Review

Genome-wide studies of mRNA synthesis and degradation in eukaryotes ☆

José E. Pérez-Ortín [a],*, Lola de Miguel-Jiménez [b], Sebastián Chávez [b]

[a] Departamento de Bioquímica y Biología Molecular, Facultad de Biológicas, Universitat de València, C/ Dr. Moliner 50, E46100 Burjassot, Spain
[b] Departamento de Genética, Facultad de Biología, Universidad de Sevilla, Avda. Reina Mercedes 6, E45012 Seville, Spain

ARTICLE INFO

ABSTRACT

In recent years, the use of genome-wide technologies has revolutionized the study of eukaryotic transcription producing results for thousands of genes at every step of mRNA life. The statistical analyses of the results for a single condition, different conditions, different transcription stages, or even between different techniques, is outlining a totally new landscape of the eukaryotic transcription process. Although most studies have been conducted in the yeast *Saccharomyces cerevisiae* as a model cell, others have also focused on higher eukaryotes, which can also be comparatively analyzed. The picture which emerges is that transcription is a more variable process than initially suspected, with large differences between genes at each stage of the process, from initiation to mRNA degradation, but with striking similarities for functionally related genes, indicating that all steps are coordinately regulated. This article is part of a Special Issue entitled: Nuclear Transport and RNA Processing.

1. Introduction

RNA polymerase (RNA pol) molecules are first recruited to gene promoters to then start transcription along the coding region of the gene. During this process, not all molecules are able to complete their journey, and they also undergo different modifications which mark their particular stage during the process [1]. Eukaryotic RNA pol II molecules are unique among transcriptases in that they have a long C-terminal tail (CTD) with repeated (from 26 in yeast to 52 in humans) heptads, Tyr-Ser-Pro-Thr-Ser-Pro-Ser. CTD is a target of many reversible modifications, including phosphorylation, prolyl-isomerization and glycosylation at various positions (see [1] for a review). These modifications are catalyzed by enzymes, kinases, phosphatases, and prolyl-isomerases, which can thus act as "writers" of an information code [2]. Given the large number of both repeats and individual alternative modification states, the theoretical number of different RNA pol II molecules is much higher than their actual number in a cell. Hence, it is possible that every transcribing RNA pol differs at a given time. Although this is merely a theoretical possibility, it illustrates the potential information stored and used during the transcription process.

As stated above, the particular way a gene is transcribed can vary somewhat from one to another. The actual differences that RNA pol II molecules have along "a model gene" are, in part, already known [2].

Unphosphorylated molecules form the PIC (promoter initiation complex). Then, Ser5 phosphorylation in CTD is required to start early elongation. Ser2 phosphorylation starts later and progressively replaces Ser5 during elongation ([1,2] for a review). The transcriptional differences between genes are, however, just emerging. Phospho-CTD dynamics was recently described to be gene-specific in yeast, and distinct patterns of CTD phosphorylation are present in different classes of RNA pol II genes in yeast [2]. Moreover, it should also be considered that differences in transcription depend not only on RNA pol II molecules, but also on a plethora of ancillary factors that accompany them along the gene. Elongating RNA pol II is accompanied by not only factors that directly affect its processivity and elongation rate, but also by many other proteins involved in co-transcriptional phenomena, which are as diverse as mRNA capping, splicing, editing, cleavage, polyadenylation and export, and also as in chromatin disassembly and reassembly (reviewed by [3]) These factors exchange more or less freely during the transcription process. In many instances, they are attracted by a specific type of CTD state (see [2] for some examples) to become "readers" of the information code. Moreover, genes differ not only in the particular amino acid sequence they encode, but also in other sequence-dependent features. Genes have a different length, G + C content, dinucleotide distribution, exon/intron organization and, especially, chromatin structure. All the cited features can influence transcription. It is likely that they are also interdependent in terms of the particular features of RNA pol II molecules and their partners.

Although much effort has been made to study RNA pol II recruitment and transcription initiation, it now comes over quite clearly that elongation is an important step in gene regulation. Currently, it is tentative to speculate that most (perhaps all) eukaryotic genes

are regulated in some way in the elongation phase and that this regulation has common aspects for the groups of genes that are functionally related. Similarly, recent years have witnessed the discovery that, in many instances, gene expression is post-transcriptionally regulated by means of changes in mRNA stability. Thus the gene regulation topic has broadened from transcription initiation to a wider field comprising all the stages of the process.

Many of the putative distinctive features of gene expression between genes can only be described by comparing a statistically relevant number of genes at the same time. This is currently possible by using the recently developed genome-wide techniques or by implementing new transcriptomics tools. Both strategies are much easier when using model organisms like the yeast *S. cerevisiae* in which a large set of molecular and genetics tools exists. Precisely for this very reason, we focus on this organism; however, the implementation of some genomics techniques in other organisms (e.g., human or *Drosophila* cells in culture) makes it possible to also review the similarities and differences between those organisms.

2. How to measure transcription rates

2.1. mRNA turnover and transcription rate

Defining the transcription rate (TR) is, in principle, quite simple. The TR is the rate at which RNA molecules are produced. It is measured as molecules per minute or per hour. Nevertheless, because there are several steps involved, from early elongation to the appearance of mature RNA in the cytoplasm, it is not so easy to define the TR. From a functional perspective, the only worthy mRNA molecule is the fully processed cytoplasmatic one. This is the molecule that can be used for translation, and may be stored in cytoplasmic bodies or substrates for cytoplasmic exonucleases. Although mRNAs at different maturation stages can be found in the cytoplasm at a given time (e.g., either containing introns or not), simplification can be performed because defective mRNAs are much less stable than correct ones [4]; therefore, they should be a minor component at steady-state levels. Methods to measure mRNA amount/concentration (RA) can discriminate, or not, between intron-containing/spliced and poly(A)/non-poly(A) molecules, which should be taken into account when analyzing results. Nuclear pre-mRNA is also a component of the total RNA extracted from cells, which is usually accepted as a minor component of measurable mRNA given its fast transit to the cytoplasm and its low stability if retained in the nucleus [5]. These considerations can be altered if working with mutants in any of the mRNA processing steps. In any case, we can assume that by using, for instance, microarrays after oligo d(T)-primed cDNA labeling, the measured RA mostly quantifies cytoplasmic functional molecules. The appearance of a processed mature mRNA in the cytoplasm can be considered, in practice, as the actual or effective TR for the cell (the mature TR).

It should also be noted that, in fact, the relevant parameter for mRNA functionality is not its amount, but its concentration. Chemical reactions are governed by the concentrations of reactants. If when comparing mRNA amounts or concentrations cell volume is the same, then it does not matter. However in some instances, two compared samples may have a different average cell volume, which should be taken into account.

Since mRNA is an unstable molecule by nature it has a relatively high turnover, which is much higher in lower eukaryotes that in higher ones [4]. Actual RA depends not only on the mature TR, but also on the disappearance of mRNA molecules. In a nongrowing cell, disappearance of mRNA is only due to the degradation caused by RN-ases. We can define a degradation rate (DR) as the rate at which RNA molecules are degraded. $DR = k_d RA$, where k_d is the degradation constant. In a growing cell, as in most experiments with cultures, and especially with microorganisms (e.g., yeasts), the cells' total volume increases, which provokes a dilution effect on mRNA [6] and adds to the DR. Although the dilution factor is the same for all genes, it is proportionally more important for those genes with a low mRNA turnover [7]. On the other hand, the TR is more affected in genes with a large RA because the TR should compensate dilution proportionally to RA in order to keep the mRNA concentration constant. To date, this last point has not been experimentally demonstrated. Yet it seems logical that actively growing cells with steady-state mRNA conditions (as demonstrated for *S. cerevisiae*, see [8]) would attempt to keep mRNA concentrations constant by devoting part of their TR to compensate dilution (see [7]). It is predicted that the measured TR varies depending on the growth rate. The problem for such an experimental demonstration is that many genes have different RA, and probably different TR, as a function of the growth rate [9,10].

2.2. Methods to measure RNA polymerase density: the nascent TR

In spite of the previous considerations related to the mature TR, most genomic methods used to measure the TR have been developed to measure the nascent TR. It can be defined as the rate at which an RNA molecule is synthesized on its template. The reason for this preference is that the nascent TR is much easier to measure. RNA pol positional information is obtained by means of a nucleic acid sequence by either sequencing or specific probe hybridization. Some methods detect RNA pol by means of the DNA sequences they are sited at, while others use the nascent RNA they are transcribing. The former does not guarantee that the RNA pol is actually transcribing and does not offer information on the DNA strand being transcribed. The latter may, or may not, ensure that they are being polymerized. Some of them derive from previous single-gene methods which have been upgraded to a genomic scale, while others have been directly developed for genomics studies.

Historically speaking, the first method to detect RNA pol on specific sequences was the transcription run-on (TRO) method [11]. Since 2002, several genomic upgrades of TRO have appeared for different organisms. The most comprehensive studies were published by the groups of Myriam Gorospe [12,13] and Jack D. Keene [14] in mammalian cells, and by Legen et al. in plants [15]. In 2004 [16], a genomic upgrade for the yeast *S. cerevisiae* called "genomic run-on" (GRO) was developed. All these protocols use total RNA isolation and its subsequent hybridization onto nylon macroarrays. In 2008, Core et al., [17] developed a GRO-seq protocol for human cells by adapting it for high-throughput sequencing (HTS) technologies. Fig. 1B depicts the outline of the GRO method. The signal corresponding to a given sequence (the probe in the DNA array, the number of hits in HTS) is an indication of the average RNA pol density in it. If we consider a constant speed (elongation rate, see Section 3.1) for the RNA pol, RNA pol density is directly proportional to its transcription rate. Use of PCR-amplified probes provides an average for elongating RNA pol (green ovals in Fig. 1) density in the probe region (usually the whole ORF) in both the sense + antisense orientations. Use of single oligonucleotide probes or HTS provides a local RNA pol density in only one orientation. Therefore, when using tiling arrays or GRO-seq experiments, a precise map of RNA pol distribution throughout the genome is obtained [18; Jordán-Pla et al., in preparation]. Employing GRO in yeast is very easy and can be used in many situations, even during rapid changes because it needs a relatively low number of cells (10^9) and does not need to isolate nuclei. This allows the physiological state of cells to be frozen at the time of sarkosyl addition. Other techniques do not allow this simplicity and quickness.

Another commonly used technique is the immunoprecipitation of RNA pol II cross-linked to its bound DNA by means of specific antibodies (Fig. 1C), which is a mere variation of the general chromatin immunoprecipitation method (ChIP) adapted for RNA pol II. Generation of RNA pol II density precisely maps where RNA pol is seated. The method can be used in combination with all kinds of arrays (ChIP-

chip) [7,18,19,2] or HTS (ChIP-seq) [21]. The results similarly depend on the protocol type in the same way described for GRO. This methodology is also quite simple and fast, so it can be also used in rapid physiological changes for all kinds of cell cultures [22]. A critical feature is that RNA pol II molecules can be distinguished from other RNA polymerases, and even their phosphorylation state can be discriminated using specific antibodies [19,2]. Other genome-wide methods employed to study transcription elongation by means of nascent mRNA isolation (Fig. 1) have not yet been implemented to obtain nascent TR data, although they are susceptible to that.

A common feature for all nascent TR methods is that they should be converted into absolute units by comparing their datasets with the indirect TR ones. Thus, the obtained values correspond to "productive" transcription (see Section 2.5). That is to say, the real nascent TR is not known. Only the transcription events which become a mature mRNA in the cytoplasm are counted. The real nascent TR should be higher because many unproductive events exist. The comparison made between individual genes is still possible because a common factor is used to multiply them all [7].

Therefore, the combination of genome-wide results obtained by different methodologies and with conventional mRNA quantification techniques, either high- or low-resolution, is a potential tool to discriminate the transcriptional features of different genes.

2.3. A snapshot of eukaryotic transcription

The GRO protocol enables the first whole TR dataset to be obtained for an organism [7]. Thus, 90% of yeast genes show TRs per cell of between 2 and 30 mRNAs/h, with a median per cell of 7 mRNAs/h. This corresponds to 0.078 RNA pol II molecules/kb and 0.1 molecules/gene. Given the much skewed distribution toward highly transcribed genes, RNA pol II molecules are scarce in yeast genes: only 14% of genes have any actively transcribing RNA pol II molecule. Hence, transcription onto canonical genes does not seem to be a common feature of the yeast genome. In spite of the 20,000 RNA pol II molecules/cell [23] only around 700–1400, according to our calculations, are actually transcribing those genes encoding proteins at a given time. One possibility is that part of these molecules is unable to productively transcribe. However, the amount of CTD-phosphorylated molecules (12,000, according to ref. [23]) suggests that they can transcribe in other regions outside canonical genes (see the next section), or perhaps that the mRNA molecules which reach the cytoplasm are merely a fraction of the nascent ones, as previously explained. The most transcribed genes are those that encode histones, with an average TR of ~200 mRNAs/h in the S phase, whereas ribosomal protein (RP) genes are less transcribed than previously thought (see refs. [7,24] for a discussion). The analysis done by P. Cramer's group using a different method (see Section 2.5) obtained similar results: a skewed distribution of transcripts, from 0.7 to 400 mRNAs/h, and a low global transcription rate with a median mRNA synthesis of 12 mRNAs/h [20].

A similar study was recently performed in mouse fibroblasts by [25]. It shows that the median TR in those cells is 2 mRNAs/h, with a distribution of between 0.2 and 20 mRNAs/h for most genes. Given that fibroblasts are 20–50 times larger than a yeast cell, TRs are relatively much lower than in yeast if we expect them to support the higher synthesis required to achieve similar mRNA concentrations. However, the mRNA steady-state concentrations seem to be lower because the actual mRNA amounts in fibroblasts (average ~30 mRNAs/cell) are only 10 times higher than in yeast (average ~3 mRNAs/cell, in [7]). This partially accounts for their lower TR. On the other hand, the mRNAs of higher eukaryotes are more stable (median half-life ~40 min in [26], but higher values in other references [25]) than in yeast (median half-life of 23 min [27]), which explains the remaining difference in the TR. Thus, it seems that the TRs of higher eukaryotes are even lower than in yeast because they have

more stable mRNAs and lower mRNA concentrations. In this case, the snapshot reveals an even scarcer RNA pol II outlook.

The conclusion drawn from these studies is that the productive transcription of canonical mRNAs in a eukaryotic cell is quite low, although real total transcription may be higher (see the next section). Given the translation rate for both yeasts (~6 proteins per mRNA per hour, [28]) and mouse fibroblasts (40 proteins per mRNA per hour, [25]), and the much higher number of mRNAs/cell (in the range of 10^4 in yeast and of 10^5 in fibroblasts) if compared with 1–2 gene copies/cell, the cost in translation is much higher than in transcription for a eukaryotic cell, as previously anticipated using nongenomic data [29].

2.4. Influence of cryptic transcription in nascent TR determinations and gene regulation

A potential artifact of nascent TR methods is interference of cryptic transcription. Cryptic transcripts are defined as those which are difficult to detect given their small amount and because they lack correspondence with predicted canonical genes. Since functional genomics techniques came about, a plethora of different kinds of cryptic transcripts was detected in all the eukaryotes studied. There is much controversy as to the magnitude and significance of cryptic transcription [30,32]. In any case, there is no doubt that RNA pol II transcribes more RNAs than previously assumed.

The RNA pol II molecules placed anywhere along the genome regions contained in microarray probes can be detected irrespectively of them transcribing "canonical" transcripts or cryptic transcripts. Cryptic transcripts account for both the sense and antisense transcripts overlapping canonical transcripts. Some nascent TR methods are intrinsically nondiscriminating for sense or antisense transcripts: ChIP, because of the double-strand nature of the immunoprecipitated DNA; the original GRO method, given the dsDNA nature of the macrochip probes used [33]. For these reasons, both sense (canonical and cryptic) and antisense transcriptions are summed in the signal. A huge number of cryptic transcripts was demonstrated in many organisms, including yeast (see [31]). In yeast, two main types were defined: cryptic unstable transcripts (CUT), which are only detectable in the absence of nuclear exosome activity and stable uncharacterized transcripts (SUT) [34]. Another class of cryptic unstable transcripts was recently described (XUT, Xrn1-sensitive unstable transcripts) and it was proposed to be degraded by the Xrn1 5′-exonuclease in the cytoplasm. Therefore, they increase in an *xrn1* mutant [35]. Some authors argue that cryptic transcription can be responsible for the differences observed in some genes' response to stress situations when comparing mRNA data and RNA pol II ChIP data [18,34,35]. This contrasts with recent evidence for a small amount of antisense transcripts compared to the canonic mRNA pool in the bidirectional promoters context [31,37].

We analyzed several of the different cryptic transcription datasets published [34–36,38] in relation to the uncoupling between the TR and RA profiles. We found that there is no statistical correlation between the presence of a cryptic transcript in a gene and the lack of the correspondence of the RA profiles and TR profiles during a stress response (García-Martínez et al., submitted). Therefore, we conclude that the main reason why RA profiles differ from the TR ones lies in the changes occurring in the DR (mRNA stability), as explained below. We believe that the nascent TR reflects mainly the "canonical transcription" for most genes. In fact, as there is very little overlapping between the different datasets, it seems that most yeast genes have cryptic transcripts. However, the majority of them are in a very low proportion to their canonical counterpart (discussed in [31]); thus, the mere presence of an overlapping cryptic transcript is not a way to classify genes. Nonetheless, we observed some statistically significant biases in those genes whose expression increases during the stress response, as well as some kinds of cryptic transcripts
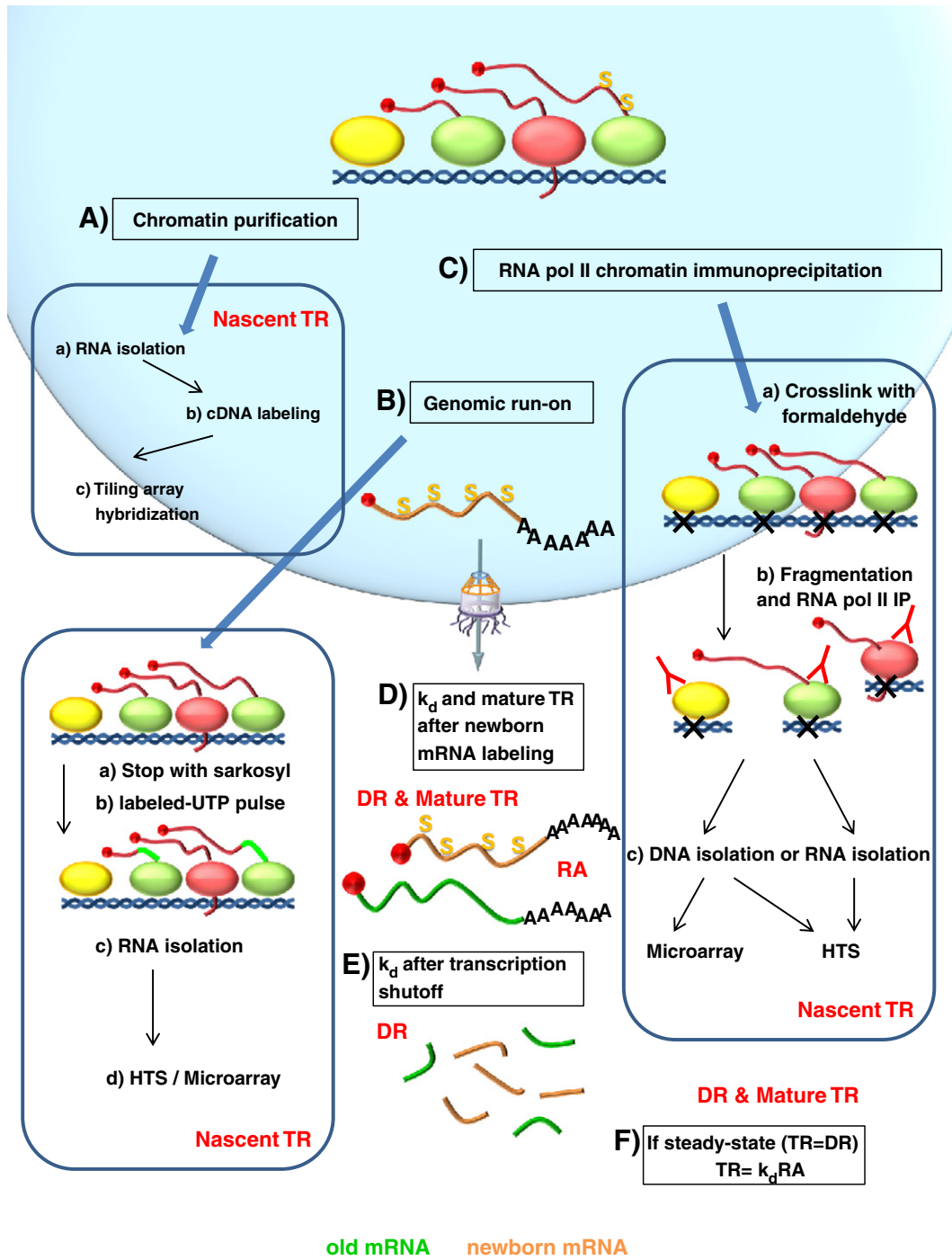
old mRNA   newborn mRNA

Fig. 1. Scheme of the mRNA turnover in a eukaryotic cell indicating the different genomic techniques used for TR, elongation rates and mRNA stability determinations. Determination of the nascent TR in the nucleus is based on the detection of either RNA polymerases or nascent RNAs isolation. A) Native chromatin can be purified and nascent RNA can be isolated from it. Then it is converted into cDNA and used for tiling array hybridization [55]. B) Alternatively, cells can be used for the GRO analysis. During run-on, cells are permeabilized with sarkosyl, which abruptly decreases NTP pools and stops all elongating RNA pol. The detergent also disrupts the chromatin structure, thus avoiding a further initiation event. A subsequent pulse of externally added ribonucleotides, including labeled UTP (^{33}P-UTP, thio-UTP or biotin-UTP), induces a short elongation (about 200–300 nucleotides for ^{33}P-UTP) of these RNA pol (green ovals). Backtracked RNA pol II (red ovals) does not incorporate nucleotides. Nascent RNA becomes labeled in this way to be then used directly as a hybridizable molecule in a DNA array or, alternatively, it can be purified based on the features conferred by the UTP analog used. The latter could be used for high-throughput sequencing (HTS) or for DNA microarray hybridization. C) Alternatively, cells can be fixed with formaldehyde, and subjected to chromatin fragmentation and immunoprecipitation with different RNA pol II antibodies (ChIP). In this ternary complex, either DNA or RNA can be isolated from polymerases and analyzed. Recovered DNA is suitable for array hybridization (ChIP-chip) [8,19,69] or HTS (ChIP-Seq), whereas RNA can be converted into cDNA and subjected to HTS [37]. D) Appearance of recently synthesized mRNA (newborn, orange) in the cytoplasm can be followed by thio-uridine *in vivo* labeling (yellow "S" onto nascent RNA), which is purified, quantified and compared with the nonlabeled old mRNA (green). With this technique, it is possible to calculate the mature TR and mRNA half-lives at the same time (DTA; [20]). E) mRNA half-lives can also be calculated from a transcription shutoff experiment [27], or F) from TR and RA data by assuming the steady state ($k_d = TR/RA$). In this last case, the mature TR can be alternatively calculated from RA and k_d data.

and genes, which lower RA during the stress response (mainly translation-related genes), which are enriched in other kinds of cryptic transcripts [García-Martínez et al., submitted]. These results argue that although cryptic transcription is not a main factor to alter the TR evaluation during stress responses, it proves to be a contributor to the regulation of some specific gene groups [35,36]. Along these lines,

there is recent evidence that meiosis-specific genes are regulated by antisense transcription in both *S. cerevisiae* [39] and *S. pombe* [38].

2.5. Methods for evaluating the mature transcription rate

Cell control on mRNA concentration should be implemented in relation to the valuable molecules that are useful for translation; that is, the cytoplasmatic ones. The rate at which newborn mRNA appear in the cytoplasm can be termed the mature TR (Fig. 1D). Current methods for determining the mature TR are based on the *in vivo* metabolic labeling of transcripts with uracil or uridine analog pulses. If the pulse is short enough and the mature TR is considered constant during this pulse, the subsequent isolation of labeled (newborn) mRNA is a method to evaluate the mature TR [40]. Quantification of non-labeled (old) or total mRNA allows measuring RA in the same experiment. The recently developed dynamic transcriptome analysis (DTA) methodology in yeast [20] and mammalian cells [25,26] is based on the purification of newborn mature transcripts after a metabolic pulse of 4-thiouridine with varying times (6 min in yeast, and from 10 min to 2 h in mammals). Given the simultaneous calculation of RA, it can also determine mRNA half-lives at the same time, thus assuming a steady-state situation (see Section 4.1).

On the other hand, the mature TR can also be calculated from RA and mRNA stability datasets indirectly (indirect TR, $TRi = k_d \, RA$) by assuming a steady state [41]. Alternatively, mRNA half-lives (or k_d) can be calculated indirectly from experimental mature TR and RA datasets using the same equation (see Section 4.1). Both indirect calculations involve the problem of a higher error being caused by the mathematical process (see [7,42] for a discussion).

2.6. Differences between the nascent and mature transcription rates

Genes may be clustered according to their profiles when studying the expression of the whole gene set of a eukaryote using genomic techniques. Clustering may reflect the existence of a common regulation of these genes. This is usually assumed to be controlled by a transcription factor (TF), and it is named a regulon [6]. Most studies use RA determinations to compare the expression of genes. However, one change in RA can be the outcome of not only a change in the TR, but also in mRNA stability, or even in both. The most common direct effect of a TF is noted on the nascent TR. Therefore, according to the nascent TR, clustering should be the best method to detect regulons. Even the rate at which mRNA appears in the cytoplasm (the mature TR) may be less suitable for regulon detection because some post-transcriptional events, like mRNA export, can affect the mature TR profile. In fact, an independent meta-analysis of previously published data [16,43,44] reveals that TR profiles are more suitable to predict functional relatedness than RA profiles [45]. Although only formally demonstrated in *S. cerevisiae*, it seems reasonable to extend this conclusion to other eukaryotes.

The nascent TR also seems the best way to classify genes according to active chromatin marks. The passage of RNA pol II molecules along the chromatin template is affected by a particular nucleosomal organization (see Section 3.4). Since the nascent TR measures the actual elongation rate, which is affected by nucleosome structure and positioning, it should better correlate with the characteristic chromatin marks of active genes than the mature TR. Comparing the level of the different active chromatin features, such as H3-K_{36} trimethylation or enrichment in Esa1p or Gcn5p histone acetyltransferases, revealed that the correlation with the nascent TR (calculated by GRO) is greater than when compared with the RA or with the indirect mature TR [7].

The mature TR, however, is best suited for kinetic studies because it is directly related with the cytoplasmic mRNA steady-state and is not affected by cryptic transcription, as previously mentioned. So it is that each kind of TR dataset is appropriate for a given goal. A recent review [46] provided a compendium of features and advantages of each TR determination method. Comparing all existing yeast TR datasets with each other and with a standardized RA dataset revealed that they all correlate quite well [24]. Those corresponding to the nascent TR correlate better with each other, and the same may be said for those corresponding to the mature TR, which also has better correlates with the RA dataset. These results are logical because nascent mRNAs should be processed and exported to the cytoplasm where the mature TR is measured. Thanks to the comparative analyses of both nascent and mature TRs, datasets should provide further insights into differences in the post-transcriptional steps between the different groups of genes.

3. Functional insights into transcription elongation obtained from genome-wide studies

3.1. How to study elongation

How to approach transcription elongation quantitatively is an open problem with different partial solutions. Studying transcription elongation *in vivo* has involved the use of very diverse techniques, including depletion of intracellular pools of ribonucleotide triphosphates by drugs like 6-azauracile [47] and mycophenolic acid [48], or the comparison of reporter genes of varying lengths [49,50]. By using antibodies against RNA pol II (Fig. 1C), chromatin immunoprecipitation enables the measurement of elongation rates and processivity [51]. The elongation rate is the speed at which RNA pol runs along its template. It is measured in kb per minute or per second. The same genomic techniques followed to study transcription elongation can be used to measure the nascent TR (Section 2.2 and Fig. 1), although the output data differ. Some of their drawbacks are also similar: the ChIP of RNA pol II lacks specificity against the active, elongation-competent form of RNA polymerases as they are detected regardless of the transcriptional states (active, paused, backtracked; see Fig. 1). *In vitro* studies showed that RNA pol II is often arrested (red ovals in Fig. 1) during elongation in the chromatin context [52], while molecular modeling indicated that backtracking during elongation is indeed a frequent phenomenon *in vivo* [53]. Therefore, the presence of polymerase in a region should not be directly assumed to be an actual transcription. Moreover, ChIP also detects the RNA pol II molecules sited at the promoter before elongation (yellow ovals in Fig. 1). On top of this, ChIP cannot discriminate the sense/antisense transcripts, even when using strand-specific probes or HTS.

In addition to run-on (see Section 2.2), other techniques can partially skirt this drawback (Fig. 1). Native elongating transcript sequencing (NET-seq) is a HTS-based technique which provides information on the position of elongating RNA pol II within highly expressed genes on a genome-wide scale, thus avoiding the problems mentioned. It involves the isolation and HTS of the nascent transcript associated with immunoprecipitated RNA pol II [37,54]. Nascent mRNAs were also directly purified from isolated chromatin converted into cDNA and hybridized to tiling arrays (Fig. 1A) [55]. In both these cases, only the RNA pol with a bound transcript is located. This result comes closer to that of GRO, but is not identical because non-elongating or backtracked RNA pol II molecules are detected. Thus, the existence of several alternative methods to map RNA pol allows research into transcription elongation by distinguishing all the RNA pol II states in its template (see Fig. 1).

3.2. Genomic ChIP approaches to elongation

By utilizing antibodies that specifically recognize RNA pol II, ChIP-chip and ChIP-seq analyses demonstrated that transcriptional regulation at the post-initiation stages is a common phenomenon across metazoan genomes. At least 40% of the genes present a promoter-proximal accumulation of RNA polymerase II (Fig. 2A) in mammalian,
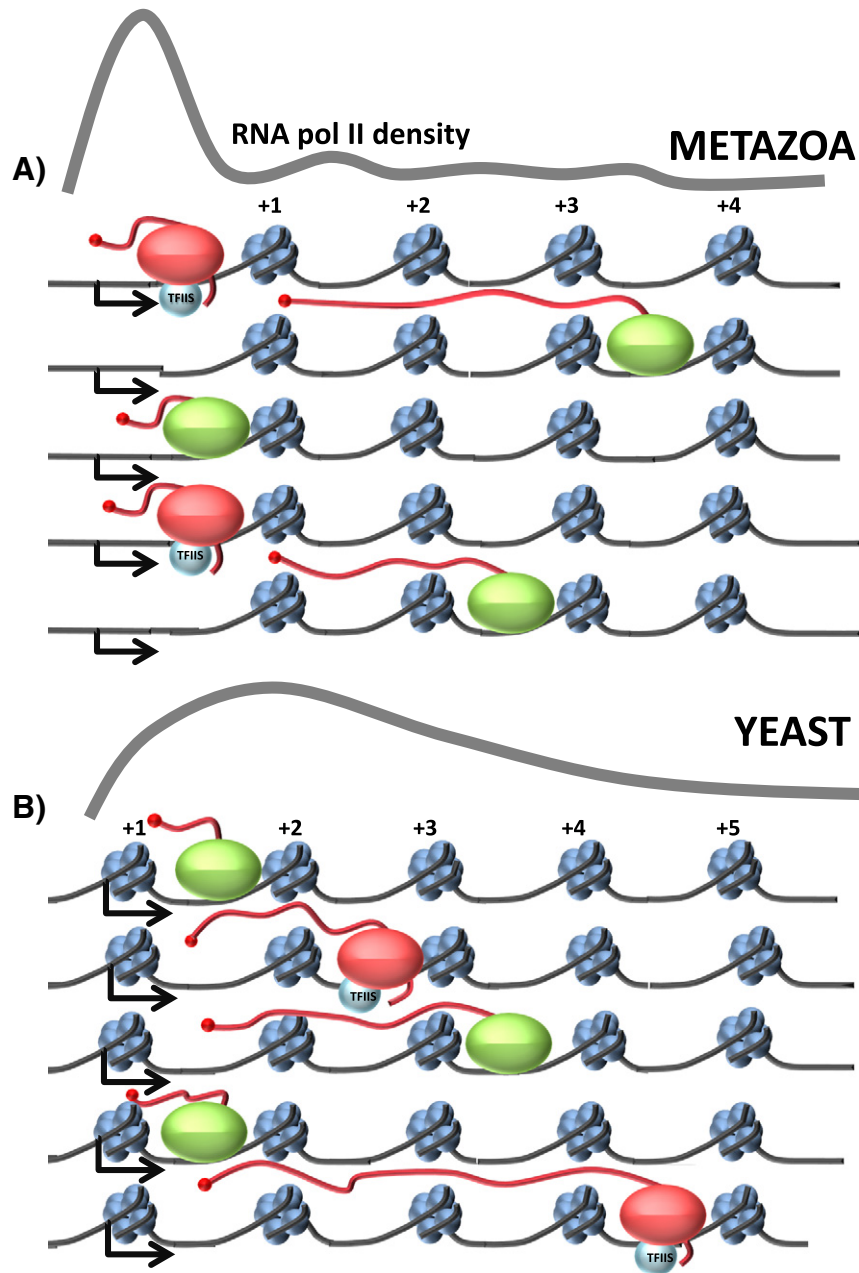
Fig. 2. Transcription elongation features revealed by genomic studies. Genomic high resolution mapping of RNA pol II across the genome, either by ChIP or run-on, has revealed that promoter proximal pausing is a common phenomenon in metazoa (A), but not in the yeast *S. cerevisiae* (B). Statistical accumulation of RNA polymerase II shortly downstream of the transcription start site may result from either a longer resident time or premature termination. The comparison made between RNA pol II distribution and nucleosome positioning maps indicates that promoter-proximal pausing takes place at nucleosome +1 (A). In yeast, nucleosome +1 frequently covers the transcription start site, while the NET-seq data suggest that RNA pol II pauses at any positioned nucleosome throughout the transcribe region (B). Evidence from certain metazoan individual genes and from yeast NET-seq indicates that paused RNA pol II becomes easily backtracked (red ovals) and that TFIIS acts during backtracked RNA pol II reactivation.

Drosophila and *Caenorhabditis* cells [56–59]. The combination of RNA pol II chip and histones ChIP approaches revealed that a subgroup of the so-called bivalent genes, bound by *Polycomb* repressive complex 2, displays stalled RNA pol II. The bivalent genes in embryonic stem cells display characteristic chromatin markers of silenced (histone H3K27 tri-methylation) and active transcription (histone H3K4 tri-methylation) [60]. Other mammalian examples of promoter-proximal accumulation of RNA pol II are growth-promoting, pro-oncogenic genes like *FOS* and *MYC*, which preferentially reside in compact chromatin, and are inefficiently transcribed under basal conditions [61].

In yeast (Fig. 2B), ChIP-chip analyses revealed that almost 2500 repressed genes show poised RNA pol II in the stationary phase [62], but

that only a few, like *CYC1* and those encoding NTP-biosynthetic enzymes, display an accumulation of RNA pol II in their 5′ region under repressive conditions in exponentially growing cells [63,64]. For NTP genes, transcription regulation works at the initiation level through an attenuation mechanism [64,65]. It is not clear whether the accumulation of RNA pol II at the 5′ end in the other cases responds to a pausing phenomenon. In any case, RNA pol II pausing at the promoter-proximal sites is not a frequent phenomenon in exponentially growing yeast [66]. ChIP has been also used at the genomic level to distinguish the different forms of CTD phosphorylation [19,2]. Ser5 phosphorylation shows a characteristic peak at the beginning of the transcribed region [19,2] whereas Ser2 and Ser7 present more uniform patterns [19,2]. Moreover, a constant delay of Ser2

phosphorylation onset in relation to that of Ser5, regardless of gene length, was found [2].

3.3. GRO and nascent RNA isolation as approaches to elongation

GRO is also a useful tool to study elongation. The relative distribution of the run-on signals across the genome also proved most useful to detect promoter proximal pausing (Fig. 2A), as well as other important features of eukaryotic gene expression, such as divergent transcription [17,67]. Our group used GRO profiles, quantified as 3′/5′ ratios, to detect the gene-specific transcription elongation patterns and the functional roles of general transcription factors in elongation [68].

A comparison of the genomic patterns of run-on signals, reflecting elongation-proficient RNA polymerases, with the genomic distribution of the total RNA pol II (Fig. 2B) measured by ChIP is an efficient way of detecting RNA pol II arrest during elongation [69]. The correlation between run-on and ChIP signals is a generally good, but some gene ontology categories exhibit lower run-on values than expected depending on their ChIP signals [69]. This difference is particularly striking for those genes encoding ribosomal proteins (RP), indicating the accumulation of arrested RNA pol II in these genes. One particular behavior of RNA pol II during transcription elongation of RP genes is also reflected by the 5′-biased distribution of the run-on signals in this specific group of genes [68].

The run-on signals of RP genes lowered when comparing the yeast cells exponentially growing in a glucose-containing medium to those cells exponentially growing in galactose-containing medium. However, this drop was significantly more marked in the total RNA pol II levels detected by ChIP than in the run-on signals; accordingly, the run-on/ChIP ratio for the RP genes diminished from glucose to galactose. The simplest interpretation of these results is a specific reduction in the number of backtracked RNA pol II complexes in response to the carbon source. In fact, the run-on/ChIP ratio depends on the integrity of the Ras–PKA pathway and is mediated by the silencing domain of Rap1, an essential transcription factor for RP genes transcription [69]. In short, comparison of genomic RNA polymerase II ChIP and GRO indicates that RNA pol II arrests are an important ingredient of transcriptional regulation.

Native elongating transcript sequencing (NET-seq) is a complementary genomic tool to study elongation [37]. It provides transcriptional profiles at a nucleotide resolution and allows the detection of pervasive RNA pol II pausing through the body of highly transcribed genes, with prominent peaks at all the first four positioned nucleosomes (Fig. 2B) [37]. The NET-seq patterns in a mutant lacking the RNA-cleavage factor TFIIS display a large-scale downstream shift of 5–18 bp at the position of the paused RNA polymerases, suggesting that most of the pausing in the wild type results in arrest and backtracking, and that a slow step follows cleavage before transcription resumes [37]. TFIIS also plays a key role in the regulation of transcription elongation. Involvement of TFIIS in the activation of stalled RNA pol II was first demonstrated for those occupying the promoter proximal region of *Drosophila* Hsp70 [70]. *In vivo* TFIIS depletion using RNA interference provoked a delay in Hsp70 induction in response to heat shock which, in this case, was only possible after a new round of transcription initiation [70]. A combination of RNA polymerase and TFIIS ChIPs evidenced a transcription repression mechanism based on the inhibition of TFIIS recruitment during early elongation in pro-oncogenic genes like *FOS* and *MYC* [61]. A series of experiments carried out in our lab suggested that TFIIS is also specifically involved in the regulation of RNA pol II elongation on RP genes in yeast, at least under the strong transcriptional stress caused by 6 AU [Gómez-Herreros et al., submitted].

The importance of RNA pol II arrest, TFIIS and other arrest-counteracting factors, like Ccr4-Not [71], in regulating transcription elongation is likely related to the interplay between elongating RNA pol II and positioned nucleosomes. *In vitro* experiments demonstrated that TFIIS promotes pol II transcription through the nucleosome [52], and that TFIIS modifies the mechanical performance of RNA pol II, allowing it to operate against higher loads, like those expected in the positioned nucleosomes context [72]. Studies of this interplay between RNA pol II elongation and chromatin structures are still scarce. A clear correlation was found in higher eukaryotes between promoter-proximal pausing and the chromatin configuration of the 5′ regions [72]. Unlike *S. cerevisiae*, *Drosophila* does not bury its transcriptional start site in the +1 nucleosome (Fig. 2A). At thousands of *Drosophila* genes, RNA pol II initiates in a nucleosome-free region, engages +1 nucleosome, and pauses [73].

Additionally to the previous results, a NET-seq study in yeast [37] restricted the importance of cryptic transcription (see Section 2.4) and determined the role of the Rpd3S histone deacetylase complex in it. Moreover, it revealed that the elongation process is far from being uniform: it is punctuated by pauses related with nucleosomes (see above), which can spend times that are comparable to those that the RNA pol II employs to move forward. Another reason for non-uniform elongation stems from the work of Carrillo-Osterreich et al. [55] by means of nascent RNA isolation after chromatin purification in yeast (Fig. 1A). These authors found that RNA pol II molecules pause within the terminal exons in those genes containing introns to allow co-transcriptional splicing.

These results are a good example of how transcription elongation dynamics is more complicated than initially suspected and of how it results from the interplay among the elongating RNA polymerases, chromatin and factors associated with the elongation complex. Interpreting the biological meaning of the genome-wide comparison between elongating RNA profiles with chromatin, elongation factors and other associated factors binding maps [74] is an exciting challenge for the near future.

4. The importance of mRNA stability

4.1. Genomic-wide methods to evaluate mRNA stabilities.

At any time, RA is the result of the balance between the TR and the DR (Fig. 1F). When RA needs to be changed, it is possible to vary the TR, the DR, or both. Traditionally, most studies into gene expression at either the single gene level or the functional genomics level have unconditionally assumed that changes in RA are due to changes in the TR. In other words, a given gene is expressed because a TF binds its promoter and attracts RNA pol II to transcribe it (increasing the TR) which, in turn, increases RA. Thus, RA profiles were implicitly considered the direct result of TR profiles. However, the same effect on RA can be obtained by an inverse change in the DR rather than in the TR. mRNA half-lives were found to vastly differ among mRNAs, physiological conditions and organisms [4]. Thanks to the development of genomics methods [27], it is now clear that the mRNAs encoding those proteins belonging to the same pathways or functions tend to have similar half-lives. This suggests that regulons also exist for the control of mRNA DRs [75]. Two types of *trans* factors which control these post-transcriptional regulons have been described to date: RNA-binding proteins (RBP) [83] and micro RNAs (miRNA) [76].

Genomic-wide mRNA stabilities can be determined directly by using transcription shut-off methods (Fig. 1E) [27,77]. These methods pose some problems as to the potential artifacts they cause in cells (discussed in [42]). In some cases, the DR is not directly measured, but RA and the TR are. Moreover, mRNA stability can be estimated from them (Fig. 1F). We used the GRO method to determine mRNA stabilities from TR and RA data by assuming steady-state conditions [16] or by using a mathematical approximation when assuming that they are not [42–44,78]. Other authors used mature TR evaluation methods (see above) in both yeast [20] and higher eukaryotic cells [26], which are able to determine the DR at the same time.

4.2. Changes in mRNA stability during transcriptional responses

Although post-transcriptional regulons were proposed on the basis of the differences in the mRNA stability found when comparing different genes situations, and also based on the existence of RBPs [27,75], the fact that functionally-related groups of genes follow common DR profiles during transcriptional responses was first demonstrated by a GRO analysis of the yeast response to a carbon source change [16]. As the GRO technique is able to indirectly determine mRNA stabilities in steady-state situations, it revealed that the genes belonging to functionally-related groups behave coordinately with the DR. Similar studies in other organisms obtained comparable conclusions, although not at the whole genome-scale level (e.g., [12,14]). In the carbon source change experiment, the times selected were separated by hours, and the steady-state conditions can apply to each one [16]. However, fast responses, typical of stress situations, do not meet the steady-state conditions. For these cases, a mathematical algorithm based on the use of chemical kinetic laws was developed [42], which made it possible to determine approximate k_d profiles in response to different stresses for most genes and to quantify the influence of DR changes on RA profiles. Our studies [43,44,78], and those of others [79,80], revealed that changes in the DR take place during stress responses for many genes which respond positively or negatively to stress. The mRNA stability of many other genes, however, does not change substantially. In line with this, interesting differences were noted between various stresses [79], which probably depend on stress intensity. It is interesting to note that some genes, which do not exhibit a coordinated behavior at the TR level, and are, therefore, not part of a regulon, actually display coordinated behavior at the mRNA stability level. This is the case of mitochondrial RP genes, which cluster in the mRNA stability analysis, but not in the TR analysis, during the shift from a glucose to a galactose medium [16]. These genes do not show obvious common regulatory elements in their promoters, but a Puf3-binding sequence in their 3′-UTR, which was verified to coordinate their stability during changes in respiratory behavior [81].

Both the TR and DR can cooperate in the transcriptional response. There are many cases in which genes increase RA by rising their TR and lowering their DR (increase in mRNA stability) for several minutes; for instance, after osmotic stress in *S. cerevisiae* [44]. A similar result in the same stress response was reported by other authors using a different method to measure mRNA stabilities [80], and by Shalem et al. [79] during the yeast oxidative stress response. After some minutes of stress, these genes reverse the change in the DR in parallel to the transcription shut-off. This effect was interpreted as a DR change which precedes changes in RA [80]. In other instances, the DR acts as an intuitively opposite counter to the TR. mRNA stability decreases while the TR increases. This strategy is energy-consuming because of the futile cycle involved, but it speeds up the response because the kinetics of the transcriptional change is inversely proportional to mRNA stability [42]. Several authors working in yeast [79] or mammalian cells [26] postulated that these kinds of changes in the DR thus contribute to sharpen response peaks. It is interesting to note that many of the short-lived mRNAs in mammalian cells are targets of miRNAs [26], which is indicative of the important role of those regulatory *trans* factors in mRNA stability and, therefore, in gene expression responses in higher eukaryotes.

4.3. The respective roles of both the TR and DR in gene expression

Altogether this published evidence indicates the importance of the DR in controlling mRNA levels during gene regulation. However, despite both the TR and DR being theoretically equivalents, they do not play the same role in determining the amount of mRNA in stress responses or under steady-state conditions. When comparing the TR, the RA and the mRNA half-life datasets, it appears that the TR and RA always present a positive significant correlation, whereas the mRNA half-lives do not correlate with RA and/or the TR. Moreover, mRNA half-lives reveal slightly negative correlations with RA and the TR in yeast [16,20] or mammalian [25] cells. On the other hand, when measuring the number of genes with response profiles that are significantly affected by DR changes, the vast majority of genes show that TR changes are the main determinant of RA profiles [16,25,26,43,77]. Thus, it seems that the DR is not utilized by the cell to control most of the mRNA levels in the majority of situations, but to distinguish between rapid or less-rapid responsive genes according to the half-lives of their mRNA [26,42]. Many mRNAs have a relatively constant DR and are, perhaps, not part of a post-transcriptional regulon. Other gene categories (stress-induced and RP yeast genes in responses to stress [43,44], yeast mitochondria-related genes when carbon source shifts from glucose to galactose [16] or inflammatory and immune signaling genes in dendritic cells [26]) show a highly dynamic DR. This is probably due to the *cis* elements present in their 3′-UTRs, which make them targets of specialized RBPs [78,82,83]. A general change in DR is also observed under nongrowth conditions, like the stationary phase [84] or after a change from glucose to galactose medium in yeast [16,85]. This phenomenon probably results from a general change in the DR machinery and/or p-body organization.

The corollary of these results is that each particular group of genes has a coordinated TR through their common promoter and chromatin organization, and a coordinated DR through their common 3′-UTR sequences, which are subjected to (transcriptional) regulons and post-transcriptional regulons [82]. It would be interesting to check if the same genes belong to the regulons and post-transcriptional regulons, or whether some combination exists.

5. How the transcription rate and mRNA stability can be coordinated

5.1. Crosstalk between the transcription and degradation rates

Maintaining proper mRNA levels is critical for the proper functionality of any living cell. These levels are controlled by the balance between the synthesis and degradation rates. It is quite likely that all the processes with more than one controlling point must have mechanisms to coordinate them. Failure to coordinate these processes might result in an improper level. Nevertheless, it is known that in different circumstances, the total mRNA level varies; however, this variation is never excessive. For instance, in yeast, we found that the total poly(A) level varies between strains or physiological situations within a less than three-fold range (Pérez-Ortín et al., unpublished observations). There seems to be a homeostatic control of the total mRNA level in the cell, which is, obviously, independent of the known wide variations in many individual mRNAs levels that respond to environmental circumstances. In this sense, and as stated above, the particular level of each mRNA is controlled by acting at both the transcription and degradation rates. Moreover, functionally related groups of genes are coordinated by using both regulons and post-transcriptional regulons. Following the previous reasoning, coordination between both kinds of regulons should exist.

How does the cell sense the mRNA level and how does it coordinate general and particular TRs and DRs? M. Choder's group demonstrated that a class of proteins, which they call "mRNA coordinators", is loaded onto mRNA during its transcription, remains bound to mRNA and accompanies it through the different stages of its life. An mRNA coordinator can, thus, regulate mRNA export, translation, its assembly with a p-body and its stability [86–88]. These authors proposed that the Rpb4/7 heterodimer, which was originally viewed merely as subunits of RNA pol II to function as a subunit of the translation and decay apparatus, acts as a prototype of mRNA coordinators. In yeast, deletion of *rpb4* alters not only

the transcription of many genes, as expected, but also their export [89], mRNA half-lives [86,90], p-body association, and even translation [88]. Thus, this mRNA coordinator acts as a general guard of many mRNAs and can, therefore, transmit information from the transcription process to later stages of mRNA life. Conversely, some components of cytoplasmic mRNA degradation machines, such as Xrn1, the Dcp1/2 Pat1/Lsm1-7 complex, among others, can travel to the nucleus and act as transcription factors. In *xrn1* mutants, the TR of every gene lowers in such a way that it seems to correlate with the effect of the mutations on mRNA stability. In this mutant, most mRNAs' levels do not change much despite their abnormally high stability [Haimovich et al. submitted]. So, it seems that Xrn1 and other decay factors also deliver information, in this case, from the cytoplasm (DR) to the nucleus (TR). Other recent examples of shuttling mRNA binding factors include the Ssd1 protein [91] and the Ccr4–Not complex. Ccr4/Pop2 is the major cytoplasmic mRNA deadenylase, whereas Not proteins are components of the mRNA processing bodies. Ccr4–Not was established to act as a positive elongation factor for RNA pol II, probably by promoting the resumption of the elongation of arrested RNA pol II [71]. Consistently with their dual role in two separate compartments, these proteins shuttle back and forth between the nucleus and the cytoplasm. Rpb4/7 subunits also travel back to the nucleus after mRNA decay [92], whereas decay factors travel back to the cytoplasm after activating the TR. Therefore, the gene expression process seems to be a circular system [Haimovich et al. submitted]. Although not formally demonstrated, the functional similarities between Rpb4/7 and some other RNA-binding proteins in other eukaryotes mean that it is quite likely that similar mechanisms exist in organisms other than yeast [93].

In light of this, recent studies compared transcription and mRNA decay in two related yeasts, *S. cerevisiae* and *S. paradoxus*, to find that the evolution of transcription and mRNA decay processes is coordinated because the genes that diverge in mRNA degradation tend to also diverge in transcription [94]. An analysis of a hybrid of these two species showed that changes in transcription and mRNA degradation are likely driven by the same individual mutations. The factors involved in both processes (e.g., Rpb4/7 and Ccr4–Not) seem to be associated with the observed changes. Thus,

a single mutation in this kind of factors can result in a simultaneous effect on both mRNA synthesis and decay, which might have accelerated yeast evolvability.

It was also proved that the ability of a mutant yeast strain (*rpb6*^{Q100R}) impaired in Rpb4/7 recruitment to module mRNA decay rates during a stress response, in addition to impaired transcription, is compromised. As a result, while the wild type features counteraction coupling between production and degradation of mRNA, this coupling is lost in the mutant [95].

An alternative form of crosstalk between transcription and mRNA stability is mediated by miRNAs, a class of 21–23 nucleotide-long noncoding RNAs that target specific mRNAs by Watson–Crick base-pairing and promote their degradation or modify their translation rates. Genome-wide analyses proved most useful in the experimental identification of miRNA targets (reviewed in [96]). Classical transcriptome profiling in response to specific miRNA gain- or loss-of-function helps define the set of mRNA targets that can be co-regulated by miRNAs with homologous seed regions. Immunoprecipitation of the *Argonauta* protein, which specifically binds miRNAs, in combination to massive RNA sequencing, allows the definition of direct targets of miRNA. The resulting databases of miRNA target genes were utilized to include miRNAs in the definition of gene-expression networks. A good example is the transcriptional response of breast cancer cells to estrogen signaling [67]. Hah et al. [67] found a statistically significant overlapping between estradiol-regulated mRNAs and mRNAs, which are targets of estradiol-regulated miRNAs. Yet they found no evidence of a concerted transcriptional regulation of the miRNAs' precursor genes with the direction (either up or down) of their potential targets mRNAs' regulation by either GRO-seq or mRNA profiling. Evidence for both coordinated and compensatory regulations was found, indicating the additional level of complexity introduced by miRNA into gene expression regulation [67].

5.2. Models for trans factor recognition of mRNA targets

How do mRNA-stability regulatory proteins recognize their mRNA targets? There are, in principle, two places where a given factor can bind to an mRNA: in the nucleus (Fig. 3, blue ovals) and in the cytoplasm (Fig. 3, yellow ovals). In the nucleus, a protein can be loaded
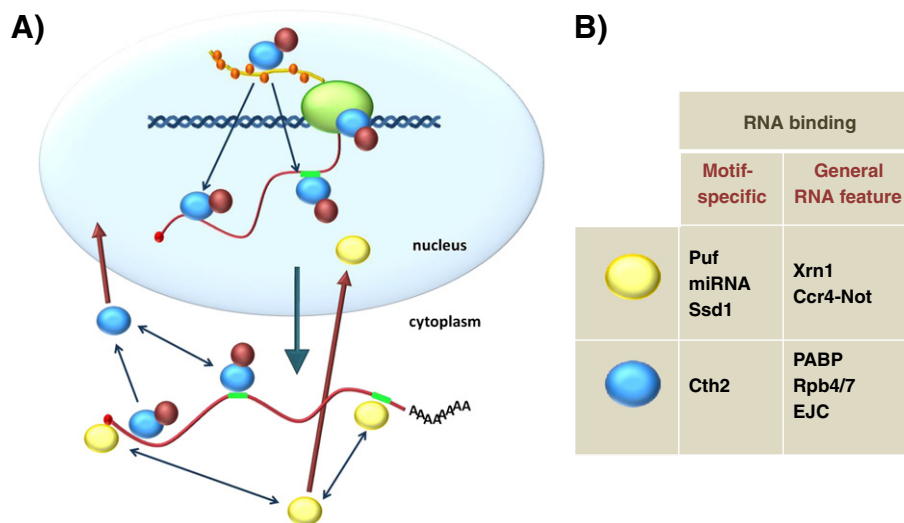

Fig. 3. Models for the *trans* factor recognition of their mRNA targets. A) A transcribing RNA pol II (green oval) may have associated factors (blue ovals) that travel with it (e.g., bound to its CTD, yellow line), which may jump to nascent mRNA (red wave line) or to a target sequence (green box), normally placed at the 3′-UTR, or to any part of the mRNA molecules. Secondary factors (brown ovals) can use primary ones to indirectly bind nascent mRNA. The mRNA charged with factors is exported to the cytoplasm, where they can influence its stability, translation or p-body association. In the cytoplasm, other factors (yellow ovals) may bind mRNAs. They can also be sequence-specific, like miRNA, or not. If factors are bound irrespectively of a target sequence, they can only dissociate from mRNA. The factors that recognize target sequences during transcription, or only in the cytoplasm, establish an association equilibrium. All kinds of factors can be imported into the nucleus when released from the mRNA. This can be used as a quantitative signal for the titration of the cytoplasmic mRNA amount. B) The table shows putative or demonstrated examples of every kind of factor: nuclear or cytoplasmic, which bind specific or general features on mRNAs.

onto mRNA during the transcription process without it recognizing any sequence motif (nucleus left arrow). This seems to be the mechanism of the Rpb4/7 subunits, which can bind only mRNAs in the RNA pol II context [87]. Thus, some general factors which bind many mRNAs may discriminate them to bind on the basis that these mRNAs are current targets of the transcriptional machinery they form part of. More gene-specific factors can be indirectly loaded via their contact with the general ones (Fig. 3, brown ovals) or can bind specific sequences (Fig. 3, nucleus right arrow). All these factors are exported to the cytoplasm as ribonucleoproteins and can be later released from their targets. Sequence nonspecific binding avoids any possible re-binding in the cytoplasm. Those factors can only be imported back to the nucleus or be degraded. A similar mechanism was described for other factors, such as the exon junction complex (EJC), which binds exon junctions only after intron splicing which, once released by the first translation round, cannot return to its target [97]. Cytoplasmatic RNA-binding factors (yellow ovals) can be sequence-specific or not. The sequence motif for both nuclear and cytoplasmatic factors can be more or less specific, but should evolve together with the factor (protein or RNA) that binds it. By using this system, an RBP or an miRNA can bind a target in the nucleus or the cytoplasm at any time following an equilibrium reaction, which can be controlled by other factors or metabolites. To date, many of the RBPs described for post-transcriptional regulons follow this scheme. For instance, the Puf factors described in *Drosophila* and yeast, and other related factors in human cells, bind specific 3′-UTR sequences and promote mRNA degradation [82,83,98]. In contrast, Ssp1 seems to bind 5′-UTR sequences [91]. In mammals, many miRNAs are able to destabilize their target mRNAs, and do so similarly to RBPs [99]. Other cytoplasmatic factors are not sequence-specific, although they can bind special mRNA structures such as the cap [101] or the poly(A) tail. Nonspecific recognition has the advantage of: broad specificity and can be used for general mRNA homeostasis. Cytoplasmatic factors can titrate the total level of cytoplasmic mRNAs and can be travelers in the opposite sense, from the cytoplasm to the nucleus, by carrying this information to the transcription machinery [71].

It is possible that some post-transcriptional regulons can be based on sequence-specific RBPs, which bind their specific target sequences during their transcription of them. One such example is the Cth2 protein which binds its mRNA targets in the nucleus, probably co-transcriptionally, to then promote their degradation during iron deficiency [100]. In this case, RBP's specificity for a group of functionally related genes is based on both its target sequence and the transcriptional machinery reading it, including the transcription factors that control initiation and/or elongation. In this way, transcriptional and post-transcriptional regulons would be interconnected.

Acknowledgements

References

[1] L.A. Selth, S. Sigurdsson, J.Q. Svejstrup, Transcript elongation by RNA polymerase II, Annu. Rev. Biochem. 79 (2010) 271–293.

[2] H. Kim, B. Erickson, W. Luo, D. Seward, J.H. Graber, D.D. Pollock, P.C. Megee, D.L. Bentley, Gene-specific RNA polymerase II phosphorylation and the CTD code, Nat. Struct. Mol. Bio. 17 (2010) 1279–1286.

[3] R. Perales, D. Bentley, "Cotranscriptionality": the transcription elongation complex as a nexus for nuclear transactions, Mol. Cell 36 (2009) 178–191.

[4] R. Parker, H. Song, The enzymes and control of eukaryotic mRNA turnover, Nat. Struct. Mol. Biol. 11 (2004) 121–127.

[5] M.K. Doma, R. Parker, RNA quality control in eukaryotes, Cell 131 (2007) 660–668.

[6] U. Alon, An introduction to systems biology, Design Principles and Biological Circuits, Chapman and Hall/CRC, London, 2007 (ed.).

[7] V. Pelechano, S. Chávez, J.E. Pérez-Ortín, A complete set of nascent transcription rates for yeast genes, PLoS One 5 (2010) e15442.

[8] V. Pelechano, J.E. Pérez-Ortín, There is a steady-state transcriptome in exponentially growing yeast cells, Yeast 27 (2010) 413–422.

[9] E.M. Airoldi, C. Huttenhower, D. Gresham, C. Lu, A.A. Caudy, M.J. Dunham, J.R. Broach, D. Botstein, O.G. Troyanskaya, Predicting cellular growth from gene expression signatures, PLoS Comput. Biol. 5 (2009) e1000257.

[10] N. Slavov, D. Botstein, Coupling among growth rate response, metabolic cycle, and cell division cycle in yeast, Mol. Biol. Cell 22 (2011) 1997–2009.

[11] K. Hirayoshi, J.T. Lis, Nuclear run-on assays: assessing transcription by measuring density of engaged RNA polymerases, Meth. Enzymol. 304 (1999) 351–362.

[12] J. Fan, X. Yang, W. Wang, W.H. Wood III, K.G. Becker, M. Gorospe, Global analysis of stress-regulated mRNA turnover by using cDNA arrays, Proc. Natl. Acad. Sci. U. S. A. 99 (2002) 10611–10616.

[13] J. Fan, M. Zhan, J. Shen, J.L. Martindale, X. Yang, T. Kawai, M. Gorospe, En masse nascent transcription analysis to elucidate regulatory transcription factors, Nucleic Acids Res. 34 (2006) 1492–1500.

[14] S.A. Tenenbaum, C.C. Carson, U. Atasoy, J.D. Keene, Genome-wide regulatory analysis using en masse nuclear run-ons and ribonomic profiling with autoimmune sera, Gene 317 (2003) 79–87.

[15] J. Legen, S. Kemp, K. Krause, B. Profanter, R.G. Herrmann, R.M. Maier, Comparative analysis of plastid transcription profiles of entire plastid chromosomes from tobacco attributed to wild-type and PEP-deficient transcription machineries, Plant J. 31 (2002) 171–188.

[16] J. García-Martínez, A. Aranda, J.E. Pérez-Ortín, Genomic run-on evaluates transcription rates for all yeast genes and identifies gene regulatory mechanisms, Mol. Cell 15 (2004) 303–313.

[17] L.J. Core, J.J. Waterfall, J.T. Lis, Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters, Science 322 (2008) 1845–1848.

[18] T.S. Kim, C.L. Liu, M. Yassour, J. Holik, N. Friedman, S. Buratowski, O.J. Rando, RNA polymerase mapping during stress responses reveals widespread nonproductive transcription in yeast, Genome Biol. 11 (2010) R75.

[19] A. Mayer, M. Lidschreiber, M. Siebert, K. Leike, J. Soding, P. Cramer, Uniform transitions of the general RNA polymerase II transcription complex, Nat. Struct. Mol. Biol. 17 (2010) 1272–1278.

[20] C. Miller, B. Schwalb, K. Maier, D. Schulz, S. Dumcke, B. Zacher, A. Mayer, J. Sydow, L. Marcinowski, L. Dolken, D.E. Martin, A. Tresch, P. Cramer, Dynamic transcriptome analysis measures rates of mRNA synthesis and decay in yeast, Mol. Syst. Biol. 7 (2011) 458.

[21] P. Lefrancois, G.M. Euskirchen, R.K. Auerbach, J. Rozowsky, T. Gibson, C.M. Yellman, M. Gerstein, M. Snyder, Efficient yeast ChIP-Seq using multiplex short-read DNA sequencing, BMC Genomics 10 (2009) 37.

[22] J. Nelson, O. Denisenko, K. Bomsztyk, Profiling RNA polymerase II using the fast chromatin immunoprecipitation method, Methods Mol. Biol. 703 (2011) 219–234.

[23] K. Struhl, Transcriptional noise and the fidelity of initiation by RNA polymerase II, Nat. Struct. Mol. Biol. 14 (2007) 103–105.

[24] J.E. Pérez-Ortín, D.A. Medina, A. Jordán-Pla, Genomic insights into the different layers of gene regulation in yeast, Genet. Res. Int. (2011) 989303.

[25] B. Schwanhausser, D. Busse, N. Li, G. Dittmar, J. Schuchhardt, J. Wolf, W. Chen, M. Selbach, Global quantification of mammalian gene expression control, Nature 473 (2011) 337–342.

[26] M. Rabani, J.Z. Levin, L. Fan, X. Adiconis, R. Raychowdhury, M. Garber, A. Gnirke, C. Nusbaum, N. Hacohen, N. Friedman, I. Amit, A. Regev, Metabolic labeling of RNA uncovers principles of RNA production and degradation dynamics in mammalian cells, Nat. Biotechnol. 29 (2011) 436–442.

[27] Y. Wang, C.L. Liu, J.D. Storey, R.J. Tibshirani, D. Herschlag, P.O. Brown, Precision and functional specificity in mRNA decay, Proc. Natl. Acad. Sci. U. S. A. 99 (2002) 5860–5865.

[28] Y. Arava, Y. Wang, J.D. Storey, C.L. Liu, P.O. Brown, D. Herschlag, Genome-wide analysis of mRNA translation profiles in *Saccharomyces cerevisiae*, Proc. Natl. Acad. Sci. U. S. A. 100 (2003) 3889–3894.

[29] J.R. Warner, The economics of ribosome biosynthesis in yeast, Trends Biochem. Sci. 24 (1999) 437–440.

[30] J. Berretta, A. Morillon, Pervasive transcription constitutes a new level of eukaryotic genome regulation, EMBO Rep. 10 (2009) 973–982.

[31] W. Wei, V. Pelechano, A.I. Jarvelin, L.M. Steinmetz, Functional consequences of bidirectional promoters, Trends Genet. 27 (2011) 267–276.

[32] H. van Bakel, C. Nislow, B.J. Blencowe, T.R. Hughes, Most "dark matter" transcripts are associated with known genes, PLoS Biol. 8 (2010) e1000371.

[33] T.M. Alberola, J. García-Martínez, O. Antúnez, L. Viladevall, A. Barceló, J. Arino, J.E. Pérez-Ortín, A new set of DNA macrochips for the yeast *Saccharomyces cerevisiae*: features and uses, Int. Microbiol. 7 (2004) 199–206.

[34] Z. Xu, W. Wei, J. Gagneur, F. Perocchi, S. Clauder-Munster, J. Camblong, E. Guffanti, F. Stutz, W. Huber, L.M. Steinmetz, Bidirectional promoters generate pervasive transcription in yeast, Nature 457 (2009) 1033–1037.

[35] E.L. van Dijk, C.L. Chen, Y. d'Aubenton-Carafa, S. Gourvennec, M. Kwapisz, V. Roche, C. Bertrand, M. Silvain, P. Legoix-Ne, S. Loeillet, A. Nicolas, C. Thermes,

A. Morillon, XUTs are a class of Xrn1-sensitive antisense regulatory non-coding RNA in yeast, Nature 475 (2011) 114–117.

[36] M. Yassour, J. Pfiffner, J.Z. Levin, X. Adiconis, A. Gnirke, C. Nusbaum, D.A. Thompson, N. Friedman, A. Regev, Strand-specific RNA sequencing reveals extensive regulated long antisense transcripts that are conserved across yeast species, Genome Biol. 11 (2010) R87.

[37] L.S. Churchman, J.S. Weissman, Nascent transcript sequencing visualizes transcription at nucleotide resolution, Nature 469 (2011) 368–373.

[38] H. Neil, C. Malabat, Y. d'Aubenton-Carafa, Z. Xu, L.M. Steinmetz, A. Jacquier, Widespread bidirectional promoters are the major source of cryptic transcripts in yeast, Nature 457 (2009) 1038–1042.

[39] A. Lardenois, Y. Liu, T. Walther, F. Chalmel, B. Evrard, M. Granovskaia, A. Chu, R.W. Davis, L.M. Steinmetz, M. Primig, Execution of the meiotic noncoding RNA expression program and the onset of gametogenesis in yeast require the conserved exosome subunit Rrp6, Proc. Natl. Acad. Sci. U. S. A. 108 (2011) 1058–1063.

[40] M. Kenzelmann, S. Maertens, M. Hergenhahn, S. Kueffer, A. Hotz-Wagenblatt, L. Li, S. Wang, C. Ittrich, T. Lemberger, R. Arribas, S. Jonnakuty, M.C. Hollstein, W. Schmid, N. Gretz, H.J. Grone, G. Schutz, Microarray analysis of newly synthesized RNA in cells and animals, Proc. Natl. Acad. Sci. U. S. A. 104 (2007) 6164–6169.

[41] F.C. Holstege, E.G. Jennings, J.J. Wyrick, T.I. Lee, C.J. Hengartner, M.R. Green, T.R. Golub, E.S. Lander, R.A. Young, Dissecting the regulatory circuitry of a eukaryotic genome, Cell 95 (1998) 717–728.

[42] J.E. Pérez-Ortín, P.M. Alepuz, J. Moreno, Genomics and gene transcription kinetics in yeast, Trends Genet. 23 (2007) 250–257.

[43] M.M. Molina-Navarro, L. Castells-Roca, G. Bellí, J. García-Martínez, J. Marín-Navarro, J. Moreno, J.E. Pérez-Ortín, E. Herrero, Comprehensive transcriptional analysis of the oxidative response in yeast, J. Biol. Chem. 283 (2008) 17908–17918.

[44] L. Romero-Santacreu, J. Moreno, J.E. Pérez-Ortín, P. Alepuz, Specific and global regulation of mRNA stability during osmotic stress in Saccharomyces cerevisiae, RNA 15 (2009) 1110–1120.

[45] B. Hayles, S. Yellaboina, D. Wang, Comparing transcription rate and mRNA abundance as parameters for biochemical pathway and network analysis, PLoS One 5 (2010) e9908.

[46] J.E. Pérez-Ortín, A. Jordán-Pla, V. Pelechano, A genomic view of mRNA turnover in yeast, C. R. Biol. 334 (2011) 647–654.

[47] F. Exinger, F. Lacroute, 6-Azauracil inhibition of GTP biosynthesis in Saccharomyces cerevisiae, Curr. Genet. 22 (1992) 9–11.

[48] D.A. Glesne, F.R. Collart, E. Huberman, Regulation of IMP dehydrogenase gene expression by its end products, guanine nucleotides, Mol. Cell. Biol. 11 (1991) 5417–5425.

[49] M. Morillo-Huesca, M. Vanti, S. Chávez, A simple in vivo assay for measuring the efficiency of gene length-dependent processes in yeast mRNA biogenesis, FEBS J. 273 (2006) 756–769.

[50] C. Tous, A.G. Rondón, M. García-Rubio, C. González-Aguilera, R. Luna, A. Aguilera, A novel assay identifies transcript elongation roles for the Nup84 complex and RNA processing factors, EMBO J. 30 (2011) 1953–1964.

[51] P.B. Mason, K. Struhl, Distinction and relationship between elongation rate and processivity of RNA polymerase II in vivo, Mol. Cell 17 (2005) 831–840.

[52] M.L. Kireeva, B. Hancock, G.H. Cremona, W. Walter, V.M. Studitsky, M. Kashlev, Nature of the nucleosomal barrier to RNA polymerase II, Mol. Cell 18 (2005) 97–108.

[53] M. Voliotis, N. Cohen, C. Molina-Paris, T.B. Liverpool, Fluctuations, pauses, and backtracking in DNA transcription, Biophys. J. 94 (2008) 334–348.

[54] N. Terzi, L.S. Churchman, L. Vasiljeva, J. Weissman, S. Buratowski, H3K4 trimethylation by Set1 promotes efficient termination by the Nrd1-Nab3-Sen1 pathway, Mol. Cell. Biol. 31 (2011) 3569–3583.

[55] F. Carrillo Oesterreich, S. Preibisch, K.M. Neugebauer, Global analysis of nascent RNA reveals transcriptional pausing in terminal exons, Mol. Cell 40 (2010) 571–581.

[56] M.G. Guenther, S.S. Levine, L.A. Boyer, R. Jaenisch, R.A. Young, A chromatin landmark and transcription initiation at most promoters in human cells, Cell 130 (2007) 77–88.

[57] G.W. Muse, D.A. Gilchrist, S. Nechaev, R. Shah, J.S. Parker, S.F. Grissom, J. Zeitlinger, K. Adelman, RNA polymerase is poised for activation across the genome, Nat. Genet. 39 (2007) 1507–1511.

[58] J. Zeitlinger, A. Stark, M. Kellis, J.W. Hong, S. Nechaev, K. Adelman, M. Levine, R.A. Young, RNA polymerase stalling at developmental control genes in the Drosophila melanogaster embryo, Nat. Genet. 39 (2007) 1512–1516.

[59] L.R. Baugh, J. Demodena, P.W. Sternberg, RNA Pol II accumulates at promoters of growth genes during developmental arrest, Science 324 (2009) 92–94.

[60] B.E. Bernstein, T.S. Mikkelsen, X. Xie, M. Kamal, D.J. Huebert, J. Cuff, B. Fry, A. Meissner, M. Wernig, K. Plath, R. Jaenisch, A. Wagschal, R. Feil, S.L. Schreiber, E.S. Lander, A bivalent chromatin structure marks key developmental genes in embryonic stem cells, Cell 125 (2006) 315–326.

[61] E. Shema, J. Kim, R.G. Roeder, M. Oren, RNF20 inhibits TFIIS-facilitated transcriptional elongation to suppress pro-oncogenic gene expression, Mol. Cell 42 (2011) 477–488.

[62] M. Radonjic, J.C. Andrau, P. Lijnzaad, P. Kemmeren, T.T. Kockelkorn, D. van Leenen, N.L. van Berkum, F.C. Holstege, Genome-wide analyses reveal RNA polymerase II located upstream of genes poised for rapid response upon S. cerevisiae stationary phase exit, Mol. Cell 18 (2005) 171–183.

[63] C. Martens, B. Krett, P.J. Laybourn, RNA polymerase II and TBP occupy the repressed CYC1 promoter, Mol. Microbiol. 40 (2001) 1009–1019.

[64] M. Kwapisz, M. Wery, D. Despres, Y. Ghavi-Helm, J. Soutourina, P. Thuriaux, F. Lacroute, Mutations of RNA polymerase II activate key genes of the nucleoside triphosphate biosynthetic pathways, EMBO J. 27 (2008) 2411–2421.

[65] J.N. Kuehner, D.A. Brow, Regulation of a eukaryotic gene by GTP-dependent start site selection and transcription attenuation, Mol. Cell 31 (2008) 201–211.

[66] J.T. Wade, K. Struhl, The transition from transcriptional initiation to elongation, Curr. Opin. Genet. Dev. 18 (2008) 130–136.

[67] N. Hah, C.G. Danko, L. Core, J.J. Waterfall, A. Siepel, J.T. Lis, W.L. Kraus, A rapid, extensive, and transient transcriptional response to estrogen signaling in breast cancer cells, Cell 145 (2011) 622–634.

[68] A. Rodriguez-Gil, J. García-Martínez, V. Pelechano, L. Munoz-Centeno Mde, V. Geli, J.E. Pérez-Ortín, S. Chávez, The distribution of active RNA polymerase II along the transcribed region is gene-specific and controlled by elongation factors, Nucleic Acids Res. 38 (2010) 4651–4664.

[69] V. Pelechano, S. Jimeno-González, A. Rodriguez-Gil, J. García-Martínez, J.E. Pérez-Ortín, S. Chávez, Regulon-specific control of transcription elongation across the yeast genome, PLoS Genet. 5 (2009) e1000614.

[70] K. Adelman, M.T. Marr, J. Werner, A. Saunders, Z. Ni, E.D. Andrulis, J.T. Lis, Efficient release from promoter-proximal stall sites requires transcript cleavage factor TFIIS, Mol. Cell 17 (2005) 103–112.

[71] J.A. Kruk, A. Dutta, J. Fu, D.S. Gilmour, J.C. Reese, The multifunctional Ccr4–Not complex directly promotes transcription elongation, Genes Dev. 25 (2011) 581–593.

[72] E.A. Galburt, S.W. Grill, A. Wiedmann, L. Lubkowska, J. Choy, E. Nogales, M. Kashlev, C. Bustamante, Backtracking determines the force sensitivity of RNAP II in a factor-dependent manner, Nature 446 (2007) 820–823.

[73] T.N. Mavrich, C. Jiang, I.P. Ioshikhes, X. Li, B.J. Venters, S.J. Zanton, L.P. Tomsho, J. Qi, R.L. Glaser, S.C. Schuster, D.S. Gilmour, I. Albert, B.F. Pugh, Nucleosome organization in the Drosophila genome, Nature 453 (2008) 358–362.

[74] B.J. Venters, S. Wachi, T.N. Mavrich, B.E. Andersen, P. Jena, A.J. Sinnamon, P. Jain, N.S. Rolleri, C. Jiang, C. Hemeryck-Walsh, B.F. Pugh, A comprehensive genomic binding map of gene and chromatin regulatory proteins in Saccharomyces, Mol. Cell 41 (2011) 480–492.

[75] J.D. Keene, S.A. Tenenbaum, Eukaryotic mRNPs may represent posttranscriptional operons, Mol. Cell 9 (2002) 1161–1167.

[76] O. Dahan, H. Gingold, Y. Pilpel, Regulatory mechanisms and networks couple the different phases of gene expression, Trends Genet. 27 (2011) 316–322.

[77] J. Grigull, S. Mnaimneh, J. Pootoolal, M.D. Robinson, T.R. Hughes, Genome-wide analysis of mRNA stability using transcription inhibitors and microarrays reveals posttranscriptional control of ribosome biogenesis factors, Mol. Cell. Biol. 24 (2004) 5534–5547.

[78] L. Castells-Roca, J. García-Martínez, J. Moreno, E. Herrero, G. Bellí, J.E. Pérez-Ortín, Heat shock response in yeast involves changes in both transcription rates and mRNA stabilities, PLoS One 6 (2011) e17272.

[79] O. Shalem, O. Dahan, M. Levo, M.R. Martínez, I. Furman, E. Segal, Y. Pilpel, Transient transcriptional responses to stress are generated by opposing effects of mRNA production and degradation, Mol. Syst. Biol. 4 (2008) 223.

[80] C. Molin, A. Jauhiainen, J. Warringer, O. Nerman, P. Sunnerhagen, mRNA stability changes precede changes in steady-state mRNA amounts during hyperosmotic stress, RNA 15 (2009) 600–614.

[81] M. Chatenay-Lapointe, G.S. Shadel, Repression of mitochondrial translation, respiration and a metabolic cycle-regulated gene, SLF1, by the yeast Pumilio-family protein Puf3p, PLoS One 6 (2011) e20441.

[82] A.P. Gerber, D. Herschlag, P.O. Brown, Extensive association of functionally and cytotopically related mRNAs with Puf family RNA-binding proteins in yeast, PLoS Biol. 2 (2004) E79.

[83] D.J. Hogan, D.P. Riordan, A.P. Gerber, D. Herschlag, P.O. Brown, Diverse RNA-binding proteins interact with functionally related sets of RNAs, suggesting an extensive regulatory system, PLoS Biol. 6 (2008) e255.

[84] A.D. Aragon, G.A. Quinones, E.V. Thomas, S. Roy, M. Werner-Washburne, Release of extraction-resistant mRNA in stationary phase Saccharomyces cerevisiae produces a massive increase in transcript abundance in response to stress, Genome Biol. 7 (2006) R9.

[85] G. Jona, M. Choder, O. Gileadi, Glucose starvation induces a drastic reduction in the rates of both transcription and degradation of mRNA in yeast, Biochim. Biophys. Acta 1491 (2000) 37–48.

[86] R. Lotan, V. Goler-Baron, L. Duek, G. Haimovich, M. Choder, The Rpb7p subunit of yeast RNA polymerase II plays roles in the two major cytoplasmic mRNA decay mechanisms, J. Cell Biol. 178 (2007) 1133–1143.

[87] V. Goler-Baron, M. Selitrennik, O. Barkai, G. Haimovich, R. Lotan, M. Choder, Transcription in the nucleus and mRNA decay in the cytoplasm are coupled processes, Genes Dev. 22 (2008) 2022–2027.

[88] L. Harel-Sharvit, N. Eldad, G. Haimovich, O. Barkai, L. Duek, M. Choder, RNA polymerase II subunits link transcription and mRNA decay to translation, Cell 143 (2010) 552–563.

[89] M. Farago, T. Nahari, C. Hammel, C.N. Cole, M. Choder, Rpb4p, a subunit of RNA polymerase II, mediates mRNA export during stress, Mol. Biol. Cell 14 (2003) 2744–2755.

[90] R. Lotan, V.G. Bar-On, L. Harel-Sharvit, L. Duek, D. Melamed, M. Choder, The RNA polymerase II subunit Rpb4p mediates decay of a specific class of mRNAs, Genes Dev. 19 (2005) 3004–3016.

[91] C. Kurischko, V.K. Kuravi, C.J. Herbert, F.C. Luca, Nucleocytoplasmic shuttling of Ssd1 defines the destiny of its bound mRNAs, Mol. Microbiol. 81 (2011) 831–849.

[92] M. Selitrennik, L. Duek, R. Lotan, M. Choder, Nucleocytoplasmic shuttling of the Rpb4p and Rpb7p subunits of Saccharomyces cerevisiae RNA polymerase II by two pathways, Eukaryot. Cell 5 (2006) 2092–2103.

[93] M. Choder, mRNA imprinting: additional level in the regulation of gene expression, Cell Logist. 1 (2011) 37–40.

[94] M. Dori-Bachash, E. Shema, I. Tirosh, Coupled evolution of transcription and mRNA degradation, PLoS Biol. 9 (2011) e1001106.

[95] O. Shalem, B. Groisman, M. Choder, O. Dahan, Y. Pilpel, Transcriptome kinetics is governed by a genome-wide coupling of mRNA production and degradation: a role for RNA Pol II, PLoS Genet. 7 (2011) e1002273.

[96] M.L. Wilbert, G.W. Yeo, Genome-wide approaches in the study of microRNA biology, Wiley Interdiscip Rev. Syst. Biol. Med. 3 (2011) 491–512.

[97] Y.F. Chang, J.S. Imam, M.F. Wilkinson, The nonsense-mediated decay RNA surveillance pathway, Annu. Rev. Biochem. 76 (2007) 51–74.

[98] A. Filipovska, M.F. Razif, K.K. Nygard, O. Rackham, A universal code for RNA recognition by PUF proteins, Nat. Chem. Biol. 7 (2011) 425–427.

[99] H. Guo, N.T. Ingolia, J.S. Weissman, D.P. Bartel, Mammalian microRNAs predominantly act to decrease target mRNA levels, Nature 466 (2010) 835–840.

[100] S.V. Vergara, S. Puig, D.J. Thiele, Early recruitment of AU-rich element-containing mRNAs determines their cytosolic fate during iron deficiency, Mol. Cell. Biol. 31 (2011) 417–429.

[101] S. Lahudkar, A. Shukla, P. Bajwa, G. Durairaj, N. Stanojevic, S.R. Bhaumik, The mRNA cap-binding complex stimulates the formation of pre-initiation complex at the promoter via its interaction with Mot1p in vivo, Nucleic Acids Res. 39 (2011) 2188–2209.