Dynamics of transcriptional regulation from total RNA-seq experiments

Mattia Furlan[1,2,*], Stefano de Pretis[1,*,#], Eugenia Galeota[1], Michele Caselle[2], Mattia Pelizzola[1,#]

[1] Center for Genomic Science, Fondazione Istituto Italiano di Tecnologia, Milan, Italy

[2] Physics Department and INFN, University of Turin, Turin, Italy

* These authors equally contributed

Corresponding author

Abstract

The kinetic rates of RNA synthesis, processing and degradation determine the dynamics of transcriptional regulation by governing both the abundance and the responsiveness to modulations of premature and mature RNA species. The study of RNA dynamics is largely based on the integrative analysis of total and nascent transcription, with the latter being quantified through RNA metabolic labelling. We describe here a computational method, based on mathematical modelling of intronic and exonic expression, able to derive the dynamics of transcription from steady-state or time course profiling of just total RNA, without requiring any information on nascent transcripts. Our approach closely recapitulates the kinetic rates obtained through RNA metabolic labelling, reduces the cost and complexity of the experiments, and can be adopted to study experimental conditions where nascent transcription cannot be readily profiled. We applied this method to the characterization of post-transcriptional regulation landscapes in dozens of physiological and disease conditions, and we revealed a previously unanticipated role for the kinetics of RNA processing in the modulation of RNA responsiveness.

Main

Since the development of microarrays first, and high-throughput sequencing later on, the investigation of the transcriptional activity of genes has been mostly based on the quantification of total RNA[1]. While bringing about a revolution in the field of transcriptional regulation, these approaches have failed to consider that the abundance of premature and mature RNA depends on the RNA life-cycle, whose three main steps are: premature RNA synthesis, processing of premature into mature RNA, and degradation of the latter[2]. These steps are governed by corresponding kinetic rates, which collectively determine the RNA dynamics of transcripts (Fig. 1A). At steady-state, the abundance of each premature RNA is equal to the ratio of its synthesis to processing rate, and the quantity of its mature form is given by its synthesis to degradation rate ratio (Fig. 1B). Thus, while the rate of RNA synthesis influences the abundance of both premature and mature RNAs, processing and degradation rates impact just on premature and mature forms, respectively. At the transition between steady-states, RNA kinetic rates define the speed at which the mature form of a transcript can be brought to a new level of abundance (here denoted as responsiveness). The degradation rate is typically considered the major determinant of responsiveness (the lower the transcript stability, the higher its responsiveness)[3,4].

For decades, RNA dynamics were mostly studied through transcription blockage experiments, but these were highly invasive methods that affected cell viability and could alter various pathways, RNA decay included. To overcome these limitations, new methods have been developed that are based on the integrative analysis of total and nascent RNA[5-7]. Nascent RNA can be marked by metabolic labelling with biotinylated 4-thiouridine (4sU) modified nucleotides and then purified by streptavidin to be sequenced[5-7]. Alternatively, after chemical derivatization and sequencing, reads from nascent transcripts can be *in silico* separated from pre-existing RNA[8-11]. These approaches have started to unveil how the combined modulation of the kinetic rates can determine gene-specific regulatory modes and elicit complex transcriptional responses[12-14].

Despite its key methodological advances, RNA metabolic labelling is not exempt from technical pitfalls, involving mostly 4sU incorporation, purification of labelled molecules, or chemical

derivatization. Moreover, these methods cannot be readily applied to model organisms, mammals[15] or plants *in vivo*. Finally, to correctly quantify nascent RNA after purification, twice the amount of sequencing and careful normalization between pre-existing and nascent RNA data are required. On the other hand, when opting for short 4sU pulses followed by chemical derivatization of labelled RNA molecules, a significantly larger sequencing coverage is required. For these reasons, the possibility to study RNA dynamics from just total RNA would be a valuable alternative. A few studies have moved in this direction by using the integrative analysis of premature and mature RNA abundances[3,16-18], yet they have fallen short of quantifying the full set of RNA kinetic rates.

Overcoming these limitations, we describe here a computational approach that permits to study RNA dynamics from total RNA-seq experimental data. The tool, available within the INSPEcT Bioconductor package[12], provides a full set of kinetic rates from time course RNA-seq datasets. Moreover, it enables the study of post-transcriptional regulation across steady-state conditions. In this work, we apply this method to the analysis of multiple time-course RNA-seq datasets to cover various transcriptional and post-transcriptional regulation scenarios. Moreover, we provide the first analysis of RNA dynamics in plants. Finally, we characterized post-transcriptional regulation landscapes, and shed light on the functional role of processing dynamics in dozens of tissues types and disease conditions.

Results

The modulation of RNA dynamics enables complex transcriptional responses

The regulation of the cellular abundance of mature (M) and premature (P) RNA species can simply derive from changes in the rate of synthesis of premature RNA (k_1), or can entail more complex co- and post-transcriptional mechanisms governed by the premature-to-mature RNA processing rate (k_2) and/or the mature RNA degradation rate (k_3). We have developed a Shiny user interface (included within INSPEcT) to explore how alternative modes of transcriptional regulation can be generated through the combined modulation of kinetic rates. Briefly: constant kinetic rates define steady-states

where P and M abundances are calculated as k_1/k_2 and k_1/k_3 ratios, respectively (Fig. 1B,C); modulations in the processing rate k_2 cause just transient variations in M abundance but permanent alterations of P abundance (Fig. 1D); M responsiveness to k_1 adjustments depends on the level of k_3 (compare Fig. 1E and F) and can be reduced by decreasing k_2 (Fig. 1G); k_1 and k_3 can separately generate the same type of M variation if changing in opposite directions (Fig. 1F,H), while adjustments in k_1 only are able to effect P abundance (Fig. 1F); k_1 and k_3 reinforce each other's modulation of M (compare Fig. 1E with I) when they simultaneously change in opposite directions, but generate just a modulation of P if they are simultaneously adjusted in the same direction (Fig. 1J); a transient alteration of M that has been induced by a temporary change in k_1 (Fig. 1K) can be made sharper by a concomitant change in k_3 (Fig. 1L, as discussed in [6]).

Based on these results, we reasoned that the temporal quantification of premature and mature RNA species should allow the deconvolution of the underlying RNA dynamics.

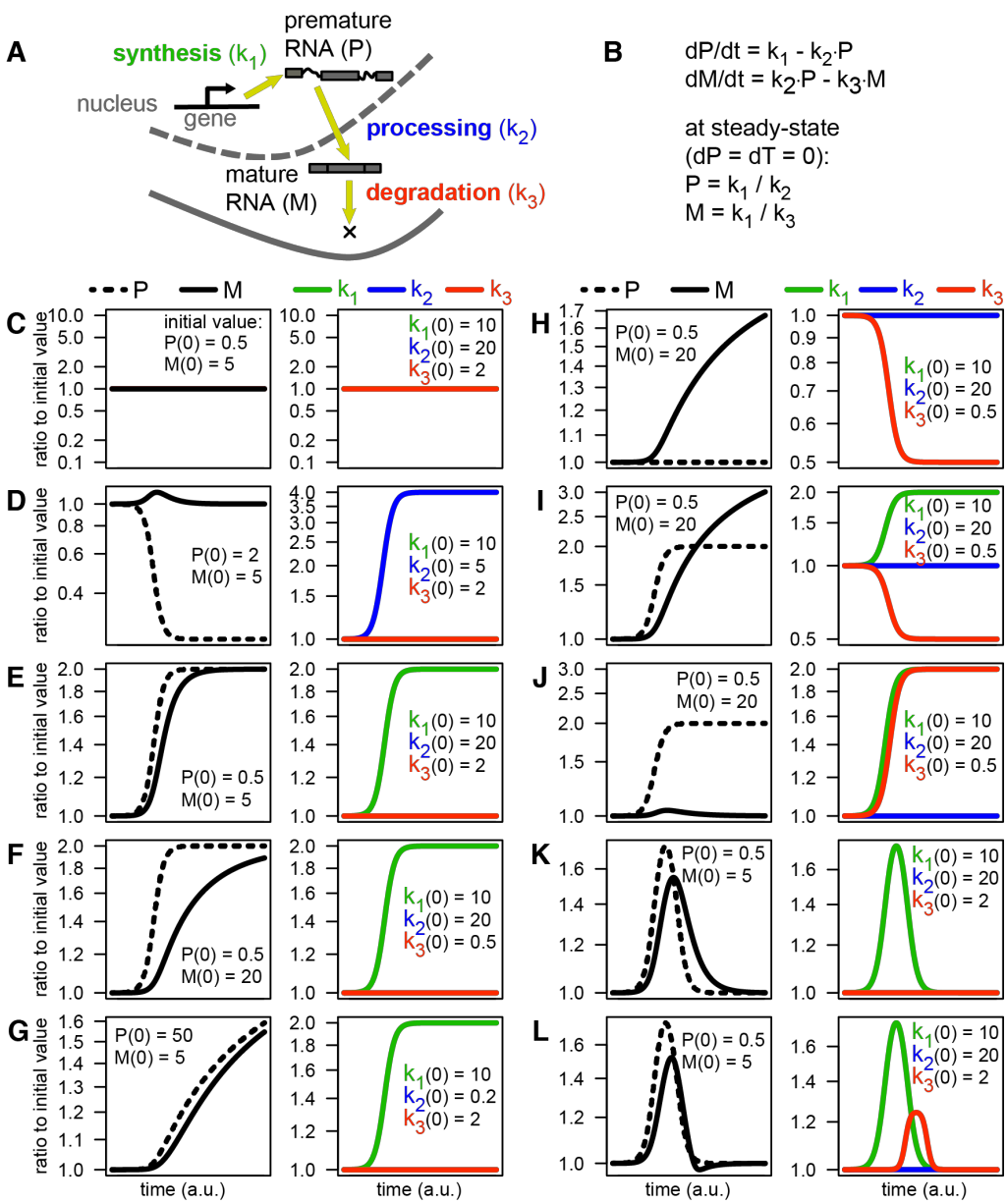

Figure 1 The influence of RNA kinetic rates on RNA abundance and responsiveness. (A) Schematic representation of the RNA life cycle, governed by the kinetics rates of synthesis, processing and degradation. (B) Deterministic mathematical model of the RNA life cycle based on Ordinary Differential Equations (ODEs), including the solution of the system at steady state. (C-L) Solution of the ODE system following the modulation of the kinetic rates: each example reports, for premature and mature RNA species (left) and for the kinetic rates (right), the ratio to the initial time point. Initial values are indicated within each panel.

RNA-dynamics from time-course total RNA-seq data

We developed a computational approach able to quantify RNA dynamics based on time-course profiling of total RNA-seq data. Briefly, eight models corresponding to all possible combinations of

each kinetic rate in two alternative analytical forms (constant or impulsive) were considered. Each model was plugged within a system of ordinary differential equations (Fig. 1B). Optimization of the free parameters associated with the rates' functional forms was performed to minimize the error in the fit of the premature and mature RNAs time-course profiles. Finally, a model was selected that gave the best trade-off between complexity and goodness of fit. As a faster alternative, we developed a derivative approach based on an analytical solution of the system, allowing the deconvolution of gene-specific RNA dynamics in 20s per core, while minimally compromising on the quality of the results (Supplementary Fig. 1).

We validated the kinetic rates based on a simulated dataset of 1000 genes. Time-course averages of synthesis rates correlated very well with expected values (0.81 Spearman correlation), while processing and degradation rates showed lower but significant correlations (0.41 and 0.59, respectively; Fig. 2A). Changes of the modelled kinetics rates over time closely recapitulated the expected response (Fig. 2B,C). The ability of our model selection procedure to correctly classify variable rates was evaluated by ROC analyses and found to perform well (AUC 0.68-0.80; Fig. 2D). These results were in line with those obtained when including nascent RNA, and gave a robust response to a reduction in the number of time points and replicates (Fig. 2D-E and Supplementary Fig. 1). Finally, INSPEcT also produced an excellent estimate of the k_2/k_3 ratio, whose value works as an indicator of post-transcriptional regulation (last row in Fig. 2). These data indicated that the rates' absolute values and their changes over time could be satisfactorily estimated even in the absence of nascent RNA data.

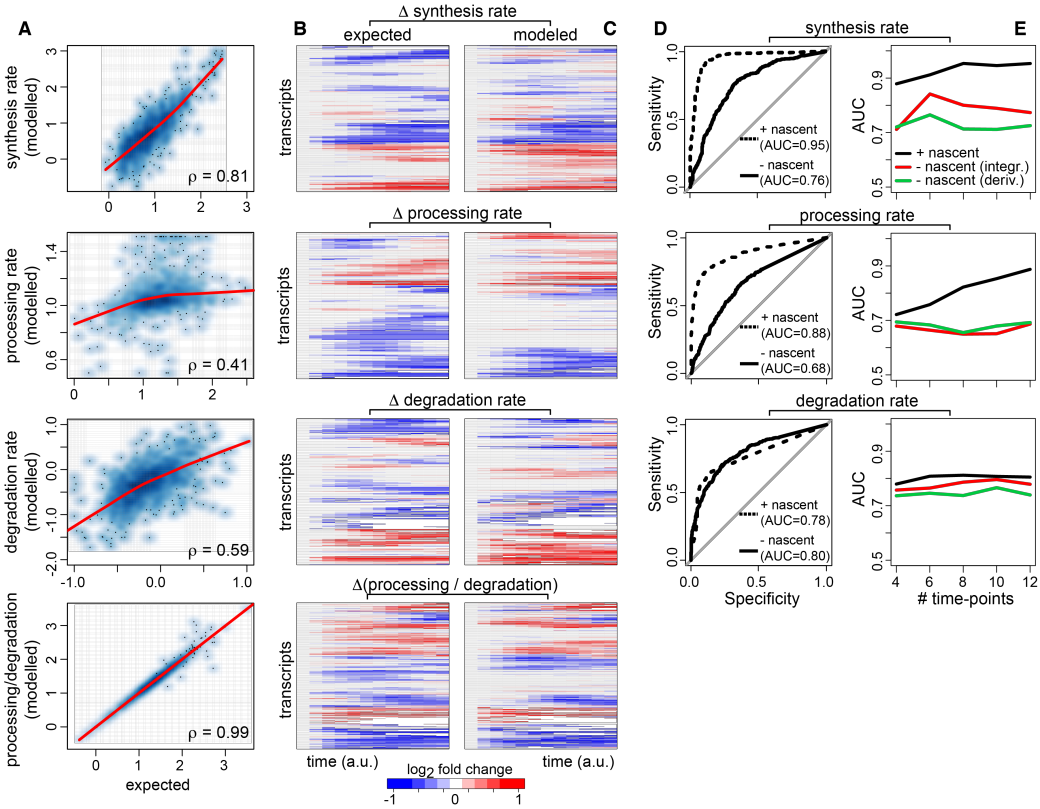

Figure 2 Validation of the kinetic rates determined with time course simulated RNA-seq data. *In silico* modulation of the (expected) kinetic rates provided the temporal expression of nascent and total (premature and mature) RNA for 1000 genes. The abundance of total RNAs was used to derive the kinetic rates (modelled) without using nascent RNAs data. (A) Modelled and expected kinetic rates were averaged over time for each gene, and compared through a density scatter plot. Regression curve and Spearman's correlation coefficient are indicated within each panel. (B) Log2 fold changes of expected kinetic rates, relative to the initial time point. (C) As in (B) for the modelled rates. (D) Sensitivity and specificity in the classification of variable kinetic rates with or without nascent RNA. The area under the curve (AUC) is reported within each panel. (E) AUCs obtained at increasing number of time points, each including three replicates. Both the integrative and the derivative modelling approach were tested.

We used INSPEcT to re-analyse three publicly available RNA-seq time-course datasets, corresponding to conditions with increasing levels of post-transcriptional regulation. Firstly, we focused on the temporal response to MYC activation in 3T9 mouse fibroblasts, which we had recently characterized[14]. MYC is expected to act primarily by modulating transcription of its target genes[14,19]. Indeed, a change in RNA synthesis was observed for 95% of the MYC modulated genes, while an alteration of either processing or degradation rates was seen in 41% of cases (Fig. 3A, Supplementary

Fig. 2). Importantly, this dataset also revealed a high correlation between estimated kinetic rates and those derived from the integrative analysis of total and nascent RNA-seq data (0.91, 0.48 and 0.69 Spearman correlation for synthesis, processing and degradation rates, respectively, Supplementary Fig. 3). Secondly, we reanalysed the temporal polarization of CD4+ cells with (Th17) or without (Th0) polarizing cytokines[20]. As expected, in comparison with the response elicited by a master transcription factor of the likes of MYC, more genes were modulated through post-transcriptional regulation (67% of genes, Fig. 3A, Supplementary Fig. 2). Key regulators of this process were permanently or temporarily modulated in Th17 cells, while changing only transiently in Th0 control cells (Fig. 3B). Our analyses revealed underlying mechanisms of regulation that rely on the control of RNA synthesis in the case of the RORC master regulator and of post-transcriptional regulation in that of SATB1 (Fig. 3B). Next, we analysed the time-course response to the activation of two microRNAs[21]. We expected to see a strong post-transcriptional regulation of the miRNA target transcripts and, indeed, these were seen to be primarily controlled at the level of their stability, while non-target transcripts remained mostly unaffected (Fig. 3C). Finally, we provided the first analysis of RNA dynamics in plants by focusing on the temporal response to ethylene in *Arabidopsis thaliana*[22]. We modelled RNA dynamics for 564 genes, 81 of which were found to be modulated at the level of total or premature RNA (Fig. 3E). Responsive genes were divided into four clusters according to their RNA kinetic rates in the untreated condition. The first cluster included genes involved in cellular respiration, with high rates of both synthesis and degradation, denoted by high responsiveness (Fig. 3D). 70 genes were found to be regulated only through changes in RNA synthesis, while 11 genes were exclusively regulated post-transcriptionally (Fig. 3E). Notably, the latter included AT1G79700 (WRI4), a newly identified factor of the ethylene signalling pathway, which we revealed to be specifically regulated through an increase in its RNA stability (Fig. 3E).

Altogether, these analyses illustrate in what ways the quantification of RNA dynamics from total RNA-seq datasets can unveil the underlying mechanisms controlling premature and mature RNA abundances as well as their variations.

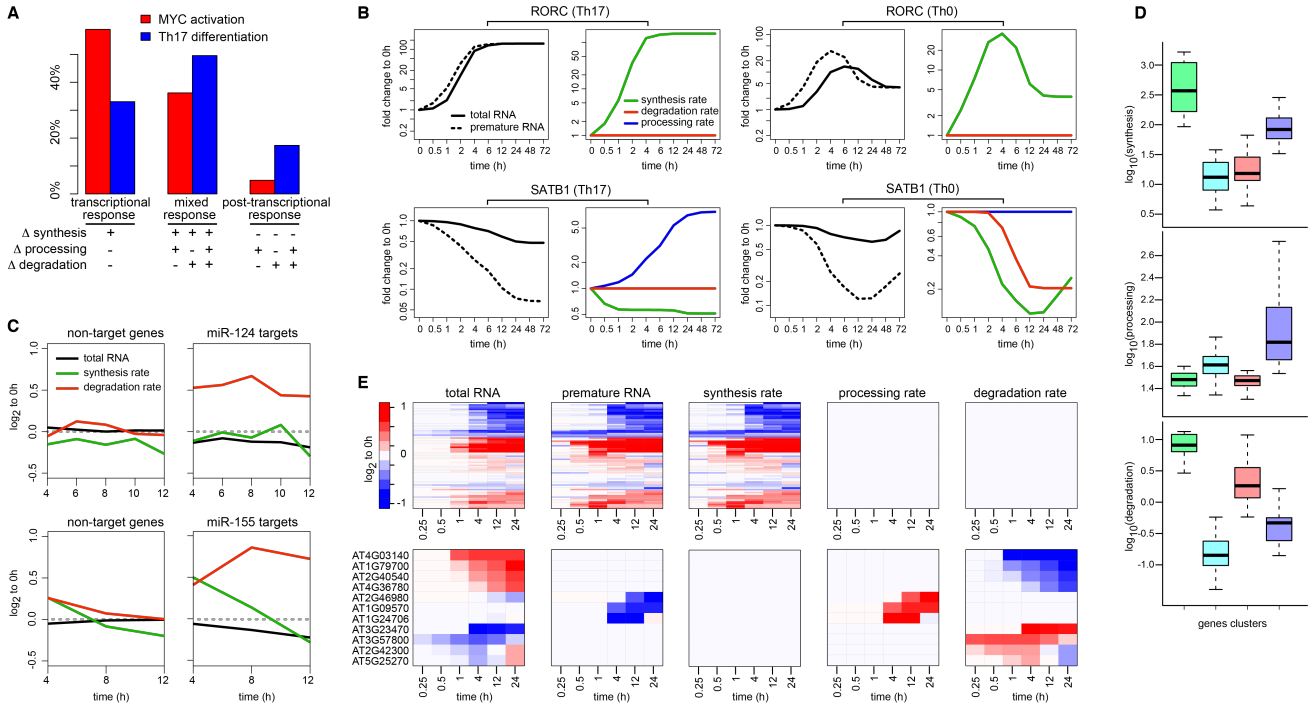

Figure 3 Characterization of time course RNA dynamics: reanalysis of four published datasets. (A) Percentage of genes that, following the activation of MYC and T cell differentiation, are subjected to the regulation of synthesis (transcriptional response), processing and/or degradation (post-transcriptional response) or both (mixed transcriptional and post-transcriptional response). (B) Fold change profiles of premature and total RNA abundance and the RNA kinetic rates for two key regulators of human T cell differentiation (RORC and SATB1) in Th17 and control Th0 cells. (C) Median Log2 fold changes for total RNA, synthesis and degradation rates in response to the expression of miR-124; miRNA targets and non-targeted genes are compared. (D) *A. thaliana* genes were divided in four groups based on the RNA kinetic rates; the distribution of each rate within each cluster is shown in the boxplots. (E) RNA dynamics of *A. thaliana* genes modulated in at least one kinetic rate following ethylene treatment; changes in premature and total RNA abundance and the kinetic rates compared to the untreated condition are shown.

RNA-dynamics from steady-state total RNA-seq data

At steady-state and in the absence of nascent RNA profiling, no information is available on the rate of synthesis. However, the ratio of premature to mature RNA abundance is equal to the ratio of processing to degradation rate (k_2/k_3, Fig. 1B). While this ratio does not allow deconvoluting the individual contributions of the two rates, its change over different conditions indicates alterations in post-transcriptional regulation.

Based on the above considerations, we used INSPEcT to characterize the landscape of post-transcriptional regulation with an unprecedented breadth, covering 35.000 genes in more than 600 samples, which we assigned to specific tissue types (26 tissues) and disease conditions (24 diseases) using the Onassis Bioconductor package[23]. We focused on RNA-seq datasets depleted of ribosomal RNA species and therefore enriched of both pre- and mature RNAs. Moreover, we relied on RNA-seq coverage data that had been homogeneously reanalysed by the recount2 project[24], thus minimizing potential batch effects due to different analysis pipelines and normalization methods. What we found is that the amount of premature RNA increases with the abundance of mature RNA following a power-law that is substantially different depending on the type of gene: protein coding, pseudo or long non-coding (Fig. 4A). Significant deviations from these trends point to post-transcriptionally regulated genes.

Differential post-transcriptional regulation heatmaps revealed that changes in the k_2/k_3 ratio automatically grouped together samples from similar types of tissues and diseases (Fig. 4B, Supplementary Fig. 4). This suggested that post-transcriptional regulation is coordinated across similar conditions, and revealed that some types of cells have a tendency to be markedly subjected to post-transcriptional regulation (Supplementary Fig. 5). Interestingly, ~30% of the information contained in the clustering derived from the post-transcriptional regulation category, and could not be obtained based on tissue-specific expression patterns (Supplementary Fig. 6). A confirmation of the consistency of our method came from the finding that several classes of miRNA targets were significantly enriched in genes found to be post-transcriptionally regulated (see methods). The frequency of post-transcriptional regulation varied significantly, with protein coding and pseudo genes being more regulated than non-coding (Fig. 4B). The 1000 protein-coding genes with the lowest frequencies of post-transcriptional regulation were found to be associated with basic cellular processes such as protein folding, organelle organization and metabolic processes. On the contrary, the 1000 genes with the highest frequencies turned out to be related either to various diseases, cancer included, or to highly specific biological processes, such as B-cell activation, autoimmune response, differentiation and morphology.

We analysed more closely the functionality of the genes undergoing post-transcriptional regulation under specific conditions. Genes altered in T-cell samples were associated with the regulation of T-cell number and proliferation and with immunodeficiency, and often targeted by the E2F1 transcription factor, the master regulator in T-cell proliferation[25]. Genes altered in heart samples were associated with cardiac hypertrophy, abnormal contractility and cardiomyopathy. Indeed, a subset of these samples could be associated with the cardiomyopathy disease. Finally, genes altered in brain samples were associated with several diseases including glioma, autism, neoplasm of the nervous system, with biological processes such as hormone secretion and synaptic transmission, and with the 5HT2 type receptor[26], a G protein-coupled receptor that binds serotonin and is widely expressed in the central nervous system where it mediates fundamental processes.

Altogether, these results illustrate how to obtain important information from the study of RNA dynamics when individual conditions are compared in the absence of nascent RNA data.

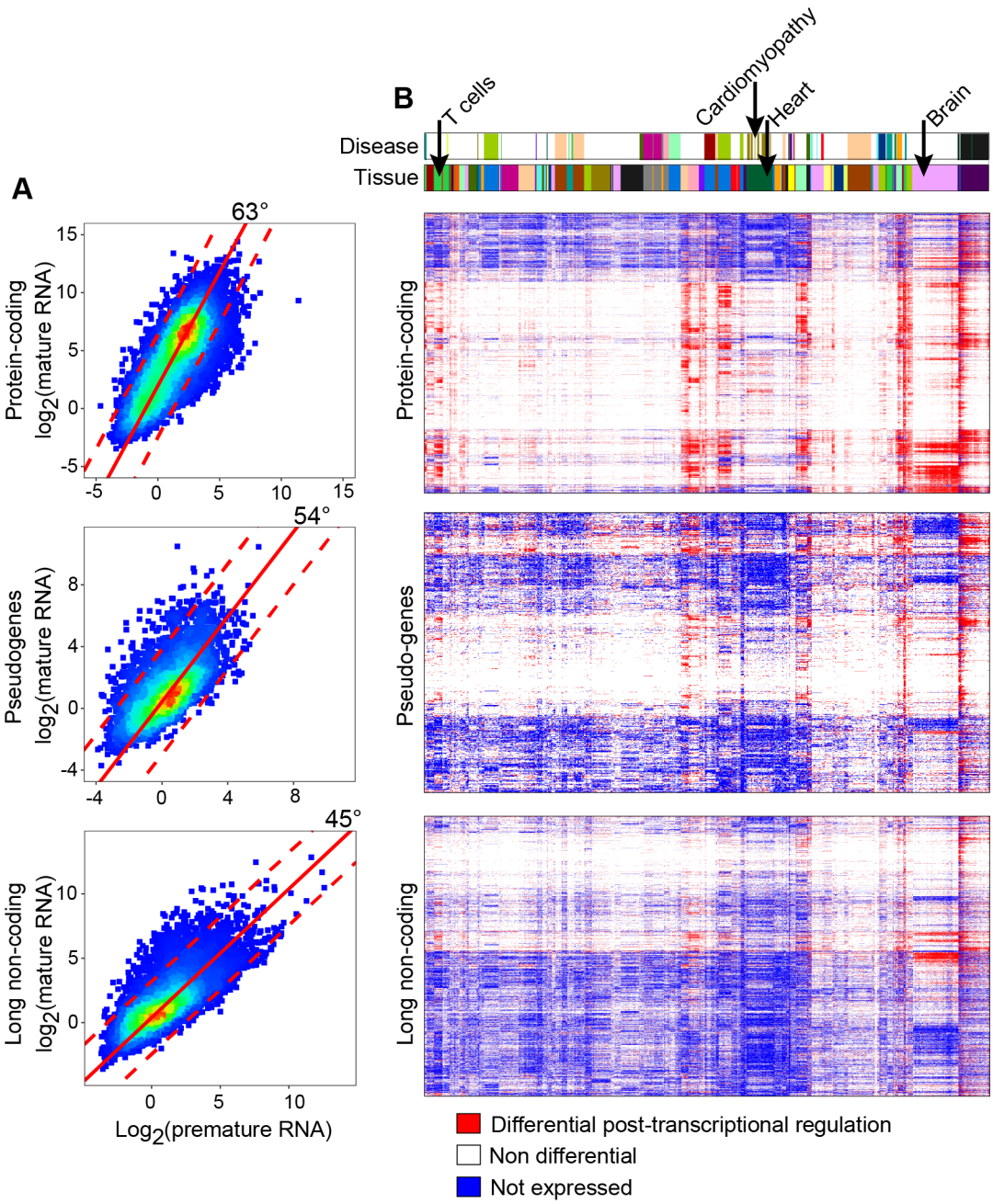

Figure 4 Characterization of steady-state RNA dynamics: reanalysis of 620 RNA-seq datasets. (A) Median abundances of premature and mature RNAs per gene were compared for protein-coding (top panel), pseudo- (middle) and long non-coding genes (bottom). Density scatter plot were fitted with a linear model, whose slope is reported. (B) Heatmaps displaying the degree of post-transcriptional regulation for each gene (row) in each sample (column). The ratio between premature and mature RNA abundance for each genes in each sample were determined and compared to the global trend depicted in (A). Each gene is either not expressed (blue), not differentially post-transcriptional regulated (white; ratio between the dashed lines in (A)), or differentially post-transcriptional regulated (red; ratio above the dashed lines). Above the heatmaps, two colourbars indicate the tissue-type and disease conditions of each sample.

The role of processing dynamics in the responsiveness of mature RNA

The integrative analysis of premature and mature RNA species can provide information on the post-transcriptional dynamics (Fig. 4), and, more specifically, on the influence that processing rates can have on RNA responsiveness. Indeed, while RNA stability is currently considered to be the major determinant of responsiveness[3], our analyses indicate that this can also be affected by RNA processing (Fig. 5A).

To measure the impact of processing on responsiveness, we devised two metrics: (i) the additional time required (τ) - in comparison to a scenario where processing occurs instantaneously - to see a two-fold increase in mature RNA levels following a doubling in synthesis rates, and (ii) the number of RNA molecules involved in this delay (Δ). These metrics can be analytically calculated in INSPEcT using the kinetic rates of individual genes, and can be approximated with high level of accuracy using premature and mature RNA abundances (τ= 1+P/M, and Δ= P/2; 0.999 and 0.993 Spearman correlation with the analytical solutions, respectively; Supplementary Fig. 7). The responsiveness of genes scoring high in both metrics (Δ>1 and τ>1.5) is thus significantly dampened by the processing step. In physiological conditions (3T9 mouse fibroblast cells[14]), processing rates were found to be extremely rapid in comparison to synthesis and degradation rates (Fig. 5B), and had to be significantly reduced for a substantial number of genes to become affected. For example, around 10% of genes were impacted by a 4-fold reduction in their processing rates (Fig. 5C). Nonetheless, the impact of a reduced processing rate was markedly dependent on the values of the other two kinetic rates. Indeed, halving processing rates impacted 10% of the genes, if combined with a two-fold increase of both synthesis and degradation rates (Fig. 5C). It is currently unclear whether physiological or pathological conditions exist where RNA maturation can be slowed down to the point of becoming an impediment to the fast response required in cases such as stress response. This prompted us to characterize this phenomenon in various cell types and disease conditions.

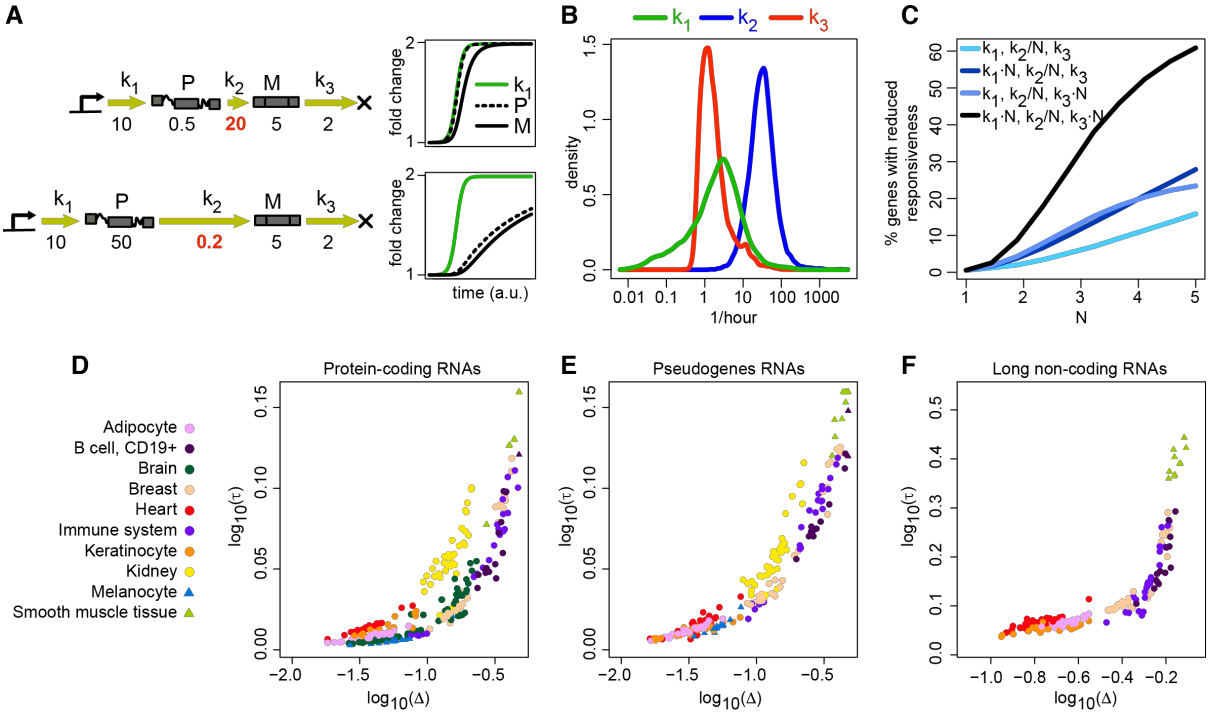

Figure 5 The role of RNA processing dynamics in the responsiveness of mature RNA. (A) Premature and mature RNA abundances and the value of the RNA kinetic rates for two genes with fast (k_2=20) and slow (k_2=0.2) processing dynamics. The corresponding profiles of increased abundance of premature and mature RNA following a doubling in the rate of synthesis (k_1) are indicated on the right. (B) Distributions of RNA kinetic rates in untreated 3T9 fibroblast cells. (C) Percentage of genes with reduced responsiveness following the indicated N-fold modulation of the kinetic rate(s). (D) RNA-seq samples are colour-coded according to the corresponding tissue-type and median values of τ and Δ for protein-coding RNAs are reported. (E-F) as in (D) for pseudogenes and long non-coding RNAs, respectively.

Median τ and Δ were quantified for each sample in the dataset examined in Fig. 4. Tissue types whose samples had similar metrics were highlighted and colour coded for three classes of genes (Fig. 5D-F). In comparison to protein-coding and pseudo genes, long non coding RNAs (lncRNAs) had higher metrics overall. For all three classes of genes, samples from smooth muscle, immune system, CD19+ B-cells, kidney and breast, consistently returned high values of τ and Δ. In other words, these tissue types are expected to be particularly sensitive to alterations of processing dynamics. Notably, it is already known that splicing alterations play a relevant role in most of the tissues with higher metrics[27-31]. Instead, lower metrics invariably were observed for hearth, adipocytes, keratinocytes and melanocytes samples, indicating that these were particularly robust to perturbations in the dynamics

of processing. Hence, specific classes of genes and specific cell types could be particularly either exposed or protected in case the spliceosome's workload becomes larger or its efficiency slightly reduced. While this could be detrimental in physiological conditions, there might be cases when conferring longer life-time to premature RNA species, resulting from reduced processing rates, could be useful. Indeed, premature RNA has been suggested having additional functional roles, potentially different from those of its mature form, and analogous to the ones played by non-coding RNAs[32].

Notably, when the susceptibility to changes in RNA processing dynamics is peculiar to disease conditions, the spliceosome could represent a target for therapeutic strategies. This was recently shown to be the case for lymphomagenesis, where a limiting spliceosome turned out to be the tumors cells Achilles' heel, as it made them more sensitive to drugs affecting RNA processing[33,34]. In agreement with this, an increasing number of genes is exposed to be burdened by RNA processing during B-cell lymphoma development compared to the normal counterpart (Supplementary Fig. 8). In line with this concept, immune system cancer cells and hepatocytes in the context of alcohol dependence resulted to be globally more exposed to changes in RNA processing compared to their normal counterparts (Supplementary Fig. 9). Indeed, therapeutic applications targeting the spliceosome are currently being considered for the treatment of these pathological conditions[30,31,35].

Finally, for each type of tissue, we analysed the functions of the genes with the higher values of τ and Δ. This analysis surprisingly revealed a strong enrichment in genes associated with RNA metabolism and processing (FDR < 0.001, Supplementary Fig. 10). Thus, spliceosome genes seemed particularly sensitive to possibly subtle alterations in their processing kinetics, suggesting a feedback control for this important machinery. On the contrary, terms involving cytoskeleton organization and cell migration were the least affected.

These results revealed how the integrative analysis of premature and mature RNA abundance from total RNA-seq experiments can provide information on the dynamics of transcriptional regulation, and unraveled for the first time the consequences of altered processing dynamics. We developed a Shiny user interface (included within INSPEcT) to explore the impact of processing on RNA responsiveness.

Discussion

The deconvolution of RNA dynamics from RNA-seq experiments is an emerging field of research, which the development of RNA metabolic labelling has fuelled by enabling the analysis of nascent transcription[5,6,11]. We recently developed INSPEcT, a Bioconductor package that, through mathematical modelling of nascent and total RNA-seq datasets, allows the quantification of the kinetic rates governing the RNA life-cycle[12]. We extensively used this tool for the analysis of RNA dynamics controlling several classes of coding and noncoding transcripts[14,36-38]. Aware of the challenges the integrative analysis of nascent and total RNA-seq data poses, we have now expanded INSPEcT to include the possibility of using total RNA-seq datasets only, without requiring any information on nascent transcripts.

To validate the RNA kinetic rates returned by INSPEcT based on total RNA-seq experiments, we analysed simulated datasets and made comparisons with previous studies that included nascent transcripts. By re-analysing various time-course datasets of total RNA-seq, we proved INSPEcT's ability to unravel underlying RNA dynamics and hence provide a deeper understanding of the resulting gene expression programs. Applying INSPEcT to the study of RNA-seq dynamics under steady-state conditions, we also managed to provide the first comprehensive analysis of post-transcriptional regulation from hundreds of publicly available datasets, covering a multitude of tissues and disease conditions. Finally, we uncovered a previously undervalued role of RNA processing dynamics. In fact, while numerous studies focus on the description of alternative splicing patterns and their functional consequences, the role of processing kinetics and their alteration is largely neglected. We and others have recently documented a link between splicing kinetics and the accuracy of splicing[14,39]. Here we revealed how slow kinetics in RNA processing could have a major impact on the responsiveness of mature RNAs.

In order to maximize intronic signal, which we use to derive the abundance of premature RNA, we conservatively decided to take into consideration total RNA-seq experiments where RNA molecules had not been poly-A selected. However, we found that standard coverage (20M aligned reads) RNA-

seq libraries prepared with various protocols including the poly-A selection step, are also suitable for these analyses (Supplementary Fig. 11; [40]), thus broadening the scope of our approaches.

In conclusion, the deconvolution of RNA dynamics can uncover the mechanistic details underlying complex transcriptional responses. INSPEcT is a unifying computational tool able to unfold these layers of regulation in most experimental scenarios, independently from the availability of information on nascent transcription and using both steady-state and time courses profiling of total RNA-seq.

Author Contributions

M.F., and S.d.P. conceived the method and wrote the software. S.d.P. developed the part on the role of RNA processing dynamics. M.F., S.d.P., and M.P. designed the study. E.G. performed the semantic annotation of the metadata of public RNA-seq experiments. M.F., S.d.P., E.G., M.C., and M.P. interpreted the data. M.F., S.d.P., and M.P. wrote the manuscript.

Competing Interests statement

The authors declare no competing interests.

References

1. Mortazavi, A., Williams, B. A., Mccue, K., Schaeffer, L. & Wold, B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Meth* **5,** 621–628 (2008).
2. Orphanides, G. & Reinberg, D. A unified theory of gene expression. *CELL* **108,** 439–451 (2002).
3. Zeisel, A. *et al.* Coupled pre-mRNA and mRNA dynamics unveil operational strategies underlying transcriptional responses to stimuli. *Molecular systems Biology* **7,** 529 (2011).
4. Friedel, C. C., Dölken, L., Ruzsics, Z., Koszinowski, U. H. & Zimmer, R. Conserved principles of mammalian transcriptional regulation revealed by RNA half-life. *Nucleic Acids Res* **37,** e115 (2009).
5. Dolken, L. *et al.* High-resolution gene expression profiling for simultaneous kinetic parameter analysis of RNA synthesis and decay. *RNA* **14,** 1959–1972 (2008).
6. Rabani, M. *et al.* Metabolic labeling of RNA uncovers principles of RNA production and degradation dynamics in mammalian cells. *Nat Biotechnol* **29,** 436–442 (2011).
7. Miller, C. *et al.* Dynamic transcriptome analysis measures rates of mRNA synthesis and decay in yeast. *Molecular systems Biology* **7,** 1–13 (2011).
8. Herzog, V. A. *et al.* Thiol-linked alkylation of RNA to assess expression dynamics. *Nat Meth*

14, 1198–1204 (2017).
9. 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.* **56,** 13479–13483 (2017).
10. Schofield, J. A., Duffy, E. E., Kiefer, L., Sullivan, M. C. & Simon, M. D. TimeLapse-seq: adding a temporal dimension to RNA sequencing through nucleoside recoding. *Nat Meth* **15,** 221–225 (2018).
11. Baptista, M. A. P. & Dölken, L. RNA dynamics revealed by metabolic RNA labeling and biochemical nucleoside conversions. *Nat Meth* **15,** 171–172 (2018).
12. de Pretis, S. *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)* **31,** 2829–2835 (2015).
13. Rabani, M. *et al.* High-Resolution Sequencingand Modeling Identifies Distinct Dynamic RNA Regulatory Strategies. *CELL* **159,** 1698–1710 (2014).
14. de Pretis, S. *et al.* Integrative analysis of RNA polymerase II and transcriptional dynamics upon MYC activation. *Genome Res.* **27,** 1658–1664 (2017).
15. Matsushima, W. *et al.* SLAM-ITseq: sequencing cell type-specific transcriptomes without cell sorting. *Development* **145,** dev164640 (2018).
16. Gray, J. M. *et al.* SnapShot-Seq: a method for extracting genome-wide, in vivo mRNA dynamics from a single total RNA sample. *PLoS ONE* **9,** e89673 (2014).
17. Gaidatzis, D., Burger, L., Florescu, M. & Stadler, M. B. Analysis of intronic and exonic reads in RNA-seq data characterizes transcriptional and post-transcriptional regulation. *Nat Biotechnol* **33,** 722–729 (2015).
18. La Manno, G. *et al.* RNA velocity of single cells. *Nature* **560,** 1–25 (2018).
19. Sabò, A. *et al.* Selective transcriptional regulation by Myc in cellular growth control and lymphomagenesis. *Nature* **511,** 488–492 (2014).
20. Tuomela, S. *et al.* Comparative analysis of human and mouse transcriptomes of Th17 cell priming. *Oncotarget* **7,** 13416–13428 (2016).
21. Eichhorn, S. W. *et al.* mRNA Destabilization Is the Dominant Effect of Mammalian MicroRNAsby the Time Substantial Repression Ensues. *Molecular Cell* **56,** 104–115 (2014).
22. Chang, K. N. *et al.* Temporal transcriptional response to ethylene gas drives growth hormone cross-regulation in Arabidopsis. *eLife* **2,** e00675–e00675 (2013).
23. Galeota, E. & Pelizzola, M. Ontology-based annotations and semantic relations in large-scale (epi)genomics data. *Brief Bioinformatics* **18,** 403–412 (2017).
24. Collado-Torres, L. *et al.* Reproducible RNA-seq analysis using recount2. *Nat Biotechnol* **35,** 319–321 (2017).
25. Zhu, J. W. *et al.* E2F1 and E2F2 determine thresholds for antigen-induced T-cell proliferation and suppress tumorigenesis. *Mol Cell Biol* **21,** 8547–8564 (2001).
26. Leysen, J. E. & PAUWELS, P. J. 5-HT2 Receptors, Roles and Regulation. *Annals of the New York Academy of Sciences* **600,** 183–193 (1990).
27. Lehmann, K.-V. *et al.* Integrative genome-wide analysis of the determinants of RNA splicing in kidney renal clear cell carcinoma. *Pac Symp Biocomput* 44–55 (2015). doi:10.1142/9789814644730_0006
28. Llorian, M. *et al.* The alternative splicing program of differentiated smooth muscle cells involves concerted non-productive splicing of post-transcriptional regulators. *Nucleic Acids Res* **44,** 8933–8950 (2016).
29. Martínez-Montiel, N., Anaya-Ruiz, M., Pérez-Santos, M. & Martínez-Contreras, R. Alternative Splicing in Breast Cancer and the Potential Development of Therapeutic Tools. *Genes* **8,** 217 (2017).
30. Yabas, M., Elliott, H. & Hoyne, G. The Role of Alternative Splicing in the Control of Immune Homeostasis and Cellular Differentiation. *International Journal of Molecular Sciences* **17,** 3 (2016).
31. Schaub, A. & Glasmacher, E. Splicing in immune cells—mechanistic insights and emerging topics. *International Immunology* **29,** 173–181 (2017).
32. Skalska, L., Beltran-Nebot, M., Ule, J. & Jenner, R. G. Regulatory feedback from nascent RNA to chromatin and transcription. *Nat Rev Mol Cell Biol* **18,** 331–337 (2017).

33. Koh, C. M., Sabò, A. & Guccione, E. Targeting MYC in cancer therapy: RNA processing offers new opportunities. *Bioessays* **38,** 266–275 (2016).
34. Koh, C. M. *et al.* MYC regulates the core pre-mRNA splicing machinery as an essential step in lymphomagenesis. *Nature* **523,** 96–100 (2015).
35. Yin, H. *et al.* Deletion of SIRT1 From Hepatocytes in Mice Disrupts Lipin-1 Signaling and Aggravates Alcoholic Fatty Liver. *Gastroenterology* **146,** 801–811 (2014).
36. Mukherjee, N. *et al.* Integrative classification of human coding and noncoding genes through RNA metabolism profiles. *Nat Struct Mol Biol* **24,** 86–96 (2017).
37. Marzi, M. J. *et al.* Degradation dynamics of microRNAs revealed by a novel pulse-chase approach. *Genome Res.* **26,** 554–565 (2016).
38. Austenaa, L. M. I. *et al.* Transcription of Mammalian cis-Regulatory Elements Is Restrained by Actively Enforced Early Termination. *Molecular Cell* **60,** 460–474 (2015).
39. Pai, A. A. *et al.* The kinetics of pre-mRNA splicing in the Drosophila genome and the influence of gene architecture. *eLife* **6,** 1123 (2017).
40. Adiconis, X. *et al.* Comparative analysis of RNA sequencing methods for degraded or low-input samples. *Nat Meth* **10,** 623–629 (2013).