Published online 28 October 2014 *Nucleic Acids Research, 2014, Vol. 42, No. 20* ***12847–12860***
doi: 10.1093/nar/gku1014

Quantification of pre-mRNA escape rate and synergy in splicing

Marie Mi Bonde[1], Sylvia Voegeli[1], Antoine Baudrimont[1], Bertrand Séraphin[2] and Attila Becskei[1,*]

[1]Biozentrum, University of Basel, Klingelbergstrasse 50/70, 4056 Basel, Switzerland and [2]Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Centre National de Recherche Scientifique (CNRS) UMR 7104, Institut National de Santé et de Recherche Médicale (INSERM) U964, Université de Strasbourg, Illkirch, Strasbourg, France

Received May 06, 2014; Revised October 05, 2014; Accepted October 07, 2014

ABSTRACT

Splicing reactions generally combine high speed with accuracy. However, some of the pre-mRNAs escape the nucleus with a retained intron. Intron retention can control gene expression and increase proteome diversity. We calculated the escape rate for the yeast *PTC7* intron and pre-mRNA. This prediction was facilitated by the observation that splicing is a linear process and by deriving simple algebraic expressions from a model of co- and post-transcriptional splicing and RNA surveillance that determines the rate of the nonsense-mediated decay (NMD) of the pre-mRNAs with the retained intron. The escape rate was consistent with the observed threshold of splicing rate below which the mature mRNA level declined. When an mRNA contains multiple introns, the outcome of splicing becomes more difficult to predict since not only the escape rate of the pre-mRNA has to be considered, but also the possibility that the splicing of each intron is influenced by the others. We showed that the two adjacent introns in the *SUS1* mRNA are spliced cooperatively, but this does not counteract the escape of the partially spliced mRNA. These findings will help to infer promoter activity and to predict the behavior of and to control splicing regulatory networks.

INTRODUCTION

Splicing can control both transcript levels and proteome diversity (1). The understanding and prediction of how splicing rate affects precursor and mature mRNA levels and how this process can be controlled are major goals of systems biology of gene expression.

Splicing of most introns proceeds with high speed and accuracy (2). However, the splicing efficiency of different mRNAs spans a very broad range (3) and some of the pre-mRNA molecules escape the nucleus without further chance of being spliced. Since a typical intron contains in frame stop codons that will be recognized as premature (premature stop codon, PMS), pre-mRNAs reaching the cytoplasm are recognized by surveillance mechanism, targeting them to nonsense-mediated decay (NMD) to limit their translation (4). Splicing can occur not only post-transcriptionally but also co-transcriptionally while the mRNA is tethered to the DNA, which may reduce the escape rate of the mRNA. Varying contributions of the co- and post-transcriptional stages have been suggested for the completion of splicing (5–9).

When an mRNA contains multiple introns, the outcome of splicing becomes more difficult to predict since the escape of the pre-mRNA can be affected by inhibitory or cooperative interactions between the splicing of introns in the same mRNA molecule.

Simple measures have been playing an important role in various fields of biology and biochemistry since they permit parsimonious model building and consistent comparison of experimental data. For example, the Hill coefficient reflects the cooperative enhancement of binding of multiple ligands to a protein relative to a single ligand, epitomized by the synergistic binding of oxygen to hemoglobin (10).

A simple measure, the mature-to-precursor (M/P) ratio was introduced to estimate the splicing rate in steady-state conditions (11). $M/P = \sigma/\mu$, where σ is the splicing rate and μ is the decay rate of the mature mRNA. Thus, splicing rate is proportional to the measured M/P ratio, also termed splicing efficiency. Since the decay of mRNAs follows typically a simple exponential kinetics (12,13), the splicing rate can be determined from the experimentally measured decay rate and steady-state M/P ratio.

In this study, we aimed at defining measures or simple expressions to quantitate the escape rate of mRNA from splicing, and to assess synergy in splicing of mRNAs with multiple introns. The estimation of escape rate is particularly important for those mRNAs that undergo alternative splicing by intron retention, yielding functionally different proteins.

*To whom correspondence should be addressed. Tel: +41 61 267 2222; Fax: +41 61 267 2189; Email: attila.becskei@unibas.ch

For these genes, the rate of splicing relative to the escape will directly affect the ratio of the protein isoforms arising from the precursor and mature mRNAs. For this reason, we have focused on the *PTC7* gene, which has a unique intron in the yeast genome since the precursor mRNA lacks PMSs and is not targeted to NMD (14). It is inefficiently spliced and the precursor and mature forms are present at comparable levels, both of them being translated to proteins, which have distinct subcellular localization: the spliced version localizes to mitochondria, while the unspliced to the nuclear envelope (14).

Similar to *PTC7*, the *SUS1* mRNA has an inefficiently spliced intron with a non-consensus sequence. In addition, it contains a second efficiently spliced intron downstream. *SUS1* is one of the 10 genes known to contain two introns in the *Saccharomyces cerevisiae* genome (15,16). The splicing of the two introns has been thought to be polar. The polar mechanism is thought to arise due to the directional nature of transcription: splicing of an intron can enhance the splicing of a downstream intron (17). The protein translated from the partially spliced *SUS1* RNA species, containing the non-consensus intron, has an impact on histone deubiquitylation, indicating that the pre-mRNA escape plays also an important role in controlling a cellular function (15). We focused on *SUS1* gene to disentangle the effects of synergistic splicing and escape from splicing.

Since processes and models have been suggested that may result in non-linear splicing efficiency (18), we first confirmed that splicing can be approximated by linear reaction terms. Subsequently, we derived simple algebraic expressions, and relying on these, we show that the M/P ratios of appropriate series of genetic constructs can be utilized to assess whether post-transcriptional splicing is dominant or co- and post-transcriptional splicing occur jointly. The model consistent with the experimental data requires the inclusion of co- and post-transcriptional splicing in addition to the NMD pathway. Subsequently, we proceeded with the central aims of this study: to estimate the escape rate of pre-mRNAs and the degree of synergy in splicing.

MATERIALS AND METHODS

Yeast strains and genetic constructs

For the insertion of introns into *GAL1*, a pair of restriction sites (*Cla*I - *Nhe*I) were inserted downstream of the 111th base of the *GAL1* ORF. These intron-in-*GAL1* constructs were integrated into the chromosome of BY4741 WT or upf1Δ::kanMX strains upon linearization with *Mfe*I (Supplementary Table S1). For titration of expression with estradiol, they were mated with the corresponding MAT alpha strain, containing gal1Δ::kanMX, MRP7::pRS303 (HIS3) P_{MRP7}-GEV (see also Supplementary Information). To maintain the *PTC7* introns in frame with the *GAL1* ORF, the introns were flanked by adding 5′G and 3′GA bases, as in the endogenous *PTC7* gene. In order to have the slowest and most reliably measurable decay rate, a PMS causing the least decrease in the M/P ratio of the *PTC7* RNA was selected.

Promoters of endogenous genes were replaced with a part of the *GAL1* promoter using pop-in/pop-out recombination (transplacement). The endogenous promoter sequences of a length indicated in Supplementary Tables S1 and S2, were deleted to keep the 5′UTR region of the endogenous gene intact. The *GAL7* terminator preceded the *GAL1* promoter to prevent transcriptional interference. When no unique restriction site for linearization was available, an *Aar*I self-eliminating element was inserted. The element consists of two *Aar*I recognition sites flanked by four random bases on each side. Four random bases are inserted between the two recognition sites to make sure both sites can be bound by the restriction enzyme.

For the transplacement, the respective plasmids were transformed into BY4742, containing the MRP7::pRS303 (HIS3) P_{MRP7}-GEV (Supplementary Figure S1). To remove the plasmid sequence, the expression of *URA3* was counterselected. Briefly, strains were grown overnight in YEP supplemented with adenine sulfate, 2% raffinose, 0.005% glucose and 0.05% galactose. Strains were plated on media with 1 g/l 5-fluoro-orotic acid (FOA), 50 mg/l uracil, 2% raffinose, 0.005% glucose and 0.05% galactose. Transplacements were screened by colony polymerase chain reaction (PCR).

To buffer the adverse effects when the expression of the promoter-replacement construct is low, a centromeric plasmid with the intron-containing gene driven by its natural promoter was transformed into the strains (Supplementary Table S2). The silent mutations to distinguish their RNA were introduced typically 100 bp downstream of the intron. The P_{GAL1}-*SUS1* strains were not supplemented with the centromeric plasmid to allow detection of alternatively spliced isoforms.

Growth conditions for steady-state and RNA decay measurements

Cells were grown aerobically at 30°C. The culture medium consisted of minimal drop-out medium to maintain selection of the transformed constructs. The carbon source was provided by 2% raffinose (sterilized by filtration) supplemented with 0.005% glucose. For induction of expression, overnight cultures were re-grown in the same medium and 0.5% galactose or estradiol was added. Induction is rapid with both estradiol (Supplementary Figure S2) and galactose (12); around half of the steady-state expression level is reached within 30 min. For steady-state experiments, cells were harvested 5 h later upon reaching an OD_{600} of 0.5–1 and stored frozen in liquid nitrogen.

For the decay measurement, the overnight cultures were supplemented with 0.5% galactose to induce expression. The cultures were then refreshed in 0.04% galactose. Transcription was shut down by addition of 2% glucose. For most of the genes, the decay rate measured in this way is equal to the ones obtained by shut down with other means, such as by doxycycline (12). The cells were harvested by adding 3 ml of cultures to equal volume of methanol sitting on dry ice as described in (18).

RNA isolation and reverse transcriptase-polymerase chain reaction

Quantitation of RNA was performed using reverse transcription with a mixture of random and oligodT primers,

as described previously (12), except that the RNeasy Mini kit (Qiagen) was also used for isolation of total RNA and Cp values (Roche Lightcycler 480 II) were used to calculate expression. Relative expression was calculated with respect to *ACT1* total mRNA, unless otherwise specified. Relative expression levels were corrected for differences in amplification efficiency of each pair of PCR primer set, which was determined from the Cp values obtained from a 1:10 dilution series of the cDNA (at maximum induction), up to at least a 1000 times dilution for all intron-in-*GAL1* constructs and genes with replaced promoters, except for the double intron genes (Supplementary Tables S3 and S4). To allow for these broad ranges, the dilution series was started from the maximally induced construct.

The precursor mRNA can have very low levels at low estradiol concentrations. To check for non-specific PCR products, we measured the precursor level in the negative control, containing the p*GAL1-GAL1* expression cassette without intron. The signal generated by primers in the negative control did not exceed an expression level of 10^{-5} relative to the *ACT1* mRNA.

Single-molecule RNA fluorescent *in situ* hybridization (sm-FISH)

Overnight culture (SC-URA with 0.5% galactose, 2% raffinose, 0.005% glucose) was refreshed ($OD_{600} = 0.2$) for 3 h before fixation in 3.7% formaldehyde. Cell preparation and hybridization of the probes were performed as described (19), with the following modifications. Coverslips were polylysine coated (20). Oligonucleotides for the *GAL1* sequence were labeled with CAL Fluor Red 610 (Alexa Fluor 594 replacement; 25 μM, Stellaris). After permeabilization of spheroplasts, the probes were applied to the samples at a dilution of 1:200 overnight at 40°C in a humid chamber. After the hybridization of the probes and washing of the samples, cells were mounted in Prolong® Gold media (Life Technologies) and let cure overnight at room temperature before imaging.

Images were acquired as stacks of optical sections at 0.2 μm intervals using a Deltavision deconvolution microscopy system (Applied Precision) under the following conditions: DAPI, FITC, Alexa 594 channels, 100% ND, bin 1, exposure time 80, 2000, 500 ms, respectively, objective 100×. Three-dimensional stacks of images were deconvolved using the Deltavision deconvolution microscopy system and further filtered (background subtraction, median, convolve) using ImageJ to enhance the signal-to-noise ratio of detected spots. Autofluorescence of the cells (FITC) was used to segment the pictures with ImageJ and define cell boundaries.

Using FISH-quant toolbox in Matlab (21), the number of mature mRNAs were scored in segmented cells after the adjustment of the detection threshold. Spots located inside the nucleus (DAPI) and displaying intensity two times higher than the average spot intensity were selected and further quantified as nascent using the point-spread function superposition approach (21). At least 200 cells were quantified for each experiment.

Data analysis

The relative promoter activity was measured by dividing reference RNA levels at a given estradiol concentration by the respective RNA levels determined at the maximal applied estradiol concentration (100 nM). The reference was *GAL1* for the promoter-replacement constructs and *GAL7* for the intron-in-*GAL1* constructs.

For the determination of the range at which splicing efficiency is constant, a Hill function was fitted to the M/P versus promoter activity data using non-linear regression with relative weighting ($1/Y^2$). The lower limit of constancy was defined to correspond to the k value of the Hill function, which in this context is devoid of any physical meaning typically associated with the Hill function.

$$\text{Hill} = b + v\frac{x^n}{k^n + x^n}$$

For fitting of exponential decay, non-linear regression with relative weighting was used. Different basal expressions were allowed to account for the non-linearity in splicing at low RNA levels. A reference gene (*GAL1, GAL7*) was used to check the consistency of the decay experiments and the timing of each decay curve was normalized according to the reference gene.

RESULTS

Genetic constructions

In this study, gene expression constructs were created that permit both the measurement of mRNA decay and the assessment of linearity of splicing rates as gene expression is varied. These constructs were placed under the control of the *GAL1* promoter and were regulated by the transcription factors Gal4p or GEV, both of which contain the Gal4 DNA-binding domains (see Methods). The endogenous Gal4p mediates the galactose and glucose signal to allow the rapid shut-down of expression and to measure RNA decay (12). GEV was used to adjust the *GAL1* promoter activity by varying the concentration of its ligand, estradiol. GEV is a fusion protein of the Gal4p DNA-binding domain, the estradiol receptor and the VP16 activation domain. Thus, it binds to the same promoters as Gal4p. Unlike Gal4p, GEV generates a graded response in wt cells (12,22), which is necessary to assess the constancy of splicing efficiency.

We have designed two classes of intron-containing constructs. In the synthetic constructs, a single intron was inserted into the *GAL1* gene, which otherwise lacks introns (Figure 1A). This intron-in-*GAL1* series allows the comparison of introns with different splicing rates so that most other parameters (such as production and mRNA decay rates) are expected to be kept constant. In order to study endogenous intron-containing genes of the yeast *S. cerevisiae*, we replaced their promoters completing two rounds of homologous recombination (Supplementary Figure S1). In this way, the promoter region upstream of the 5′UTR of the endogenous yeast gene was replaced by the region of the *GAL1* promoter containing the UAS and the TATA box of the *GAL1* promoter (Figure 1B). Since all the residual plasmid sequences were looped out, the endogenous promoter

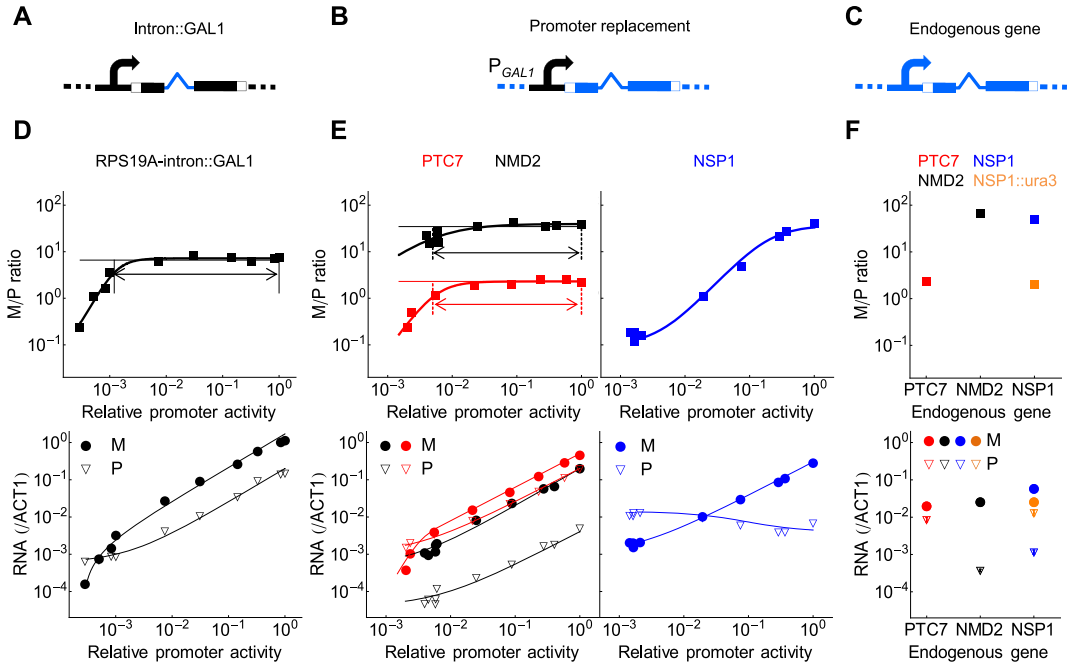

Figure 1. Genetic constructs and assessment of the constancy of splicing efficiency as mRNA expression is varied. (**A**) The intron-in-*GAL1* constructs are integrated into the *GAL1* locus. (**B**) In the promoter replacement constructs, the 5′ UTRs (empty squares) of the endogenous genes are preserved; only the promoter regions upstream of the 5′UTR are replaced by the *GAL1* promoter region. (**C**) The endogenous genes are at their original chromosomal location except for NSP1::ura3, which denotes the *NSP1* gene transferred to the ura3 locus. (**D-E**) Relative promoter activity is the ratio of RNA levels in a given condition to the maximally induced RNA level (estradiol = 100 nM). The range of constant splicing efficiency is indicated by arrows in plots showing the M/P ratios. The lower panels show the corresponding values of mature and precursor mRNAs. (D) The reference gene is *GAL7*. The following values were fitted for the Hill function: $n = 2.4$, $k = 0.0012$, $v = 7.1$ for the *RPS19A* construct. (E) The plasmid-free promoter replacement constructs (P_{GAL1}-*PTC7*, P_{GAL1}-*NMD2*, P_{GAL1}-*NSP1*) were supplemented with a centromeric plasmid expressing the respective endogenous gene with silent mutations (P_{NMD2}-*NMD2**, P_{NSP1}-*NSP1**, P_{PTC7}-*PTC7**). The reference gene is *GAL1*. The following values were fitted for the Hill function: $n = 1$, $k = 0.0052$, $v = 39$ for *NMD2* and $n = 2$, $k = 0.0054$, $v = 2.3$ for *PTC7*. (**F**) Measurements of endogenous RNAs (mean and SD, $n \geq 3$) .

is replaced precisely at the natural chromosomal location of each gene. These constructs can be directly compared to the endogenous yeast genes driven by their own promoters (Figure 1C).

Linearity of splicing efficiency

First we aimed at assessing if splicing can be described by linear reaction terms. For this purpose, we measured the ratio of mature to precursor mRNA, the M/P ratio to estimate splicing efficiency as the activity of the *GAL1* promoter was varied. We tested both intron-in-*GAL1* (with the ribosomal *RPS19A* introns) constructs and the endogenous *PTC7* gene with a replaced promoter. As promoter activity decreased, the M/P ratio was constant for both mRNAs (Figure 1D and E) in steady-state conditions (Supplementary Figure S2, Materials and Methods), suggesting that splicing can be approximated by a linear process over at least two orders of magnitude of mRNA expression levels mRNAs. Below this range, the M/P ratio declined.

The constancy of the M/P ratio permits splicing to be approximated by a first-order chemical reaction. Before proceeding with building models of splicing, we wanted to see if mRNAs exist at all that fail to display a broad range of constant splicing efficiency. We hypothesized that a sign of non-linear splicing activity could be that upon an environmental stimulus, the level of mature mRNA changes more than that of the total or precursor mRNA. Using a database on mRNA responses upon ethanol exposure (23), we identified two genes satisfying the aforementioned criteria—*NMD2* and *NSP1* (Supplementary Table S5)—and we analyzed them further upon promoter replacement. The M/P ratio of *NMD2* was constant for a broad range of expression levels, similar to the previous constructs (Figure 1E). On the other hand, the M/P ratio of *NSP1* displayed a continuous downward trend as the *GAL1* promoter activity diminished (Figure 1E). The mature mRNA level followed the promoter activity similar to the other constructs, but the precursor level of *NSP1* was nearly constant even when the promoter activity was varied over three orders of magnitude (Figure 1E). The declining splicing efficiency cannot be explained by potential adverse effect of an insufficiently expressed gene since a back-up copy of *NSP1* was expressed from a centromeric plasmid, which contained silent mutations in order to distinguish its mRNA from the one expressed by the promoter-replacement construct.

When the *NSP1* intron was inserted into the *GAL1* gene, we observed a constant M/P ratio, suggesting it is not the intron alone that generates the non-linear splicing (Supplementary Figure S3). Interestingly, when the *NSP1* gene—including its promoter and transcriptional terminator region—was integrated into the ura3 chromosomal locus, the M/P ratio was substantially lower than that of the endogenous gene, even though its expression did not change, as judged from the mature mRNA levels (Fig-

ure 1F). This finding suggests that the chromosomal context of the gene, including the promoter sequence, plays a role in generating the non-linear response. Having found one gene, *NSP1,* with variable splicing efficiency, suggests that the constancy of M/P ratios has to be assessed for all genes to be studied.

Importantly, the M/P ratios of all genes with promoter replacement were comparable to that of the respective endogenous genes, even in the case of *NSP1* when the *GAL1* promoter activity was high (Figure 1 E and F). In the subsequent experiments, we measure steady-state M/P ratios at maximum expression levels and fit mRNA decay rates from an RNA concentration range of two orders magnitude, which coincides with the range of constant splicing efficiency.

The M/P ratio can be used to distinguish models of splicing and surveillance mechanisms

Using linear reaction terms, we constructed two models to account for the escape of precursor from splicing and its cytoplasmic decay by the NMD pathway. We compared these with the existing nuclear surveillance model (Figure 2A) (11). In the simpler model, the precursor is competed for by the splicing (σ) and escape reactions (ε) (Figure 2B). In the extended model, two stages of splicing are separated so that splicing can occur co- and post-transcriptionally. The nascent precursor cannot evade splicing by escape to the cytoplasm since the nascent RNA is tethered to the DNA (Figure 2C). Only the precursor that underwent transcriptional termination can evade splicing.

In these models, the same precursor mRNA species can have multiple localizations, each of which is defined by a specific set of processes/reactions. The number of these precursor subsets is different in each model. It is one in the nuclear surveillance model (P). It is two when the surveillance mechanism is represented by NMD (P_D, P_R), and the number increases to three (P_D, P_R and P_N), when co- and post-transcriptional stages of splicing are treated separately. Henceforth, we shall refer to the three models as the one-, two- and three-precursor models.

Next, we wanted to see which of the three models describes most accurately the experimental data and what types of measurements are the most informative to distinguish the model predictions. All three models predict—nearly indistinguishably—that the mature mRNA level declines when the splicing rate is reduced (Figure 3A). On the other hand, the three models can be distinguished by focusing on the precursor so that we compare the M/P ratios measured in wild-type cells, M/P(WT), to that in cells in which one of the components of NMD is mutated, M/P(NoSurv). We assume that the decay rate of the precursor mRNA in NMD mutant cells equals the decay rate of the mature mRNA.

For the two-precursor model, we obtained the following ratio:

$$\frac{\text{M/P(WT)}}{\text{M/P(NoSurv)}} = \frac{\mu_{PR}(\mu_M + \varepsilon)}{\mu_M(\mu_{PR} + \varepsilon)} > 1 \qquad (1)$$

On the other hand, the three-precursor model yields a M/P (wt)/M/P (NoSurv) ratio that is dependent on the post-transcriptional splicing rate:

$$\frac{\text{M/P(WT)}}{\text{M/P(NoSurv)}} = \frac{\mu_{PR}}{\mu_M}\frac{\mu_M + f(\varepsilon',k,\sigma)}{\mu_{PR} + f(\varepsilon',k,\sigma)} \qquad (2)$$
$$\text{where } f(\varepsilon', k, \sigma) = \frac{k\varepsilon'}{\sigma_D + \varepsilon' + k}$$

From the above algebraic expressions, robust predictions can be made that are independent of the particular parameter values. The above ratio is always one for the nuclear surveillance model (Figure 3B, Equation (S2) in the Supplementary Information). It is always larger than one for the two-precursor model (Equation (1)). For the three-precursor model, the lumped escape rate $f(\varepsilon',k,\sigma)$ approaches zero in the limit of large splicing rates, which makes the above ratio equal to one (Equation (2)). This leads to the downward trend of the M/P (wt)/M/P (NoSurv) ratio as the splicing efficiency increases (Figure 3B).

These predictions can be examined with constructs containing introns with different splicing efficiencies. We studied a series of introns-in-*GAL1* constructs by varying splicing efficiencies. Two introns from the yeast *Kluyveromyces lactis* were also included, which are highly similar to those of *S. cerevisiae* (24), in order to assess evolutionary conservation. All these constructs encoded precursor mRNAs with PMSs in the intron.

The decay rate of the mature *GAL1* RNA was not affected appreciably by the intron of highest (*NMD2*) or lowest (*RPS11B*) splicing efficiency (Figure 3C). This permits the consistent comparison of various intron-in-*GAL1* constructs.

Then we compared the predictions with the experimental results. We obtained a series of constructs spanning M/P ratios of three orders of magnitude. The M/P ratios were measured in wt cells and in cells where the evolutionary conserved *UPF1* gene, which mediates NMD, has been deleted. Deletion of *UPF1* does not have significant adverse effects in yeast in laboratory growth conditions (4,25). The M/P(wt):M/P(upf1Δ) ratio was typically higher for introns with inefficient splicing (low M/P ratio). The three precursor-model captured clearly this downward trend as the splicing rate increased (Figure 3D, gray lines), unlike the two-precursor model (Figure 3D, orange lines). Scatter of data around the prediction may indicate variations in other parameters.

While even the single precursor model captures the behavior of efficiently spliced introns, the model with NMD and separated stages of co- and post-transcriptional splicing captures the complete trend including introns with weak splicing efficiency.

Estimation of escape rate

We proceeded with the three-precursor model to estimate the rate of pre-mRNA escape from splicing. The M/P ratio derived from the three-precursor model includes the lumped escape rate: $f(\varepsilon',k,\sigma)$

$$\text{M/P} = \frac{\text{M}}{P_N + P_D + P_R} = \frac{\sigma}{\mu_M}\frac{\mu_{PR}}{\mu_{PR} + \frac{k\varepsilon'}{\sigma + \varepsilon' + k}} = \frac{\sigma}{\mu_M}\frac{\mu_{PR}}{\mu_{PR} + f(\varepsilon',k,\sigma)} \qquad (3)$$

Thus, in addition to the M/P ratio all other parameters in the equation have to be measured to calculate the

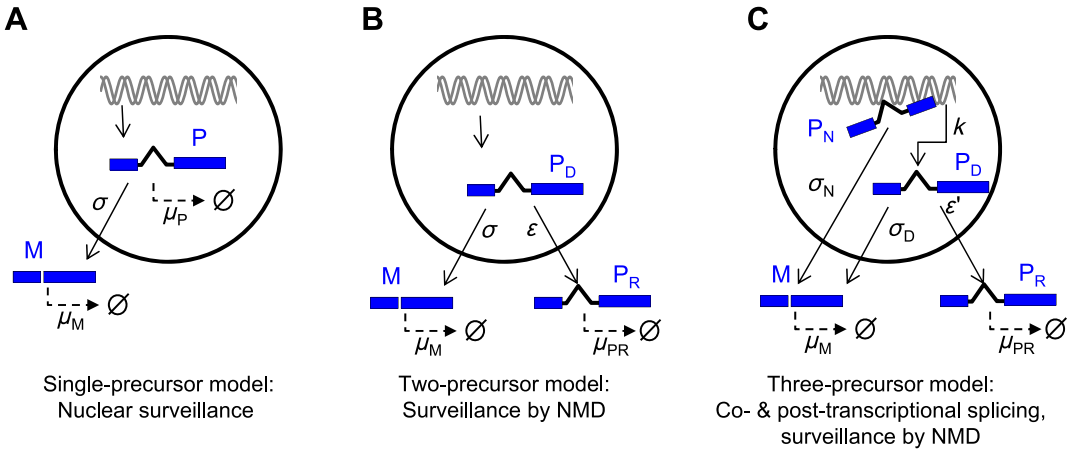

Figure 2. Models of splicing and RNA surveillance mechanisms. (**A**) In the single-precursor model, a nuclear surveillance mechanism degrades the unspliced RNA (μ_P). (**B**) In the two-precursor model, P_D is the precursor that has dissociated from the DNA and can be either spliced (σ) or exported without being spliced (ε), resulting in a precursor with retained intron, P_R. P_R is subject to fast degradation by NMD in the cytoplasm (μ_{PR}). (**C**) The nascent precursor, P_N, can be spliced co-transcriptionally or is converted to P_D upon transcriptional termination (where k is a lumped parameter for elongation and termination). The rest of the reactions are identical to that in the two-precursor model, except that the escape rate of the three-precursor model (ε') will have different fitted values than ε.

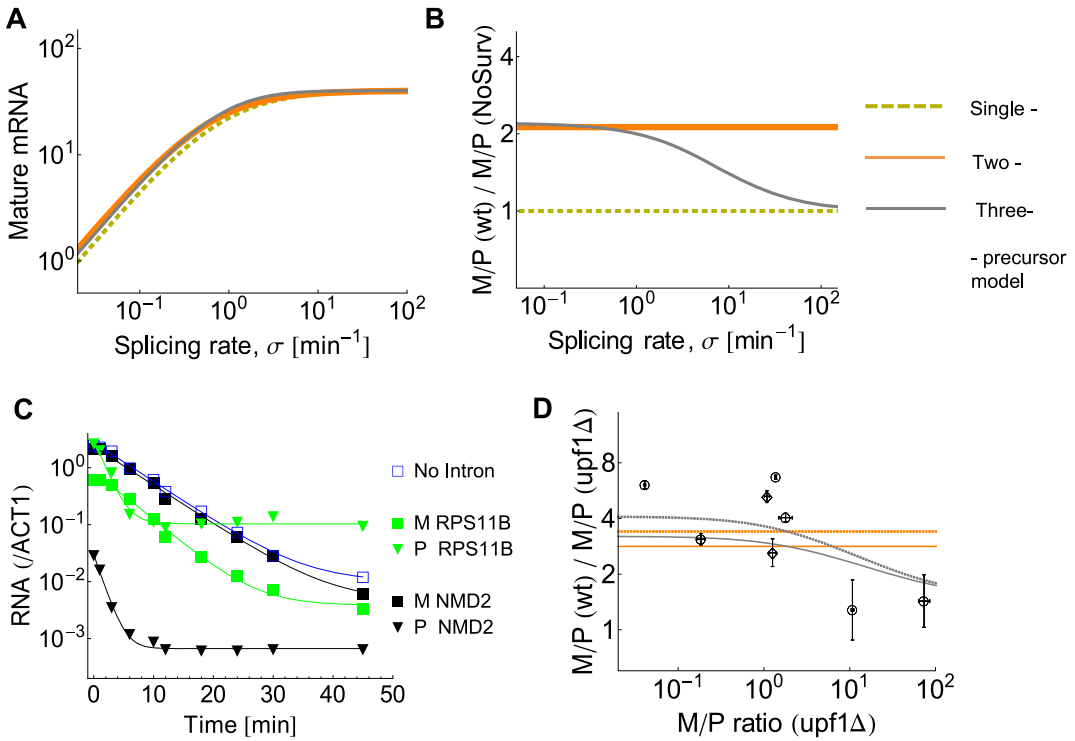

Figure 3. Distinction of models based on their predictions. (**A-B**) The response of mature mRNA in function of the splicing rate fails to distinguish the models, whereas the M/P(wt)/M/P(NoSurv) ratio allows clear cut between models. The mature mRNA (A) and the M/P(wt)/M/P(NoSurv) ratio (B) was calculated as the splicing rate was varied. In addition to p = 10 molecule * min^{-1}, the following rates (min^{-1}) were kept constant for the calculations: $\mu_M = 0.25$, $\varepsilon = 0.6$, $\varepsilon' = 2$ and $k = 1$. $\mu_P = 0.8$ min^{-1}; for the three-precursor model: $\sigma_N = \sigma_D = \sigma$. The precursor decay and escape rates were given values so that decline of the mature mRNA occurs approximately at the same M/P ratios. (**C**) Insertion of exogenous introns into *GAL1* has no impact on the mature decay kinetics. Decay of the *GAL1* RNAs without intron and the *NMD2* and *RPS11B* introns. The following half-lives were fitted for the precursor and mature mRNAs: *NMD2* (0.97 and 4.4 min), *RPS11B* (1.15 and 3.4 min), no intron (4.4 min). (**D**) Experimental verification of model prediction. The introns of the *S. cerevisiae RPS11B*, *PTC7, HNT2, RPS19A, NSP1, NMD2* (empty circles, in the order of increasing M/P ratios) and *Kluyveromyces lactis KlRPS11*, *KlHNT2* (empty diamonds) genes were inserted into *GAL1*. The *PTC7* intron in this out-of-frame construct has a lower M/P in comparison with the in-frame construct. Error bars represent standard deviations ($n = 3$; using RNA samples isolated from cells grown on different days). The M/P ratios of the same constructs were measured in wt and upf1Δ cells. For the solution of the two– and three-precursor models (orange and gray lines), $\varepsilon = 0.4$ and $\varepsilon' = 1.48$ were used; $\mu_{P(R)} = \mu_M$ when surveillance is absent. To account for the variations of precursor degradation rates in wt cells, predictions with $\mu_P = 0.9$ (full lines) and 2 min^{-1} (dotted lines) are shown. upf1Δ has a non-specific effect on *PTC7* expression, M/P(wt)/M/P(upf1Δ) = 1.53. Therefore, the predictions were multiplied with this value .

lumped escape rate. The decay rate of the mature and precursor RNAs can be accurately measured by promoter replacement constructs (13). If two unknowns (splicing and escape rate) remain in the equation, two measurements have to be performed so that the two unknowns can be calculated from two equations (two sets of Equation (3)). Therefore, we opted to measure M/P ratios of two constructs containing two forms of an intron differing only in the precursor decay rate (Figure 4A). For this purpose, the *PTC7* intron is ideal since it does not contain PMSs. To obtain a counterpart with faster precursor decay, mutations resulting in PMS were generated with precursor half-life that enabled estimation of the escape rate (see Materials and Methods). The *PTC7* intron was inserted in frame in the *GAL1* gene. The half-lives of the precursor and mature mRNAs were equal (Figure 4B). On the other hand, introduction of a PMS in the *PTC7* intron reduced the half-life of the precursor around four times without affecting the half-life of the mature mRNA (Figure 4B). Concomitantly, the mean M/P ratio doubled to 2.8. Thus, upon obtaining the two M/P ratios and two half-lives, the two equations with two unknowns yield a root of $f(\varepsilon',k,\sigma) = 0.40 \pm 0.11$ min^{-1} [mean ± standard deviation (SD)].

The escape rate may be influenced by both the *PTC7* intron and *GAL1* exon. In order to obtain specific data for the endogenous *PTC7* mRNA, the same procedure (Figure 4A) was repeated with the *PTC7* gene with the replaced promoter (Figure 4C), yielding a lumped escape rate of 0.75 ± 0.44 min^{-1}. Thus, the full *PTC7* mRNA and its intron-in-*GAL1* counterpart have similar escape rates.

The advantage of using steady-state M/P ratios as inputs to the equations is that the ratio reflects splicing rates more directly than time-course measurements (Supplementary Figures S4 and S5, Supplementary Information: simulated time-course measurements). The solution of the model with the above value of the escape rate is in good agreement with the time-course measurement (Supplementary Figure S6, Supplementary data).

The lumped escape rate, $f(\varepsilon',k,\sigma)$, includes the transcriptional termination rate k, which determines the ratio of co-transcriptional and post-transcriptional forms of precursor RNA. Thus, k has to be measured to obtain the specific value for the specific escape rate of the precursor mRNA, ε'. We have estimated this parameter by performing single-molecule RNA FISH (smFISH). In addition to measuring the total RNA count, we quantified the intensity of the active site of transcription, which reflects the production and termination rate of the transcript (26,27) (Figure 5A). The effective elongation and termination rate was estimated to be $k = 0.99 \pm 0.39$ min^{-1}. The smFISH result also confirmed that the insertion of intron into *GAL1* has no major impact on the production rate and the mean value decreased slightly with a concomitant increase in noise (Figure 5B). This is consistent with the observation that splicing operates at minimal noise level, especially at high expression levels (28,29).

Having obtained the value for k, the specific escape rate can be expressed: $\varepsilon' = 1.48$ min^{-1}. This value is in good agreement with mRNA export rates (see Discussion).

Estimated escape rate predicts the splicing rate below which mature mRNA levels decline

While all models predict that the mature mRNA declines at lower splicing rates, it is unclear what the threshold value is for the splicing rate below which the mature mRNA levels start to decline. When the splicing rate of an intron drops below that of the estimated lumped escape rate ($f(\varepsilon',k,\sigma) = 0.40$ min^{-1}), (Figure 4D, blue dashed line), the level of mature RNA is expected to decline in a linear fashion. Therefore, we examined the mature mRNA levels of the above intron-in-*GAL1* constructs. The mature mRNA levels started to decline when the splicing rate dropped below the threshold value corresponding to the previous estimate of the escape rate (Figure 4D).

This confirms the utility of using M/P ratios and decay rates to predict escape rate, especially in the light of the fact that the escape rate was predicted in a set of experiments in which the mature mRNA level was not varied (Figure 4A), and yet it correctly predicted the threshold splicing rate below which the mature mRNA declines linearly (Figure 4D). Since these results rely on various introns, we can conclude that the escape rate is conserved for pre-mRNAs with different introns, or at least that the escape rate variations of these constructs are restricted to a relatively narrow range.

Synergy in the splicing of multi-intronic mRNAs

When an RNA molecule contains multiple introns, each intron can be spliced at a rate which is either independent of or influenced by the other introns, for example by cooperative interaction.

The simplest model assumes independent splicing of each intron (Figure 6A, left panel). First, we again examined the mature mRNA levels as splicing rate is decreased, in the context of the three-precursor model. When an mRNA contains a weak intron and the splicing rate of the second intron is varied (Figure 6B, black dashed line), the decline of mature mRNA has characteristics similar to that of an mRNA with a single intron (Figure 6B, green line). However, when the splicing rate of both introns is varied, the mature mRNA declines more precipitously, imitating a cooperative process. Expressed more quantitatively, the logarithmic derivative of the mature mRNA response approaches two as the splicing rate is reduced (Figure 6B, black full line). A logarithmic derivative with a value higher than one can be a sign of cooperative processes (30–33). However, several processes are known to imitate cooperative processes, such as signaling cascades (30–33). In the case of splicing, the high logarithmic derivative of the response is due to the multiplicative influence of the escape rate and so it only imitates a cooperative process—given no cooperative interaction is defined in the model. Thus, the mature mRNA response alone is not sufficient for the identification of informative relations, similar to the earlier model selection (Figures 2 and 3). Therefore, we sought to define new relations that incorporate precursor levels in order to detect synergy in splicing.

We focused on three simple models of splicing of two closely spaced introns to identify distinctive effects of cooperative splicing (Figure 6A). In the independent (non-cooperative) splicing model, the splicing rate of an intron

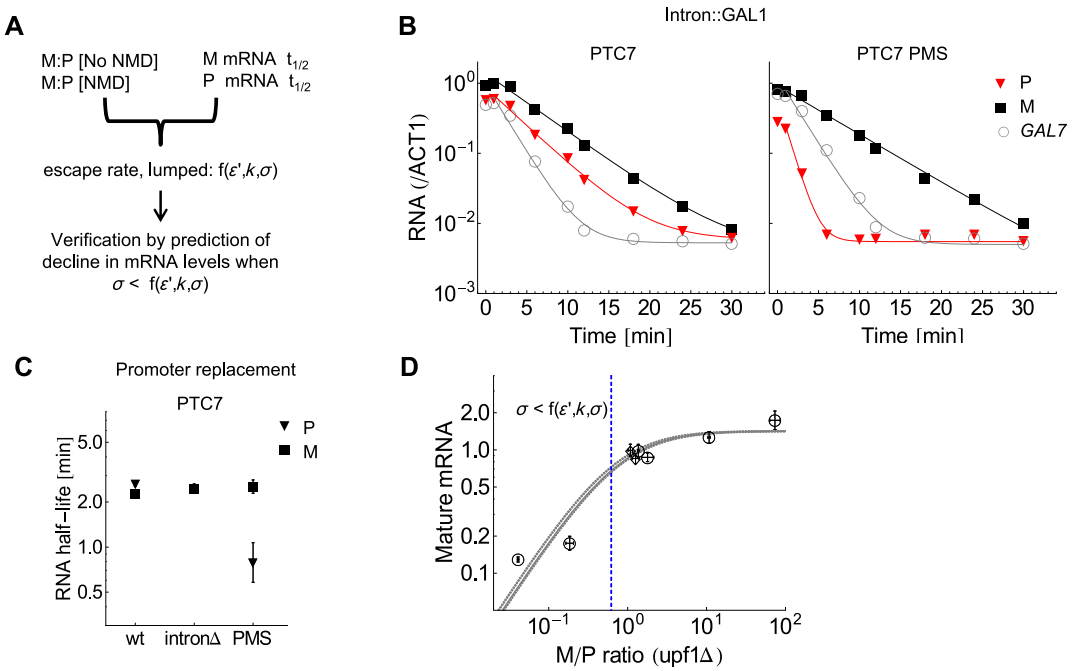

Figure 4. Estimation of escape rate from splicing and prediction of how it affects mature mRNA levels. (**A**) Flow chart for the determination and verification of the escape rate. (**B**) Measurement of half-life of precursor and mature mRNA by non-linear regresssion. The intron-in-*GAL1* constructs contain the wt *PTC7* intron (left panel) and the *PTC7* intron with PMS (right panel). The decay rates of the WT and PMS precursors are 0.24 and 0.9 min^{-1}, respectively. The respective rates for the mature mRNAs are 0.19 and 0.16 min^{-1} (raw mRNA data). (**C**) Decay rates of the *PTC7* precursor and mature mRNA measured with the promoter replacement construct. The introns are identical to that in the intron-in-*GAL1* construct. Error bars represent standard errors of parameter estimates. (**D**) Comparison of predictions with experiments. The model predicts that mature mRNA levels decline linearly when splicing rate is less than the lumped escape rate. This threshold value corresponds to an M/P ratio of 0.62 in upf1Δ cells (blue dashed line). For the solution of the three-precursor model, the splicing rate ($\sigma = \sigma_N = \sigma_D$) was varied, as expressed in Equations (S10) and (S11). The prediction is based on $k = 0.99$ min^{-1}, $\mu_M = 0.24$ min^{-1} and $\varepsilon' = 1.48 \pm 0.78$ [full (mean) and dotted gray lines (SD)]. The data were obtained from the same experiment as in Figure 3D. The mature RNA was normalized to *GAL2,* and the corresponding proportionality constant was used to adjust the mature RNA levels from the model.

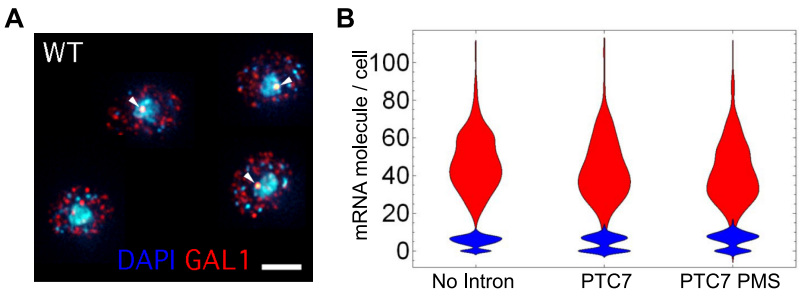

Figure 5. Single-molecule FISH images of the *GAL1* RNA. (**A**) DNA is stained by DAPI. Arrows mark the active sites of transcription (nascent RNA). The scale bar is equal to 2 μm. (**B**) Histograms of the nascent (blue) and all the other (red) RNA molecules. The following values for mean nascent (n) and single molecule (s) and coefficient of variation (CV) of the latter were measured. Control (no intron): $n = 5.0$, $s = 47$, CV = 0.30; *PTC7* intron: $n = 4.3$, $s = 43$, CV = 0.37; *PTC7* intron with PMS: $n = 5.8$, $s = 42$, CV = 0.39.

is independent of the presence of another intron. In the cooperative model, the splicing of an intron is enhanced by the presence of an adjacent intron, by a factor c. In the consecutive enhancement model, the splicing of an intron is enhanced only when the splicing of the adjacent intron had been completed.

In these models, the enhancement of splicing can be distinguished by comparing the precursor RNA containing two introns: P(2), and partially spliced RNA species. The splicing of P(2) yields partially spliced RNAs, P_1M_2 or M_1P_2, in which only one of the two introns has been spliced (Figure 6A). When the two introns are spliced at equal rates, all three models show that the levels of P_1M_2 and M_1P_2 are equal (Figure 6C, upper panel). In the independent splicing model, the levels of P(2) and partially spliced RNAs are approximately equal. On the other hand, P_1M_2 or M_1P_2 have higher levels than the precursor in the cooperative model. The opposite is seen in the consecutive enhancement model.

To express the above relations in quantitative terms, we introduced the interintronic splicing enhancement factor, ISEF, which is calculated by dividing the geometric mean

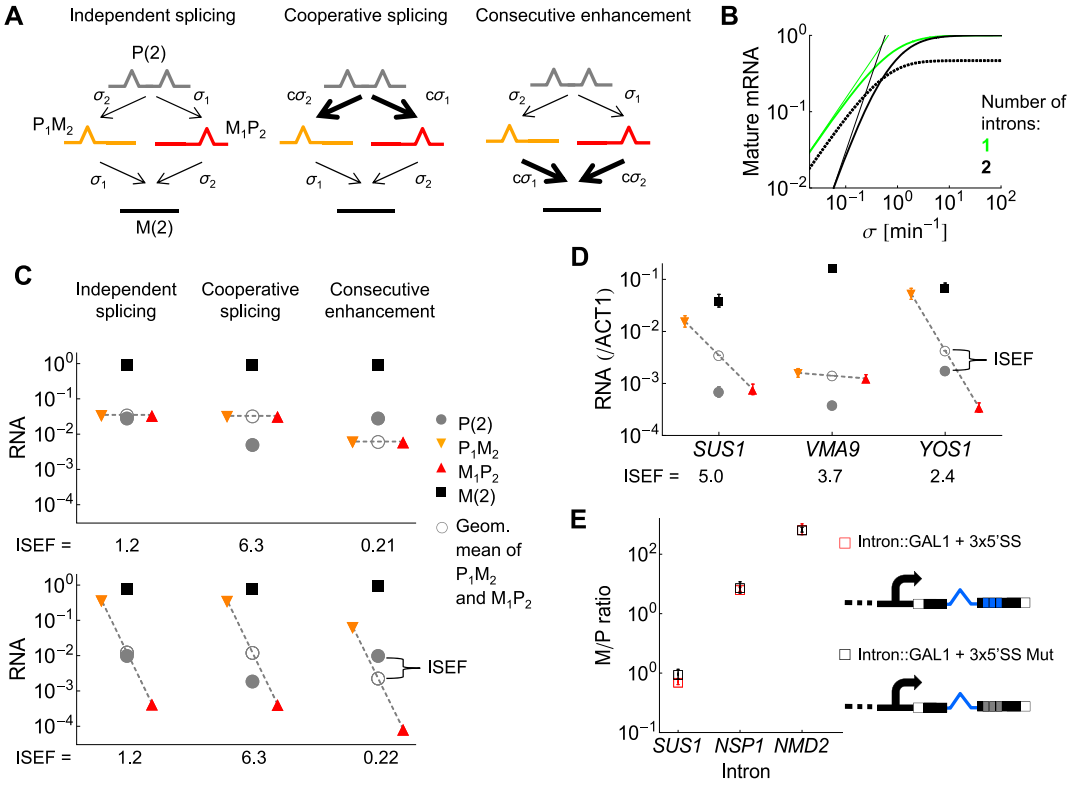

Figure 6. Models of splicing of mRNAs with two introns and experimental verification of models. The following RNA species are distinguished: P(2) and M(2) stand for the unspliced precursor and fully mature mRNA. The number in the parenthesis stands for the number of introns in the encoding gene. P_1M_2 and M_1P_2 represent splicing intermediates when either the 3′ or the 5′ intron is spliced out. (**A**) In the independent (non-cooperative splicing) model, the 1st and 2nd introns are spliced at rates σ_1 and σ_2, respectively. In the cooperative model, the splicing of each intron is enhanced by a factor of c ($c > 1$), only if the other intron is present. In the consecutive enhancement model, the splicing rate of each intron is enhanced by c when the other intron has been spliced already. (**B–C**) Solutions of the three-precursor model in a no surveillance setting. The co- and post-transcriptional splicing rates were equal and $c = 5$. The other parameters were fixed: p = 0.5 molecule * min^{-1}; $\mu_{PR} = \mu_M = 0.5$, $\varepsilon' = 2$ and $k = 1$ min^{-1}. (B) The green and black thick lines stand for mature mRNAs transcribed from genes with single and two introns. Splicing rate is varied in the independent splicing model. The thin lines represent polynomials that are asymptotic to the model solutions to indicate the logarithmic derivative, which is one for the single intron (green) and two for two-intron gene (black). For the black dashed lines, one of the two introns has a fixed splicing rate: $\sigma_1 = 0.5$ min^{-1}. (C) The two introns have equal ($\sigma_1 = \sigma_2 = 10$ min^{-1}; upper panel) and unequal ($\sigma_1 = 2$, $\sigma_2 = 50$ min^{-1}; lower panel) splicing rates. The geometric mean corresponds to the point that is positioned at equal 'distances' from both P_1M_2 and M_1P_2 levels (connected by dashed lines) when plotted logarithmically. The respective ISEF values are shown below each panel. (**D**) The RNA levels (mean and SD; $n = 3$) of the mRNAs expressed from the endogenous genes were measured in upf1Δ cells, grown in raffinose medium. (**E**) Addition of downstream 5′SSs fails to change the M/P ratio of the respective introns (mean and SD; $n = 4$). The *SUS1* intron represents the 1st (non-consensus) intron.

of the partially spliced precursors by P(2):

$$\mathrm{ISEF} = \frac{\sqrt{\mathrm{M_1P_2} \times \mathrm{P_1M_2}}}{\mathrm{P(2)}} \qquad (4)$$

In the non-cooperative model, ISEF has a value close to one since the above geometric mean is nearly identical to P(2). The small deviation from one is due to the escape of the partially spliced mRNAs (see Supplementary information). When splicing is cooperative, ISEF is approximately equal to the factor c (see Supplementary information, Equation (S12)). Conversely, ISEF is approximately equal to the reciprocal of c in the consecutive enhancement model and consequently less than one (given $c > 1$). Thus, the ISEF value is characteristic for each scenario of splicing of multi-intronic mRNA.

The utility of the ISEF becomes more evident when the splicing rates of the two introns are not equal and the two partially spliced mRNAs have different levels. Their geometric mean takes the same position relative to P(2) as in the case of introns with equal splicing rates (Figure 6C, lower panel). This can be easily explained by taking the example of the independent splicing model. When the splicing rate of the 1st intron is decreased and that of the second intron is increased by the same factor, the level of M_1P_2 drops below the level of P(2). The level P_1M_2 is also shifted by the same factor, but in the opposite direction. In this way, their geometric mean remains close to the P(2) level. Indeed, the respective ISEF values are approximately equal to the ISEF values calculated for the introns with equal splicing rates (Figure 6C).

Next, we quantified the ISEF values of endogenous genes with two introns. We were able to analyze three of the 10 genes that have two introns in the yeast genome: *SUS1*, *YOS1* and *VMA9*. All three genes have a length of around 0.5 kb and two short introns (<100 bp) interrupt the exons; this architecture enables the measurement of all RNA species by real-time PCR. The values of P_1M_2 and M_1P_2 were comparable for *VMA9*, while unequal levels of P_1M_2

and M_1P_2 indicate different splicing rates for the introns within *SUS1* and *YOS1*.

All ISEF values were clearly larger than one in upf1Δ cells: 5.0 for *SUS1*, 3.7 for *VMA9* and 2.4 for *YOS1* (Figure 6D), suggesting the presence of cooperative splicing of the two introns in all three mRNAs.

Next, we wanted to see whether the cooperative interaction requires the presence of a complete intron or complete splicing event. Each intron is defined by a 5′ and 3′ splice site (SS) and a branch point between them. To test if the binding of 5′SS splicing factors is sufficient to enhance the splicing rate, we placed three repeats of 5′ splice sites downstream of the weak *SUS1* intron inserted in *GAL1* mRNA. The splicing efficiency of these intron-[5′SS]$_3$-in-*GAL1* construct was compared to that of the intron-[5′mutSS]$_3$-in-*GAL1* constructs, in which the consensus splice site sequence was mutated (see Supplementary information). There was no significant difference in the respective M/P ratios. Similarly, when introns with higher splicing efficiency were analyzed (*NSP1, NMD2*), the M/P ratios were similar (Figure 6E). This indicates that the splicing of introns is not significantly affected by the downstream 5′ splice sites, neither by cooperative effect nor by depletion. The M/P ratios of the introns are comparable to their endogenous counterparts (Figures 1F and 6E and next paragraph for *SUS1*), revealing that the relative efficiency of introns is preserved in *GAL1* mRNA.

Reconstruction of cooperative splicing in *SUS1*

We further analyzed the *SUS*1 mRNA since it has the highest ISEF value. We applied the following strategy to reconstruct how the cooperative interaction affects the individual splicing events. In the first stage, we assessed whether the splicing rates are linearly dependent on the RNA concentrations and measured the half-lives of the mRNA species. In the second stage, we constructed two single intron derivatives of the *SUS1* gene to calculate the splicing rates of individual introns. Then using all these data we compared the predictions of models assuming either no synergy or different cooperative interactions.

First, we analyzed the *SUS1* gene with promoter replacement by varying the *GAL1* promoter activity with estradiol. The levels of mature mRNA to the precursor and partially spliced RNAs changed in a linear fashion with the promoter activity over a broad range of mRNA concentrations (Figure 7A).

The decay of all RNA species follows an exponential kinetics (Supplementary Figure S7). Prior studies indicate that the more 5′ the PMS positioned, the faster the mRNA decay by NMD (34,35). Thus, the mRNAs that contain the first intron—P(2) and P_1M_2—are expected to be more sensitive to NMD. Indeed, upon deletion of *UPF1*, their half-lives increase the most (Figure 7B). The half-lives of the precursor and mature RNAs are larger and confined into a narrower range in upf1Δ than in wt cells (Figure 7B). Thus, the splicing rates of the individual introns can be more faithfully estimated by M/P ratios in upf1Δ cells, and subsequent measurements were performed in upf1Δ cells, as well.

In the second stage, we constructed two genes with promoter replacement in which either the first or the second intron was deleted (Figure 7C, Supplementary Figure S8). In this way, splicing rate of each intron can be assessed independently of others. We obtained M/P ratios of 1 and 24 for the first and second introns, respectively, confirming that the 1st intron is inefficiently spliced and the 2nd intron is spliced efficiently (15). The relative values of the mRNA species expressed from the *SUS1* promoter replacement constructs are very similar to the ones expressed from the endogenous gene (Figures 6D and 7D).

Since the *SUS1* splicing can also be approximated by linear reaction rates (Figure 7A), we used linear models to make predictions, incorporating the splicing and decay rates obtained from the above *SUS1* constructs. For the escape rate, the estimate obtained in the previous section was used.

As already indicated, the ISEF value of 5 suggests that a cooperative model is required to explain splicing of *SUS1*. Indeed, the cooperative model with $c = 2$ predicts the level of precursor, P(2) correctly (Figure 7D, prediction with cooperative splicing compared to measured value of WT RNA) unlike the independent splicing model. The level of M_1P_2 is, however, clearly underestimated. This asymmetric behavior of one of the intermediates suggests a directional effect. Since the above model assumes that the two introns are very closely spaced, we included an additional step in the model to account for the time it takes for the second intron to be transcribed (see Supplementary information). During this time only the first intron can be spliced. Consequently, the levels of the splicing intermediates got predicted correctly (Figure 7D). Importantly, the synergy itself does not counteract intron retention by the P_1M_2, which results in having levels comparable to that of the mature mRNA. This step-wise reconstruction shows that the cooperative model can be refined upon the measurement of the intron deletion constructs.

While the synergy enhances the splicing of the first non-consensus intron from the precursor P(2), the splicing of the same intron is not enhanced from the P_1M_2 species. Consequently, the escape of this species is expected to result in a 1.28-fold reduction of the mature mRNA (Figure 7D). This implies that deletion of the inefficiently spliced intron would enhance the mature mRNA levels. We compared the mature level of wild-type *SUS1* and that of a *SUS1* in which the 1st intron is deleted. Indeed, we observed a 1.26 ± 0.30 and 1.38 ± 0.27-fold increase ($n = 4$) in mature RNA levels upon the removal of the 1st intron in upf1Δ and WT cells, respectively, in good agreement with the prediction.

DISCUSSION

Constancy of splicing efficiency and its promoter dependence

The constancy of splicing efficiency with respect to RNA concentration facilitates and supports model building (18,36). Conversely, non-linear splicing rates may play a role in creating complex regulatory responses. We found that splicing can be approximated by a first-order reaction over a broad range of expression levels for most of the examined constructs in our study, including *PTC7* and *SUS1*. Low splicing efficiency or mis-splicing is restricted to low expression levels. The decrease of the M/P ratio at very low expression levels can have at least two causes. At low expression level, the presence of non-coding, partial or over-

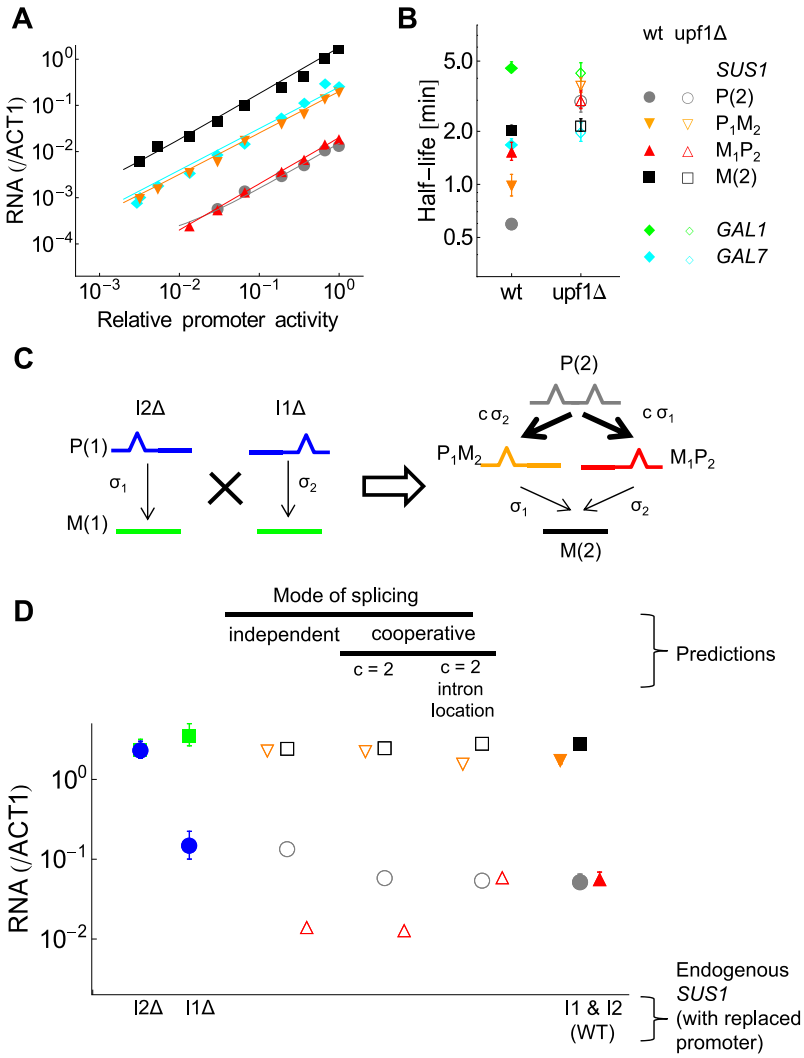

Figure 7. Reconstruction of synergy in the splicing steps of the *SUS1* mRNA. (**A**) The proportions of the *SUS1* mature and precursor mRNAs remain constant as the *GAL1* promoter activity is varied with estradiol in wt cells. (**B**) Half-life of the *SUS1* mRNAs in wt and upf1Δ cells. Error bars represent standard errors of parameter estimates. (**C**) The procedure for reconstruction of the splicing rates of mRNAs with two introns. First, the M/P ratios of mRNAs are measured in which either the 1st (I1Δ) or the 2nd (I2Δ) intron is deleted. P(1) and M(1) stand for the precursor and mature mRNAs expressed from the genes with a single intron. (**D**) The filled symbols stand for measurements in upf1Δ cells (mean and SD; $n = 4$) and the empty ones for predictions/reconstruction. The colors represent RNA species as indicated in (C). The following values were taken for the parameters from the previous measurements: p = 1.17 molecule * min^{-1}, $\varepsilon' = 1.48$ and $k = 0.99$ min^{-1}. The half-life of the mature mRNA is 2.17 min, and for the half-lives of the precursor and partially spliced mRNAs 3 min was taken since they have similar values (B). From the M/P ratios of the I1Δ and I2Δ constructs, the following splicing rates were obtained: $\sigma_1 = 0.95$ min^{-1} for the 1st intron and $\sigma_2 = 11.3$ min^{-1} for the 2nd intron. These values were fed in to the reconstruction, for which only the values for the enhancement factor, c were allowed to vary.

lapping transcripts can represent a significant portion of RNAs. For example, the *ACT1* intron can initiate a non-coding transcript into the 3′ direction at low expression levels (37). Another possible source of non-linearity would be inherent to the splicing mechanism itself. For example, a positive feedback has been suggested to occur during the second step of the splicing (18). Positive feedback may in principle result in a precipitous decline in splicing efficiency at low concentrations of RNA (e.g. (12)). Our experiments do not distinguish the two mechanisms.

Our targeted search for genes with non-linear splicing retrieved the *NSP1* and *NMD2* genes. This strategy may be generally applicable to identify genes with non-linear splicing. The intron of *NSP1* is not responsible for the non-linear behavior, but rather the chromosomal sequence-context of the *NSP1* gene. What could be the benefit of non-linear splicing? It is possible that the non-linear splicing itself contributes shaping the response generated by the environmental signal, such as ethanol (23). Since splicing and transcription can be regulated by shared factors (38), non-linearity in splicing may amplify transcriptional responses. The other retrieved gene, *NMD2* did not display non-linear M/P ratios, but it is possible that a non-linear response in splicing rates or precursor level may arise only in specific environmental conditions. Interestingly, a feedback regulation has been identified in the NMD pathway in mammalian cells (39). In the cells we constructed, a possible feedback was prevented by expressing an extra copy of *NMD2* with silent

mutations in order to focus on the non-linearity in the splicing rate.

The linearity of splicing rates is important also for reporting gene activities. Because of their fast decay rates, precursor mRNAs have been often used as a proxy to detect temporal variations in promoter activity (40–42). Our results suggest that these measurements are indeed generally suitable to report promoter activities, but linearity of splicing rate has to be confirmed for each intron containing reporter.

A model incorporating co- and post-transcriptional splicing and RNA surveillance by NMD reproduces trends in M/P ratios

From mathematical models of splicing and RNA surveillance, we derived simple algebraic expressions, which predict trends robustly, independent of specific parameter values. In particular, it was possible to answer the question whether splicing occurs in one stage or separated into post- and/or co-transcriptional stages solely based on M/P ratios measured in a series of genetic constructs (Figure 3B and D). The model including post- and co-transcriptional stage splicing and NMD was required to explain the trend in the M/P ratios across the entire range of splicing efficiencies (Figure 3D). This trend is in agreement with previous genome-wide study showing that the increase in precursor mRNA levels in upf1Δ cells is more pronounced for those introns that have a reduced splicing rate (4).

Furthermore, our results (Figure 3B and D) suggest that the studies addressing the question whether splicing of different mRNAs is more transcriptional or post-transcriptional should be performed also in upf1Δ cells, which can help in drawing quantitative conclusions.

Estimation of the escape rate

Knowing the escape rate is particularly important when both the unspliced and spliced forms of the mRNAs are translated and the respective proteins influence a cellular function, epitomized by *PTC7* and *SUS1*. In this case, the ratio of escape and splicing rates define the proportions of two protein isoforms. Introns in regulatory genes, especially those involved in controlling gene expression typically have weak introns whose retention plays an active role in the regulation of relevant cellular processes (15,16,43).

In addition to adjust the ratio of protein isoforms, the major biological function of the splicing rate is to control mature mRNA levels. Therefore, it is of central importance to estimate the escape rate. The escape rate we estimated was in accordance with the splicing rate below which mature mRNA levels declined for *GAL1* (Figure 4D) and also for *SUS1* (Figure 7D). Reducing splicing rates may act as means to regulate gene expression, by diverting pre-mRNAs to NMD instead of splicing them into mature mRNA. It has been observed that in several developmental programs, including meiosis and granulocyte differentiation, the splicing rate reduces globally or for a group of mRNAs by changing the concentration of splicing factors, with consequent changes in protein levels (38,44). In this way, escape of pre-mRNAs and the concomitant NMD serves as an inherent control mechanism and not just as surveillance mechanism to limit the accumulation of aberrant proteins. Similar global mechanism may affect the escape rate. Our approach (Figure 4A) is expected to facilitate the detection of such changes.

The precursor mRNA is a single entity in terms of sequence, but can participate in different reactions as it proceeds from being synthesized to being exported; thus, the precursor carries information on both splicing, evasion and NMD. Therefore, measuring decay rates and steady-state M/P ratios provides sufficient information to calculate the escape rate with the help of simple algebraic expressions.

The estimated escape rate ($\varepsilon' = 1.48\ \mathrm{min}^{-1}$) corresponds to an average period of 28 s for a pre-mRNA to be exported from the nucleus after the termination of transcription. Most of the time is required for the mRNA to encounter the nuclear pore complex (NPC) in the nucleus; the translocation across the NPC is fast (45–47). While no direct measurements have been available even with the most advanced microscopic technology, it has been estimated that the complete period of mRNA export relying on Brownian motion occurs in the range of seconds in lower eukaryotes, such as yeast, and in the order of minutes in higher eukaryotes (45). The escape rate calculated in our work is specifically the export rate of pre-mRNAs. Components have been identified that can lower the export rate of pre-mRNA to promote the nuclear retention as an effort to give a last chance for splicing (48,49). Even if we take this nuclear retention into account, assuming it reduces the export of pre-mRNAs two to five times (49), it would give a half-time for the export rate (without retention) of around 6–14 s, which agrees well with the suggested range of values (45).

The value of the escape rate also reveals the limitations of time-course experiments to infer splicing rates in yeast. Splicing rates are typically at least 10 times higher than decay rates, given most M/P ratios are higher than 10 (3) (see also Supplementary information, Supplementary Table S6). In this range of rates, the transient response that reflects splicing is small and is restricted to less than a minute (Supplementary Figure S4), which would be masked by the kinetics of the induction/de-induction process (Supplementary Figure S2). Moreover, if splicing is slow and its rate drops below the escape rate, most of the pre-mRNAs will escape with a retained intron, whose level will be largely determined by the NMD, masking the effect of variations in splicing rates (Supplementary Figure S5). Thus, it is the steady-state M/P ratio that can provide a faithful measure of splicing rate. These constraints highlight the benefit of our approach (Figure 4A), relying on steady-state M/P ratios and decay rates obtained from time-course measurements. It is not clear if the same constraints apply to mammalian cells, where splicing rates are typically slower. Furthermore, genes are considerably longer than in yeast and the longer periods of transcription allow splicing to be studied without interference from the precursor escape and subsequent NMD (50).

The escape rate is relevant for biotechnological engineering since gene expression can be regulated by introduction of aptamers in introns in mRNAs (51). In an optimal scenario, the aptamer in the intron would not change the mature mRNA level in the absence of ligand and the mature mRNA would linearly decline upon addition of increasing

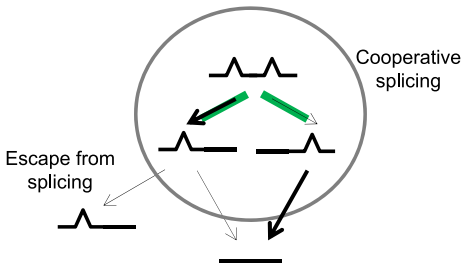

Figure 8. Integrated model of the *SUS1* mRNA splicing. The splicing efficiency of the first intron is low (thin arrow) and that of the second intron is high (thick arrow). Splicing is enhanced cooperatively only when both introns are present (green lines). The escape and splicing rates for the RNA species containing only the first non-consensus intron are similar.

concentration of the ligand. For this purpose, the splicing rate of an ideal intron should be just above the escape rate, which will minimally affect mature mRNA levels in the absence of ligand, but still ensures responsiveness to the ligand. The insertion of two introns with aptamers can successfully extend the range of mature mRNAs reduction (51), possibly by the multiplicative effect of escape rates (see below).

Synergy in splicing of multi-intronic mRNA

No measure has been available—to our knowledge—to quantify synergistic effects in splicing. The ISEF as defined in this study can be easily calculated from the precursor and partially spliced mRNAs. The larger than one ISEF values of all of the examined two-intron genes (*SUS1*, *VMA9*, *YOS1*) suggest they are subject to independent or cooperative splicing (Figure 6C and D). While the ISEF values provide a rapid assessment, the actual splicing rates have to be reconstructed from the set of intron deletion constructs. This approach revealed that both introns profit from a 2-fold enhancement of splicing rate, rather than only the second intron mediated by the polar effect, as hypothesized earlier based on measuring the decline of the mature mRNA levels (15). The mature RNA level in the construct with deletion of the first intron is higher simply because the escape of the partially spliced mRNA is reduced (Figures 7D and 8). This synergy does not depend on the RNA concentration since the precursor and different splicing intermediates change proportionally as the promoter activity is varied (Figure 7A).

It remains to be determined which factors mediate the synergy. Cooperative splicing may arise due to the interaction of identical or different components. An example for the former would be the interaction of splicing factors bound to the different introns. *In vitro* evidence suggests that such interaction can occur since native spliceosomes assemble with introns in the pre-mRNA to form supraspliceosomes (52). Our results indicate that more than the 5′splice site sequence is required to see synergistic effects. An example for the interaction of different protein is the family of hnRNP proteins: some combinations of hnRNP proteins exhibit significant synergy, whereas others act antagonistically during alternative splicing (53,54). In principle, cooperative effect can also arise if transcriptional elongation is coupled to splicing. In this case, the upstream intron has more time to complete splicing, which would result in asymmetric enhancement of splicing rates (8).

Interestingly, multiple introns can have a very strong effect on mature mRNA even without cooperative splicing (Figure 6B). Thus, simple intuitive reasoning may not suffice to disentangle the effects of escape rate of mRNAs with multiple introns from the actual cooperative interactions between the introns. Answering the question whether splicing of multiple introns is cooperative requires systematic model comparison.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGMENTS

We thank Mélusine Bleu, Valerie Crotet, Jakob Käppler, Charlotte Simonin and Stephanie Müller for the help with the experiments, and Chieh Hsu for the *GAL1* expression plasmid.

FUNDING

Systems X foundation [RTD grant, StoNets]; postdoctoral fellowship by the Carlsberg Foundation [to M.M.B.]. Funding for open access charge: Systems X [RTD grant, StoNets].
Conflict of interest statement. None declared.

REFERENCES

1. Nilsen,T.W. and Graveley,B.R. (2010) Expansion of the eukaryotic proteome by alternative splicing. *Nature*, **463**, 457–463.
2. Chin,K. and Pyle,A.M. (1995) Branch-point attack in group II introns is a highly reversible transesterification, providing a potential proofreading mechanism for 5′-splice site selection. *RNA*, **1**, 391–406.
3. Pleiss,J.A., Whitworth,G.B., Bergkessel,M. and Guthrie,C. (2007) Transcript specificity in yeast pre-mRNA splicing revealed by mutations in core spliceosomal components. *PLoS Biol.*, **5**, e90.
4. Sayani,S., Janis,M., Lee,C.Y., Toesca,I. and Chanfreau,G.F. (2008) Widespread impact of nonsense-mediated mRNA decay on the yeast intronome. *Mol. Cell*, **31**, 360–370.
5. Carrillo Oesterreich,F., Preibisch,S. and Neugebauer,K.M. (2010) Global analysis of nascent RNA reveals transcriptional pausing in terminal exons. *Mol. Cell*, **40**, 571–581.
6. Tardiff,D.F., Lacadie,S.A. and Rosbash,M. (2006) A genome-wide analysis indicates that yeast pre-mRNA splicing is predominantly posttranscriptional. *Mol. Cell*, **24**, 917–929.
7. Lopez,P.J. and Seraphin,B. (2000) Uncoupling yeast intron recognition from transcription with recursive splicing. *EMBO Rep.*, **1**, 334–339.
8. Alexander,R.D., Innocente,S.A., Barrass,J.D. and Beggs,J.D. (2010) Splicing-dependent RNA polymerase pausing in yeast. *Mol. Cell*, **40**, 582–593.
9. Girard,C., Will,C.L., Peng,J., Makarov,E.M., Kastner,B., Lemm,I., Urlaub,H., Hartmuth,K. and Luhrmann,R. (2012) Post-transcriptional spliceosomes are retained in nuclear speckles until splicing completion. *Nat. Commun.*, **3**, 994.
10. Cornish-Bowden,A. (2004) *Fundamentals of Enzyme Kinetics*. 3rd edn. Portland Press, London.
11. Pikielny,C.W. and Rosbash,M. (1985) mRNA splicing efficiency in yeast and the contribution of nonconserved sequences. *Cell*, **41**, 119–126.
12. Hsu,C., Scherrer,S., Buetti-Dinh,A., Ratna,P., Pizzolato,J., Jaquet,V. and Becskei,A. (2012) Stochastic signalling rewires the interaction map of a multiple feedback network during yeast evolution. *Nat. Commun.*, **3**, 682.

13. Passos,D.O. and Parker,R. (2008) Analysis of cytoplasmic mRNA decay in Saccharomyces cerevisiae. *Methods Enzymol.*, **448**, 409–427.
14. Juneau,K., Nislow,C. and Davis,R.W. (2009) Alternative splicing of PTC7 in Saccharomyces cerevisiae determines protein localization. *Genetics*, **183**, 185–194.
15. Hossain,M.A., Rodriguez,C.M. and Johnson,T.L. (2011) Key features of the two-intron Saccharomyces cerevisiae gene SUS1 contribute to its alternative splicing. *Nucleic Acids Res.*, **39**, 8612–8627.
16. Cuenca-Bono,B., Garcia-Molinero,V., Pascual-Garcia,P., Dopazo,H., Llopis,A., Vilardell,J. and Rodriguez-Navarro,S. (2011) SUS1 introns are required for efficient mRNA nuclear export in yeast. *Nucleic Acids Res.*, **39**, 8599–8611.
17. Fededa,J.P., Petrillo,E., Gelfand,M.S., Neverov,A.D., Kadener,S., Nogues,G., Pelisch,F., Baralle,F.E., Muro,A.F. and Kornblihtt,A.R. (2005) A polar mechanism coordinates different regions of alternative splicing within a single gene. *Mol. Cell*, **19**, 393–404.
18. Aitken,S., Alexander,R.D. and Beggs,J.D. (2011) Modelling reveals kinetic advantages of co-transcriptional splicing. *PLoS Comput. Biol.*, **7**, e1002215.
19. Raj,A., van den Bogaard,P., Rifkin,S.A., van Oudenaarden,A. and Tyagi,S. (2008) Imaging individual mRNA molecules using multiple singly labeled probes. *Nat. Methods*, **5**, 877–879.
20. Trcek,T., Chao,J.A., Larson,D.R., Park,H.Y., Zenklusen,D., Shenoy,S.M. and Singer,R.H. (2012) Single-mRNA counting using fluorescent in situ hybridization in budding yeast. *Nat. Protoc.*, **7**, 408–419.
21. Mueller,F., Senecal,A., Tantale,K., Marie-Nelly,H., Ly,N., Collin,O., Basyuk,E., Bertrand,E., Darzacq,X. and Zimmer,C. (2013) FISH-quant: automatic counting of transcripts in 3D FISH images. *Nat. Methods*, **10**, 277–278.
22. Kelemen,J.Z., Ratna,P., Scherrer,S. and Becskei,A. (2010) Spatial epigenetic control of mono- and bistable gene expression. *PLoS Biol.*, **8**, e1000332.
23. Pleiss,J.A., Whitworth,G.B., Bergkessel,M. and Guthrie,C. (2007) Rapid, transcript-specific changes in splicing in response to environmental stress. *Mol. Cell*, **27**, 928–937.
24. Neuveglise,C., Marck,C. and Gaillardin,C. (2011) The intronome of budding yeasts. *C R Biol.*, **334**, 662–670.
25. Gatfield,D., Unterholzner,L., Ciccarelli,F.D., Bork,P. and Izaurralde,E. (2003) Nonsense-mediated mRNA decay in Drosophila: at the intersection of the yeast and mammalian pathways. *EMBO J.*, **22**, 3960–3970.
26. Raj,A., Peskin,C.S., Tranchina,D., Vargas,D.Y. and Tyagi,S. (2006) Stochastic mRNA synthesis in mammalian cells. *PLoS Biol.*, **4**, e309.
27. Zenklusen,D., Larson,D.R. and Singer,R.H. (2008) Single-RNA counting reveals alternative modes of gene expression in yeast. *Nat. Struct. Mol. Biol.*, **15**, 1263–1271.
28. Waks,Z., Klein,A.M. and Silver,P.A. (2011) Cell-to-cell variability of alternative RNA splicing. *Mol. Syst. Biol.*, **7**, 506.
29. Melamud,E. and Moult,J. (2009) Stochastic noise in splicing machinery. *Nucleic Acids Res.*, **37**, 4873–4886.
30. Shah,N.A. and Sarkar,C.A. (2011) Robust network topologies for generating switch-like cellular responses. *PLoS Comput. Biol.*, **7**, e1002085.
31. Savageau,M.A. (2011) Biomedical engineering strategies in system design space. *Ann. Biomed. Eng.*, **39**, 1278–1295.
32. Zhang,Q., Bhattacharya,S. and Andersen,M.E. (2013) Ultrasensitive response motifs: basic amplifiers in molecular signalling networks. *Open Biol.*, **3**, 130031.
33. Kholodenko,B.N., Hoek,J.B., Westerhoff,H.V. and Brown,G.C. (1997) Quantification of information transfer via cellular signal transduction pathways. *FEBS Lett.*, **414**, 430–434.
34. Hagan,K.W., Ruiz-Echevarria,M.J., Quan,Y. and Peltz,S.W. (1995) Characterization of cis-acting sequences and decay intermediates involved in nonsense-mediated mRNA turnover. *Mol. Cell. Biol.*, **15**, 809–823.
35. Decourty,L., Doyen,A., Malabat,C., Frachon,E., Rispal,D., Seraphin,B., Feuerbach,F., Jacquier,A. and Saveanu,C. (2014) Long open reading frame transcripts escape nonsense-mediated mRNA decay in yeast. *Cell Rep.*, **6**, 593–598.
36. Schmidt,U., Basyuk,E., Robert,M.C., Yoshida,M., Villemin,J.P., Auboeuf,D., Aitken,S. and Bertrand,E. (2011) Real-time imaging of cotranscriptional splicing reveals a kinetic model that reduces noise: implications for alternative splicing regulation. *J. Cell Biol.*, **193**, 819–829.
37. Irniger,S., Egli,C.M., Kuenzler,M. and Braus,G.H. (1992) The yeast actin intron contains a cryptic promoter that can be switched on by preventing transcriptional interference. *Nucleic Acids Res.*, **20**, 4733–4739.
38. Munding,E.M., Igel,A.H., Shiue,L., Dorighi,K.M., Trevino,L.R. and Ares,M. Jr (2010) Integration of a splicing regulatory network within the meiotic gene expression program of Saccharomyces cerevisiae. *Genes Dev.*, **24**, 2693–2704.
39. Huang,L., Lou,C.H., Chan,W., Shum,E.Y., Shao,A., Stone,E., Karam,R., Song,H.W. and Wilkinson,M.F. (2011) RNA homeostasis governed by cell type-specific and branched feedback loops acting on NMD. *Mol. Cell*, **43**, 950–961.
40. Zeisel,A., Kostler,W.J., Molotski,N., Tsai,J.M., Krauthgamer,R., Jacob-Hirsch,J., Rechavi,G., Soen,Y., Jung,S., Yarden,Y. *et al.* (2011) Coupled pre-mRNA and mRNA dynamics unveil operational strategies underlying transcriptional responses to stimuli. *Mol. Syst. Biol.*, **7**, 529.
41. Koike,N., Yoo,S.H., Huang,H.C., Kumar,V., Lee,C., Kim,T.K. and Takahashi,J.S. (2012) Transcriptional architecture and chromatin landscape of the core circadian clock in mammals. *Science*, **338**, 349–354.
42. Stavreva,D.A., Wiench,M., John,S., Conway-Campbell,B.L., McKenna,M.A., Pooley,J.R., Johnson,T.A., Voss,T.C., Lightman,S.L. and Hager,G.L. (2009) Ultradian hormone stimulation induces glucocorticoid receptor-mediated pulses of gene transcription. *Nat. Cell Biol.*, **11**, 1093–1102.
43. Preker,P.J. and Guthrie,C. (2006) Autoregulation of the mRNA export factor Yra1p requires inefficient splicing of its pre-mRNA. *RNA*, **12**, 994–1006.
44. Wong,J.J., Ritchie,W., Ebner,O.A., Selbach,M., Wong,J.W., Huang,Y., Gao,D., Pinello,N., Gonzalez,M., Baidya,K. *et al.* (2013) Orchestrated intron retention regulates normal granulocyte differentiation. *Cell*, **154**, 583–595.
45. Oeffinger,M. and Zenklusen,D. (2012) To the pore and through the pore: a story of mRNA export kinetics. *Biochim. Biophys. Acta*, **1819**, 494–506.
46. Becskei,A. and Mattaj,I.W. (2005) Quantitative models of nuclear transport. *Curr. Opin. Cell Biol.*, **17**, 27–34.
47. Smith,A.E., Slepchenko,B.M., Schaff,J.C., Loew,L.M. and Macara,I.G. (2002) Systems analysis of Ran transport. *Science*, **295**, 488–491.
48. Rutz,B. and Seraphin,B. (2000) A dual role for BBP/ScSF1 in nuclear pre-mRNA retention and splicing. *EMBO J.*, **19**, 1873–1886.
49. Galy,V., Gadal,O., Fromont-Racine,M., Romano,A., Jacquier,A. and Nehrbass,U. (2004) Nuclear retention of unspliced mRNAs in yeast is mediated by perinuclear Mlp1. *Cell*, **116**, 63–73.
50. Singh,J. and Padgett,R.A. (2009) Rates of in situ transcription and splicing in large human genes. *Nat. Struct. Mol. Biol.*, **16**, 1128–1133.
51. Weigand,J.E. and Suess,B. (2007) Tetracycline aptamer-controlled regulation of pre-mRNA splicing in yeast. *Nucleic Acids Res.*, **35**, 4179–4185.
52. Azubel,M., Habib,N., Sperling,R. and Sperling,J. (2006) Native spliceosomes assemble with pre-mRNA to form supraspliceosomes. *J. Mol. Biol.*, **356**, 955–966.
53. Huelga,S.C., Vu,A.Q., Arnold,J.D., Liang,T.Y., Liu,P.P., Yan,B.Y., Donohue,J.P., Shiue,L., Hoon,S., Brenner,S. *et al.* (2012) Integrative genome-wide analysis reveals cooperative regulation of alternative splicing by hnRNP proteins. *Cell Rep.*, **1**, 167–178.
54. Chen,M., David,C.J. and Manley,J.L. (2012) Concentration-dependent control of pyruvate kinase M mutually exclusive splicing by hnRNP proteins. *Nat. Struct. Mol. Biol.*, **19**, 346–354.