtained through purification (Figure 3B) could be in principle adopted for the quantification of the RNA kinetic rates.

A novel method, Dyrec-seq, was recently proposed in which cells are first pre-cultured for 12 h in a medium with 5'-bromouridine (BrU), and then RNAs are collected at several time points after switching the cells to a 4sU-containing medium. On one hand, BrU containing RNAs are collected by immunoprecipitation; on the other hand, SLAM-seq is applied to identify 4sU containing RNAs. Eventually, the abundance of BrU and 4sU decreases and increases over time, respectively, allowing the quantification of degradation and synthesis rates [42]. The advantage of this method is that it allows the joint quantification of two kinetic rates. The disadvantage resides in the complexity of the experimental setup and in the number of samples that have to be collected for each condition.

### Purification of labeled RNA versus chemical derivatization: pros and cons

While the physical separation between nascent and pre-existing reads is convenient, it presents a number of drawbacks. In fact, it leads to the contamination of the labeled fraction with unlabeled transcripts (up to 30% for short labeling pulses, [3, 35]), and it requires sequencing of both labeled and unlabeled fractions thus increasing costs, especially when time course experiments are performed. Moreover, given the small amount of nascent RNA within the total population of transcripts (especially when short labeling times are used), substantial RNA has to be obtained to retrieve enough nascent RNA for its subsequent sequencing [27].

Finally, the integration of total and nascent RNA-seq data introduces normalization issues. These can be solved computationally based on the overdetermination of the ODEs system when synthesis rates are introduced as proxy from nascent RNA profiling, as in INSPEC<sup>T+</sup> [7] and DRiLL [5], resulting in the quantification of a global scaling factor for each condition. Alternatively, these can be solved experimentally by introducing the profiling of spike-ins, as in cDTA [33] and pulseR [34], or by adding the profiling of the unlabeled fraction in addition to the labeled fraction and total RNA, as in pulseR [34].

The key advantage of the methods relying on chemical derivatization is that reads from nascent transcripts can be *in silico* separated from those from their pre-existing counterparts. Hence, a single sequencing run is enough to characterize both transcripts populations, while there is no need for normalization between them. However, these methods typically require longer pulse times (60' or several hours, depending on the cell type).

Moreover, while they allow the identification of nascent RNAs with high specificity, they have low sensitivity [41].

### Comparison between degradation rates quantified through different approaches

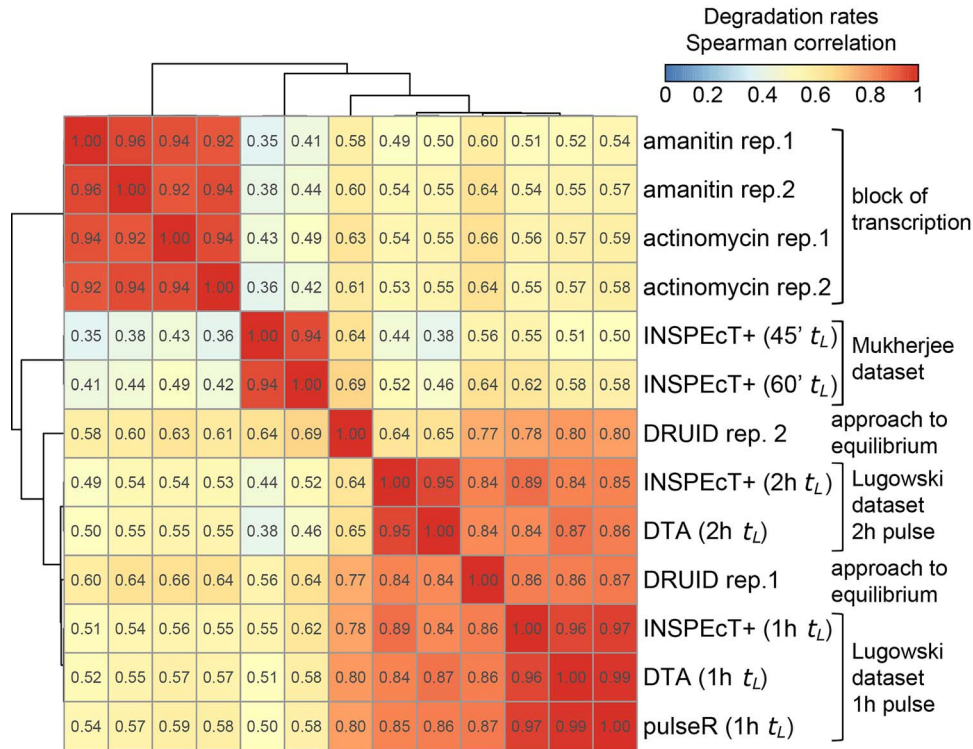
The various tools presented above differ markedly in terms of pre-requisites and outputs, so that comparing their results is not a trivial matter. To reach this aim, we took advantage of the fact that all tools permit to quantify RNA degradation rates. Hence, we looked for the cell type that had been characterized with the highest number of experimental approaches, including block-of-transcription experiments and different approaches based on RNA metabolic labeling, and identified the widely used HEK293 cell line [12]. For the tools and experimental settings already exploited to quantify the degradation rates in this cell type, we considered the data available in literature. For the other approaches, we applied them on the raw data from the same studies.

Pairwise Spearman's correlations among HEK293 degradation rates revealed that in general, each method had high internal reproducibility, in the 0.8–1.0 range (Figure 4). We took the degradation rates quantified through the approach-to-equilibrium method as our gold standard, as discussed above, and we expected the rates determined through block-of-transcription to have the lowest correlations, when compared with those from RNA metabolic labeling data. Indeed, two main clusters could be identified, separating the results obtained by blocking transcription from those obtained by RNA metabolic labeling. In particular, approach-to-equilibrium data analyzed by DRUID showed average correlations in the order of 0.6 with block-of-transcription results. The degradation rates determined by INSPEC<sup>T+</sup>, pulseR, and DTA were highly correlated with those reported by DRUID, in the order of 0.7–0.9, with higher correlation when data generated with the same labeling time were considered. Rather, the analysis of approach-to-equilibrium data for the same cells from an independent study reduced the correlation between INSPEC<sup>T+</sup> and DRUID to the order of 0.6–0.7. This suggests a strong dependence on specific settings of the protocol, the experimenter and/or the adopted reagents. Finally, the correlation between degradation rates determined by GRAND-SLAM, based on SLAM-seq data for mouse embryonic stem cells, was reported being in the order of 0.5–0.7 with block-of-transcription experiments using actinomycin D [40].

Noteworthy, both DTA and pulseR required the profiling of unlabeled RNA, which was not publicly available for the HEK293 cells. To overcome this limitation, we took advantage of the fact that both tools can work, as an alternative, on the ratio between the labeled and the total fractions, thus allowing their normalization. This ratio was set assuming a 25% yield in the labeled RNA purification step, and its setting had a minor impact on the reported correlations.

### Quantifying RNA kinetic rates at single cell level

A pioneering study conducted by La Manno *et al.*, allowed for the first time the characterization of RNA metabolism in single cells, and demonstrated the importance of characterizing the RNA kinetic rates for the interpretation of single cells transcriptomes [43]. Relying on the combination of intronic and exonic signals from total RNA data in single cells, the authors were able to estimate gene-level rates of synthesis and degradation.



**Figure 4.** Similarity between RNA degradation rates quantified by methods relying on block-of-transcription or RNA metabolic labeling. Heatmap of pairwise Spearman's correlations comparing degradation rates obtained for HEK293 cells using different approaches. Amanitin and Actinomycin D drugs were used to block transcription. Several tools for the analysis of RNA metabolic labeling data were compared, covering various experimental designs: approach to equilibrium, and single metabolic labeling performed at the indicated times ( $t_L$ ). All degradation rates derive from the Lugowski dataset [12], or from the reanalysis of the same data, with the exclusion of the HEK293 degradation rates determined in an independent study (Mukherjee dataset [68]).

However, the degradation rate of each gene was borrowed across different cells, and the rate of RNA processing was neglected and globally set to one. This allowed the authors to determine how far individual transcripts were from their closer steady state and the direction, i.e. velocity, they will take to reach it. When this information is aggregated for all transcripts, it provides an estimate of the overall direction that each single cell is taking in the transcriptomic space, being able to reconstruct connections between adjacent cells and temporal trajectories of the process under study. Yet, time scales remain in the domain of pseudotime. This approach, named RNA velocity, was recently improved by a novel method, scVelo, which, overcoming important limitations of the former approach, does not assume steady states and provides gene-level processing rates. Moreover, scVelo latent time better captures the ordering of events along temporal trajectories compared with pseudotime [44].

Recently, three studies were able to profile both total and nascent RNA in individual cells using RNA metabolic labeling through chemical derivatization of modified nucleotides. In the study presenting NASC-seq, they relied on 1 h metabolic labeling with 4sU followed by the alkylation of extracted RNA. This allowed detecting nascent transcription for  $\sim 700$  genes and the analysis of their differential transcription in stimulated cells [45]. In the study introducing Sci-fate, they quantified changes in nascent RNA following the temporal response to glucocorticoid receptor activation. This allowed determining gene-level half-lives; however, these were assumed to be constant across cells and conditions. Eventually, both changes in nascent RNA and constant degradation rates were used to infer cell states transitions. Sci-fate relies on combinatorial indexing and *in situ* 4sU chemical conversion in bulk fixed cells,

thus providing a more efficient and cheaper quantification of nascent and pre-existing RNA compared with competing methods [46]. A third study, introducing scSLAM-seq, was applied to cells subjected to mock or viral infection, uncovering the (epi)genomic determinants of intercellular transcriptional heterogeneity, and linking it to transcriptional bursting. In fact, the variability among cells of the gene-level ratio between nascent and total RNA was found to be strictly associated to the methylation status of CpG islands and to the presence of the TATA box, the latter being formerly implicated in transcriptional bursting [47].

Few approaches were recently proposed that integrate the profiling of nascent RNA through metabolic labeling into the dynamical models that infer the trajectories of single cells. A first set of methods enabled increasing accuracy of the inferred models and allowed predicting cells' trajectories during a dynamic process over extended time periods [48, 49]. Another approach, adopting pulse and chase experiments through the incorporation of 5-ethynyl-uridine in single cells (scEU-seq), allowed the estimation of constant synthesis and degradation rates. By placing individual cells along cell cycle or differentiation trajectories, scEU-seq could infer how synthesis and degradation rates change over time along those processes. However, the rate estimates were aggregated across cells [50].

The analysis of single cells transcriptomes revealed the importance of transcriptional bursting, in which intermittent bursts of expression alternate with transcriptional inactivity [51]. While this offers unprecedented opportunities for the study of transcriptional activity at the level of individual genes, it also poses major issues in the inference of transcripts half-lives in single cells. In fact, steady state assumptions are too loose to

be considered valid in single cells, and the ratio between the nascent and pre-existing molecules could be more indicative of the transcriptional status of a gene (on versus off) rather than informing on transcript stability. As a consequence, novel computational techniques are needed to estimate bursting kinetics unbiased half-lives in single cells, likely incorporating a dynamical component for the transcriptional status of individual genes.

## Current limitations

Experimental and computational methods in the field are nowadays sufficiently consolidated and streamlined, so that the quantification of RNA kinetic rates is accessible to every lab with standard equipment and access to high-throughput sequencing. In fact, as discussed above, user-friendly software applications exist, some of which only require standard RNA-seq data as input. However, a number of improvements are necessary to deal with the current limitations and to further increase the power of these approaches.

## Issues affecting the quantification of all RNA kinetic rates

Two are the issues that are detrimental to the quantification of all RNA kinetic rates, and they both involve the contamination of the labeled fraction. The first is due to the non-specific pull down of pre-existing transcripts, and the second results from the inability to precisely distinguish the labeled portion within nascent transcripts.

The first issue specifically affects RNA metabolic labeling approaches based on the purification of nascent transcripts. This issue directly contributes to inflating the rates of synthesis by returning higher levels of nascent transcription due to the contaminating unlabeled RNAs. Consequently, this inflates the quantification of the processing and degradation rates. Sophisticated normalization procedures [5, 7] and ways of measuring the level of contamination have been developed [3, 5, 31]. However, avoiding physically separating nascent and pre-existing RNA fractions is currently the most effective way to overcome this issue, as in the approaches that are based on the chemical derivatization of modified nucleotides [35]. Noteworthy, while the methods relying on the chemical derivatization of labeled transcripts are not affected by the contamination effect, they are however afflicted by errors in the quantification of nascent RNAs due to their low sensitivity. However, the amount of nascent transcription and the rate of synthesis are typically underestimated rather than inflated by these methods.

The second issue is partially solved by fragmenting the labeled transcripts before their purification, a key improvement first introduced in TT-seq [30]. This additional step maximizes the enrichment of nascent RNAs in the labeled fraction and reduces the incorporation of RNAs synthesized before the metabolic labeling by capturing only the parts that include modified nucleotides. This leads to a more precise measurement of the nascent RNA and is expected to improve the quantification of all RNA kinetic rates. Noteworthy, the difficulty in identifying the pre-existing portion in a metabolically labeled transcript, if any, is also shared with the methods relying on chemical derivatization. Indeed, modified nucleotides have a low rate of incorporation (one 4sU can be found every 43 uridines [36]), indicating that defining what part of a labeled transcript existed prior to the pulse is not easy.

## Issues in RNA processing quantification

The quantification of processing rates can be affected by two different issues. The first issue is the overestimation of premature RNA abundance, which can lead to shorter processing rates. This is expected to occur particularly for long mammalian introns, as the RNA polymerase can take a significant time transcribing an entire intron, before this becomes available for processing by the spliceosome machinery. Consequently, the RNA-seq signal in long introns decreases from the 5' to the 3' end. The quantification of intron abundance based on the 3' end region would provide a more robust measurement of premature RNA abundance [21]. However, this is not trivial to model [21], it is complicated by the typical reduced intronic coverage, and it is expected to affect no more than 10% of introns, having a length greater than 10Kb [52]. Therefore, most approaches prefer integrating the coverage across the whole intron.

The second issue, which also particularly affects long mammalian genes, is related to co-transcriptional processing [53, 54]. In the model represented in Figure 1, the rate of processing represents the speed at which premature RNA, following its synthesis, is converted into mature RNA. However, the time this process takes does not necessarily coincide with the time the actual splicing of the transcripts takes. Introns that are spliced almost immediately after the completion of their synthesis are expected to have similar processing and splicing rates. Rather, introns that are spliced post-transcriptionally may see their actual splicing being delayed. These can be due to the time the RNA polymerase takes to reach the gene end, which depends on the distance to be covered and the elongation speed of the complex, to the time the transcript takes to terminate, and to the time necessary for the recruitment and proper assembly of the spliceosomal complex. All these steps contribute to the processing rate. This issue is likely to become relevant when one is interested in estimating the speed of splicing, which can be important for mechanistic studies on this process.

Eventually, one way to solve these issues is calculating the RNA kinetic rates of individual introns, rather than averaging them over entire transcriptional units [5, 52].

## Issues in premature and mature RNA degradation

The quantification of degradation rates is based on two key assumptions. First, all the approaches described in Figure 2B assume that premature RNAs do not undergo degradation, as only mature ones do. This implies that all premature transcripts are, sooner or later, processed into a mature form. This is obviously a simplification, as an RNA exosome complex controlling the degradation of premature RNAs is known to exist in the nucleus [55]. However, the prevalence of this mechanism at the cellular level is difficult to determine. In general, without this assumption, the RNA life cycle model would include more terms and be more complicated, thus requiring additional information to be solved. This issue specifically affects the quantification of synthesis and processing rates.

Second, those approaches, such as INSPECT+ and DRILL, that include the modeling of the RNA processing step, assume that no degradation of the labeled transcripts occurs during the pulse. This assumption is likely to hold when short pulse times are used. Instead, when pulses are long, for example in the range of hours, the likelihood that some of the labeled transcripts are processed into their mature forms and eventually degraded increases. As a result, the quantification of the labeled fraction would lead to the underestimation of the amount of nascent transcription and, consequently, of the rates of synthesis. To



**Table 1.** Overview of methods currently available for the genome-wide quantification of RNA kinetic rates. The following columns are included. ‘Experimental framework’ indicates the experimental focus. ‘Method’ reports the name of the method or, when not available, the first author of the corresponding study. ‘# Conditions’ indicates the minimum number of samples that are necessary, N and 2 N indicating methods which require the sequencing of one or two conditions, respectively, in multiple data points (the higher the N, the higher the quality of the results). ‘Input’ columns indicate the RNA species that have to be quantified (+) or not (-). ‘Output’ columns indicate whether the method allows the quantification of synthesis (k1), processing (k2) or degradation (k3) rates, and specifically whether their absolute values and/or their temporal modulation can be quantified. The ~ symbol indicates that specific limitations occur, as reported in the ‘Limitations’ column. ‘Key concepts’ provides a short overview of the main features. ‘Limitations’ reports key drawbacks. ‘Software’ indicates whether a software is available. ‘Refs’ reports the corresponding reference.

Experimental framework	Method	Conditions	Input		Output			Key concepts	Limitations	Software	Ref.	
			Total	Pre-cent existing	k1	k2	k3					Absolute rates
Total RNA	Block of transcription	N	+	-	-	-	+	-	Models k3 based on RNA exponential decay following the block of transcription	Invasive, disregards RNA processing, requires time course data, provides steady state rates	-	[11]
	EISA	2	+	-	-	+	-	-	Returns k1 and k3 changes between conditions	No absolute rates, disregards RNA processing	+	[19]
	Rembrandts	2	+	-	-	+	-	-	Models the coupling between k1 and k2	No absolute rates, assumes constant k2	+	[20]
	Snapshot-Seq	1	+	-	-	+	~	-	Gene-level k1 and k3, aggregated k2	Assumes constant elongation rate, requires normalization factor to estimate absolute rates	-	[21]
	Zeisel et al.	N	+	-	-	+	+	+	Temporal quantification of k1 and k3	Requires an independent characterization of k2 that is assumed to be constant	-	[6]
Nascent RNA	Wang et al.	4	+	-	-	+	+	-	Compares histone modifications with total RNA-seq data	No info on k2, requires multiple data types	-	[22]
	INSPECt (steady state)	2	+	-	-	-	~	-	Gene-level changes in post-transcriptional regulation between conditions	No info on k1, aggregated quantification of k2 and k3	+	[3]
	INSPECt (time course)	N	+	-	-	+	+	+	Absolute values of all kinetic rates and their modulation over time	Requires time course RNA-seq data	+	[3]
RNA metabolic labeling	Chromatin associated RNA	1	-	-	+	~	-	-	Provides an experimental proxy of k1.	No info on k2 and k3, includes pre-existing transcripts still on chromatin	-	[26]
	GRO-seq	1	-	-	+	~	-	-	Provides an experimental proxy of k1.	No info on k2 and k3, biased by elongation rates.	-	[28]
RNA metabolic labeling	PRO-seq	1	-	-	+	~	-	-	Provides an experimental proxy of k1.	No info on k2 and k3, biased by elongation rates.	-	[29]
	DTA	2-3	+	-	+	+	+	-	k1 and k3 absolute values	No info on k2, requires 3 data types (or 2 if the nascent over pre-existing ratio is provided)	+	[32]
	cDTA	2-3	+	-	+	+	+	-	k1 and k3 absolute values, suitable for comparative studies.	No info on k2, requires 3 data types (or 2 if the nascent over pre-existing ratio is provided)	+	[33]
	pulseR	2-3	+	-	+	+	+	-	Statistical model, handles pulse and chase experiments	No info on k2, requires 3 data types (or 2 with spike-ins)	+	[34]

Table 1. Continued

Experimental framework	Method	Conditions	Input		Output			Key concepts	Limitations	Soft-ware	Ref.
			Total	Pre- existent	k1	k2	k3				
Single cell	DRUID	N	+	+	-	-	+	-	Robust quantification of k3 absolute value without spike-ins	No info on k1 and k2, provides constant rates, laborious, requires multiple conditions	[12]
	DRILL	2 N	+	+	+	+	+	+	Absolute values of all kinetic rates and their modulation over time	Requires time course RNA-seq data	[5]
	INSPECT+	2	+	+	+	+	+	+	Absolute values of all kinetic rates and their modulation over conditions or over time	-	[7]
	Dyrec-seq	2	-	+	+	-	+	-	Leverages sequential labeling with BrU and 4sU	No info on k2, laborious	[42]
	SLAM-DUNK	2	+	-	-	+	+	-	Tailored to the analysis of SLAM-seq data	No info on k1 and k2	[41]
	GRAND-SLAM	1	-	+	-	+	+	-	Tailored to the analysis of SLAM-seq data	No info on k1 and k2	[40]
	RNA Velocity	1	+	-	+	-	+	~	k1 at single cell level, k3 at population level, returns the cells temporal order	No info on k3 at single cell resolution, k2 is assumed to be equal 1, exploits only steady state conditions	[43]
	scVelo	1	+	-	+	~	+	~	k1 at single cell level, k2 and k3 at population level, returns cells temporal order based on their dynamics	No info on k2 and k3 at single cell resolution	[44]
	MetabotASC-seq	1	-	+	+	-	+	-	k1 at single cell resolution	No info on k2 and k3	[45]
	labeling fate	1	-	+	+	+	+	~	k1 at single cell level, k3 at population level, returns the cells temporal progression	No info on k3 at single cell resolution, no info on k2, cells temporal progression analysis relies on a time course	[46]
Single molecule	scSLAM-seq	1	-	+	+	-	+	~	k1 at single cell level, robust temporal order of the cells, takes into account transcriptional bursting	No info on k2 and k3	[47]
	scEU-seq	2 N	-	+	~	~	+	~	k1 and k3 at population level during complex dynamical processes	No info on k1 and k3 at single cell level, no info on k2	[50]
	Metabotano-ID labeling	1	-	+	+	+	+	-	k1 and k3 absolute values at single isoform resolution from a single experiment	No info on k2	[60]

deal with this issue, INSPECT+ allows, at steady state, to consider the processing and degradation of labeled transcripts during the pulse, even though this significantly complicates the mathematical modeling and results in a longer time of analysis and a slight increase in the number of genes that cannot be resolved.

## Future perspectives and conclusions

The quantification of RNA kinetic rates pertains to individual transcripts in individual cells, meaning that each RNA molecule could be distinct, due to subtle changes in its nucleotide composition, structure, and/or set of modifications. Potentially, RNA molecules representing the same transcriptional isoform could have similar rates, compared with different isoforms of the same transcriptional unit. For example, isoforms with retained introns that originate premature termination codons are expected to be associated with high degradation rates, since they are specifically and rapidly degraded through the nonsense-mediated decay pathway [4]. However, approaches for deconvoluting the various transcripts coded by a given gene are still in their infancy [56]. Long-reads sequencing platforms, such as Nanopore, offer an unprecedented opportunity to explore entire transcripts, or large fractions of them [57]. Efforts are underway for coupling RNA metabolic labeling with these sequencing platforms [58, 59]. The nano-ID approach was recently proposed that is able to quantify synthesis and degradation at the level of individual isoforms, based on the identification of nascent RNAs metabolically labeled with 5-Ethynyluridine [60]. Recently, a novel approach was presented applying Nanopore sequencing for the profiling of the transcriptome of individual cells [61]. This advance opens the possibility to study RNA dynamics at the level of single molecules in single cells.

Quantitative modeling of the RNA life cycle is expected to be instrumental in deciphering RNAPII life cycle. Indeed, the action of the RNAPII complex can be broken down into a set of key steps, including RNAPII recruitment, pause-release, elongation and detachment from 3' end sites [62]. Similarly to what happens in the absolute quantification of RNA species by RNA-seq experiments, also RNAPII positioning and quantification by ChIP-seq data provides a static glimpse of a very dynamic process [63]. As for gene expression programs, where the increase in RNA abundance could be similarly obtained through increased synthesis or reduced degradation, an increase in RNAPII at promoters could also originate from increasing recruitment or reduced pause-release. The first attempts at comprehensively quantifying the kinetic rates governing the RNAPII life cycle have been made by taking advantage of the quantification of RNA synthesis rates [17].

Given the intertwined nature of the RNA and RNAPII life cycles, novel methods should aim at jointly modeling the corresponding kinetic rates, potentially at the level of individual cells. The development of methods to couple RNA metabolic labeling with single molecule and single cells sequencing, and the improvement of methods for the profiling of RNAPII, which can now be obtained in few cells [64], suggest that this goal could be reached soon.

In conclusion, numerous experimental and computational methods for the quantification of RNA kinetic rates are being developed. The approaches described in this review are listed in Table 1, providing a reference resource that summarizes their key features and limitations. This emerging field promises to offer an unprecedented understanding of how transcriptional

and post-transcriptional regulation generates complex gene expression programs by fine tuning both its magnitude and its variation over time or across conditions. Furthermore, the determination of RNA kinetic rates is expected to be precious in the characterization of the functional role of RNA modifications [65, 66] and RNA binding proteins [67], which are fundamental determinants in the fate of a transcript.

### Key Points

- Gene expression programs are set by the balance of cellular machineries dedicated to RNA synthesis, processing and degradation, and the abundance and modulation of RNAs are determined by the tight regulation of the rates governing the kinetics of these key steps.
- Various experimental and computational methods exist to study RNA dynamics, allowing the quantification of RNA kinetic rates through the profiling and integrative analysis of nascent and pre-existing RNAs.
- Additional approaches are available that only require the profiling of total RNA, while they depend on specific assumptions or they require time course experiments.
- Approaches are being developed that allow studying RNA dynamics at the level of single molecules and single cells.

## Funding

This work was supported by a grant from the Italian Association for Cancer Research (AIRC), project IG 2020 (ID. 24784) to M.P.

## References

1. Cramer P. Eukaryotic transcription turns 50. *Cell* 2019;179:808–12.
2. Edgar R, Domrachev M, Lash AE. Gene expression omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* 2002;30:207–10.
3. Furlan M, Galeota E, Gaudio ND, et al. Genome-wide dynamics of RNA synthesis, processing, and degradation without RNA metabolic labeling. *Genome Res* 2020;30:1492–507.
4. Houseley J, Tollervey D. The many pathways of RNA degradation. *Cell* 2009;136:763–76.
5. Rabani M, Raychowdhury R, Jovanovic M, et al. High-resolution sequencing and modeling identifies distinct dynamic RNA regulatory strategies. *Cell* 2014;159:1698–710.
6. Zeisel A, Köstler WJ, Molotski N, et al. Coupled pre-mRNA and mRNA dynamics unveil operational strategies underlying transcriptional responses to stimuli. *Mol Syst Biol* 2011;7:529.
7. de Pretis S, Kress T, Morelli MJ, et al. INSPECT: a computational tool to infer mRNA synthesis, processing and degradation dynamics from RNA- and 4sU-seq time course experiments. *Bioinformatics (Oxford, England)* 2015;31:2829–35.
8. Lam LT, Pickeral OK, Peng AC, et al. Genomic-scale measurement of mRNA turnover and the mechanisms of action of the anti-cancer drug flavopiridol. *Genome Biol* 2001;2:RESEARCH0041.

9. Raghavan A, Ogilvie RL, Reilly C, et al. Genome-wide analysis of mRNA decay in resting and activated primary human T lymphocytes. *Nucleic Acids Res* 2002;**30**:5529–38.
10. Yang E, van Nimwegen E, Zavolan M, et al. Decay rates of human mRNAs: correlation with functional characteristics and sequence attributes. *Genome Res* 2003;**13**:1863–72.
11. Chen H, Shiroguchi K, Ge H, et al. Genome-wide study of mRNA degradation and transcript elongation in *Escherichia coli*. *Mol Syst Biol* 2015;**11**:781.
12. Lugowski A, Nicholson B, Rissland OS. DRUID: a pipeline for transcriptome-wide measurements of mRNA stability. *RNA* 2018;**24**:623–32.
13. Presnyak V, Alhusaini N, Chen Y-H, et al. Codon optimality is a major determinant of mRNA stability. *Cell* 2015;**160**:1111–24.
14. Wada T, Becskei A. Impact of methods on the measurement of mRNA turnover. *Int J Mol Sci* 2017;**18**:2723–14.
15. Rabani M, Levin JZ, Fan L, et al. Metabolic labeling of RNA uncovers principles of RNA production and degradation dynamics in mammalian cells. *Nat Biotechnol* 2011;**29**:436–42.
16. Sammeth M, Foissac S, Guigó R. A general definition and nomenclature for alternative splicing events. *PLoS Comput Biol* 2008;**4**:e1000147.
17. de Pretis S, Kress TR, Morelli MJ, et al. Integrative analysis of RNA polymerase II and transcriptional dynamics upon MYC activation. *Genome Res* 2017;**27**:1658–64.
18. Adiconis X, Borges-Rivera D, Satija R, et al. Comparative analysis of RNA sequencing methods for degraded or low-input samples. *Nat Methods* 2013;**10**:623–9.
19. Gaidatzis D, Burger L, Florescu M, et al. Analysis of intronic and exonic reads in RNA-seq data characterizes transcriptional and post-transcriptional regulation. *Nat Biotechnol* 2015;**33**:722–9.
20. Alkallas R, Fish L, Goodarzi H, et al. Inference of RNA decay rate from transcriptional profiling highlights the regulatory programs of Alzheimer's disease. *Nat Commun* 2018;**8**:1–11.
21. Gray JM, Harmin DA, Boswell SA, et al. SnapShot-Seq: a method for extracting genome-wide, in vivo mRNA dynamics from a single total RNA sample. *PLoS One* 2014;**9**:e89673.
22. Wang C, Tian R, Zhao Q, et al. Computational inference of mRNA stability from histone modification and transcriptome profiles. 2012;**40**:6414–23.
23. de Pretis S, Furlan M, Pelizzola M. INSPECT-GUI reveals the impact of the kinetic rates of RNA synthesis, processing, and degradation, on premature and mature RNA species. *Front Genet* 2020;**11**:230.
24. Tesi A, de Pretis S, Furlan M, et al. An early Myc-dependent transcriptional program orchestrates cell growth during B-cell activation. *EMBO Rep* 2019;**20**:e47987.
25. Dolken L, Ruzsics Z, Radle B, et al. High-resolution gene expression profiling for simultaneous kinetic parameter analysis of RNA synthesis and decay. *RNA* 2008;**14**:1959–72.
26. Wissink EM, Vihervaara A, Tippens ND, et al. Nascent RNA analyses: tracking transcription and its regulation. *Nat Rev Genet* 2019;**20**:1–19.
27. Biasini A, Marques AC. A protocol for Transcriptome-wide inference of RNA metabolic rates in mouse embryonic stem cells. *Front Cell Dev Biol* 2020;**8**:97.
28. Core LJ, Waterfall JJ, Lis JT. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science* 2008;**322**:1845–8.
29. Kwak H, Fuda NJ, Core LJ, et al. Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. *Science* 2013;**339**:950–3.
30. Schwalb B, Michel M, Zacher B, et al. TT-seq maps the human transient transcriptome. *Science* 2016;**352**:1225–8.
31. Uvarovskii A, Naarmann-de Vries IS, Dieterich C. On the optimal design of metabolic RNA labeling experiments. *PLoS Comput Biol* 2019;**15**:e1007252.
32. Schwalb B, Schulz D, Sun M, et al. Measurement of genome-wide RNA synthesis and decay rates with dynamic transcriptome analysis (DTA). 2012;**28**:884–5.
33. Sun M, Schwalb B, Schulz D, et al. Comparative dynamic transcriptome analysis (cDTA) reveals mutual feedback between mRNA synthesis and degradation. *Genome Res* 2012;**22**:1350–9.
34. Uvarovskii A, pulseR DC. Versatile computational analysis of RNA turnover from metabolic labeling experiments. *Bioinformatics (Oxford, England)* 2017;**33**:3305–7.
35. Baptista MAP, Dölken L. RNA dynamics revealed by metabolic RNA labeling and biochemical nucleoside conversions. *Nat Methods* 2018;**15**:171–2.
36. Herzog VA, Reichholf B, Neumann T, et al. Thiol-linked alkylation of RNA to assess expression dynamics. *Nat Methods* 2017;**14**:1198–204.
37. Schofield JA, Duffy EE, Kiefer L, et al. TimeLapse-seq: adding a temporal dimension to RNA sequencing through nucleoside recoding. *Nat Methods* 2018;**15**:221–5.
38. Riml C, Amort T, Rieder D, 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**:13479–83.
39. Gasser C, Delazer I, Neuner E, et al. Thioguanosine conversion enables mRNA life-time evaluation by RNA sequencing via double metabolic labeling. *Angew Chem Int Ed Engl* 2020;**20**:631.
40. Jürges C, Dölken L, Erhard F. Dissecting newly transcribed and old RNA using GRAND-SLAM. *Bioinformatics* 2018;**34**:i218–26.
41. Neumann T, Herzog VA, Muhar M, et al. Quantification of experimentally induced nucleotide conversions in high-throughput sequencing datasets. *BMC Bioinform* 2019;**20**:258.
42. Kawata K, Wakida H, Yamada T, et al. Metabolic labeling of RNA using multiple ribonucleoside analogs enables the simultaneous evaluation of RNA synthesis and degradation rates. *Genome Res* 2020;**gr.264408**:120–41.
43. La Manno G, Soldatov R, Zeisel A, et al. RNA velocity of single cells. *Nature* 2018;**560**:1–25.
44. Bergen V, Lange M, Peidli S, et al. Generalizing RNA velocity to transient cell states through dynamical modeling. *Nat Biotechnol* 2020;**14**:e8046.
45. Hendriks G-J, Jung LA, Larsson AJM, et al. NASC-seq monitors RNA synthesis in single cells. *Nat Commun* 2019;**10**:3138–9.
46. Cao J, Zhou W, Steemers F, et al. Sci-fate characterizes the dynamics of gene expression in single cells. *Nat Biotechnol* 2020;**38**:980–8.
47. Erhard F, Baptista MAP, Krammer T, et al. scSLAM-seq reveals core features of transcription dynamics in single cells. *Nature* 2019;**571**:419–23.
48. Qiu Q, Hu P, Qiu X, et al. Massively parallel and time-resolved RNA sequencing in single cells with scNT-seq. *Nat Methods* 2020;**29**:1–35.

49. Qiu X, Zhang Y, Yang D, et al. Mapping vector field of single cells. *bioRxiv* 2019;37:L25–35.
50. Battich N, Beumer J, de Barbanson B, et al. Sequencing metabolically labeled transcripts in single cells reveals mRNA turnover strategies. *Science* 2020;367:1151–6.
51. Tantale K, Mueller F, Kozulic-Pirher A, et al. A single-molecule view of transcription reveals convoys of RNA polymerases and multi-scale bursting. *Nat Commun* 2016;7:12248–14.
52. Wachutka L, Caizzi L, Gagneur J, et al. Global donor and acceptor splicing site kinetics in human cells. *Elife* 2019;8.
53. Herzl L, Ottoz DSM, Alpert T, et al. Splicing and transcription touch base: co-transcriptional spliceosome assembly and function. *Nat Rev Mol Cell Biol* 2017;18:637–50.
54. Neugebauer KM. Nascent RNA and the coordination of splicing with transcription. *Cold Spring Harb Perspect Biol* 2019;11:a032227–15.
55. Kilchert C, Wittmann S, Vasiljeva L. The regulation and functions of the nuclear RNA exosome complex. *Nat Rev Mol Cell Biol* 2016;17:227–39.
56. Zhang C, Zhang B, Lin L-L, et al. Evaluation and comparison of computational tools for RNA-seq isoform quantification. *BMC Genomics* 2017;18:583–11.
57. Workman RE, Tang AD, Tang PS, et al. Nanopore native RNA sequencing of a human poly(a) transcriptome. *Nat Methods* 2019;16:1297–305.
58. Furlan M, Tanaka I, Leonardi T, et al. Direct RNA sequencing for the study of synthesis, processing, and degradation of modified transcripts. *Front Genet* 2020;11:394.
59. Reimer KA, Neugebauer KM. Preparation of mammalian nascent RNA for long read sequencing. *Curr Protoc Mol Biol* 2020;133:421–5.
60. Maier KC, Gressel S, Cramer P, et al. Native molecule sequencing by nano-ID reveals synthesis and stability of RNA isoforms. *Genome Res* 2020;9:1332–44.
61. Lebrigand K, Magnone V, Barbry P, et al. High throughput error corrected Nanopore single cell transcriptome sequencing. *Nat Commun* 2020;11:4025–8.
62. Jonkers I, Lis JT. Getting up to speed with transcription elongation by RNA polymerase II. *Nat Rev Mol Cell Biol* 2015;16:167–77.
63. Ehrensberger AH, Kelly GP, Svejstrup JQ. Mechanistic interpretation of promoter-proximal peaks and RNAPII density maps. *Cell* 2013;154:713–5.
64. Liu B, Xu Q, Wang Q, et al. The landscape of RNA pol II binding reveals a stepwise transition during ZGA. *Nature* 2020;587:139–44.
65. Furlan M, Galeota E, de Pretis S, et al. m6A-dependent RNA dynamics in T cell differentiation. *Genes* 2019;10.
66. Roundtree IA, Evans ME, Pan T, et al. Dynamic RNA modifications in gene expression regulation. *Cell* 2017;169:1187–200.
67. Glisovic T, Bachorik JL, Yong J, et al. RNA-binding proteins and post-transcriptional gene regulation. *FEBS Lett* 2008;582:1977–86.
68. Mukherjee N, Calviello L, Hirsekorn A, et al. Integrative classification of human coding and noncoding genes through RNA metabolism profiles. *Nat Struct Mol Biol* 2017;24:86–96.