# SUPPLEMENTAL INFORMATION

## EXTENDED DISCUSSION

For labeling introns in pre-mRNA molecules, we applied the MS2 detection technique (Beach et al., 1999; Bertrand et al., 1998), which has been used extensively to visualize mRNA molecules in living cells (reviewed in (Tyagi, 2009)). A problem with this technique is that insertion of foreign RNA domains for artificial protein tethering may interfere with the dynamics of gene expression and RNA processing. To assess the impact of the MS2 system on beta-globin pre-mRNA, we used a distinct labeling approach based on the N-peptide of bacteriophage λ. Compared to MS2 coat protein, which is composed of 129 amino acids (MW = 13 728), the λN peptide consists of 22 amino acids and therefore potentially interferes less with the folding and function of the tagged RNA (Baron-Benhamou et al., 2004; Daigle and Ellenberg, 2007). We engineered isogenic cell lines containing MS2 and λN cassettes of similar sizes integrated at the same site in each intron. The consistent observation of kinetics that are essentially the same, whether an intron is labeled with MS2 or with λN argues against a bias introduced by the fluorescent tagging procedure.

A previous study using tetracycline-regulated promoters and RNase protection assays inferred half-lives of approximately 1 minute for the splicing of beta-globin introns (Audibert et al., 2002). Splicing rates of under 2 minutes were also found for splicing of adenovirus E2A pre-mRNA by pulse-labeling (Gattoni et al., 1986). In contrast, Schmidt and colleagues used fluorescence recovery after photobleaching (FRAP) to analyze the kinetic properties of an ensemble population of pre-mRNAs synthesized from a gene cluster containing ~20 copies of a reporter artificial transgene, and determined a half-life of 105-165 s for the MINX intron, which is a small adenovirus-

derived intron with strong splicing signals (Schmidt et al., 2011). Like MINX, the beta-globin introns are short, constitutively spliced and contain strong splicing sites. Yet, we observed much faster splicing rates of beta-globin introns. The reason for this discrepancy is likely methodological: the rapid fluorescence fluctuations that we observed for a single pre-mRNA molecule (or a burst of 2-3 molecules) are probably hidden in the bulk measurement of fluorescence recovery from a multitude of introns at different stages of their life cycle.

Our finding that not all RNA stem-loops bind GFP fusion proteins in living cells is in agreement with data from previous studies (Fusco et al., 2003; Grunwald and Singer, 2010). The loading state of GFP fusion proteins onto MS2 stem-loops may depend on the concentration of these proteins available in the cell (Wu et al., 2012), suggesting that accessibility of all stem-loops for binding to GFP fusion proteins may be limited in vivo.

In agreement with the single integration site of the *HBB* gene, a unique fluorescent nuclear dot was visible in the majority of cells. However, approximately 15% of the cells had two dots that were either very closely apposed (**Figure S1B**) or further apart from each other (**Figure S1D**). We measured the distance between pairs of dots in the same nucleus and found that it remained approximately constant over time (**Figure S1C, E**). Paired dots probably result from replication of the *HBB* gene locus, with closely apposed dots corresponding to recently replicated sister loci. Less frequently, more than two dots per nucleus were observed. Aneuploidy, which is common in human cell lines, may additionally explain the appearance of nuclei with two or more beta-globin genomic loci.

## EXTENDED EXPERIMENTAL PROCEDURES

### Plasmids

The *HBB* genomic clone containing a deletion of 593 bp in intron 2 between *Rsa* I and *Ssp* I sites was previously described (Antoniou et al., 1998). A blunted *Nco*I-*Acc65*I fragment containing the *HBB* gene from initiation codon to approximately 1800 bp past the poly(A) site was inserted into the *Kpn*I (blunted) site of pcDNA5/FRT/TO (Invitrogen). The sequence corresponding to 6 MS2 stem-loops was excised from Pcβwtβ2-6MS2 (Lykke-Andersen et al., 2000) by PCR using the *Bgl*II5MS2For and *BamH*I5MS2Rev pair of primers, followed by *Bgl*II and *BamH*I digestion and insertion into the *Bgl*II and *BamH*I digested pCMV5 vector, generating pCMV5-6MS2. An array of 24 stem-loops was constructed by successive insertions of *Bgl*II - *BamH*I pCMV5-6MS2 fragments into the *Bgl*II site of pCMV5-6MS2. The sequence corresponding to 5 λN binding sites (BoxB) was excised from pβglobin.5BoxB (Gehring et al., 2003) by PCR using the *Bgl*II5BoxFor and *BamH*I5BoxRev pair of primers, followed by *Bgl*II and *BamH*I digestion and insertion into the *Bgl*II and *BamH*I sites of pCMV5, generating pCMV5-5BoxB. An array of 25 binding sites was then constructed by successive insertions. The arrays of MS2 and λN binding sites were excised by *Bgl*II and *BamH*I digestion, blunted and inserted into either the blunted *Bpi* I site of beta-globin intron 1, or the blunted *Psp5* II site of beta-globin intron 2. Plasmids expressing the MS2 coat protein fused to GFP or mCherry were a gift of E. Bertrand (Boireau et al., 2007; Fusco et al., 2003) and the λN-GFP plasmid was a gift of J. Ellenberg (Daigle and Ellenberg, 2007). The PTS1 peroxisomal targeting signal was introduced into Bsp1407I and XbaI sites of pECFP-N1 (Clontech) by PCR amplified complementary primers PTS1-Fw and PTS1-Rev encoding the peroxisomal targeting sequence type 1 from human acyl-CoA-oxidase obtained from http://www.peroxisomedb.org/ (Schluter

et al., 2010). The IgM and IgM-PY mini-genes were amplified from pμM (IgM M1-M2) and pPy-AdML-IgM derived constructs described in (Guth et al., 1999; Guth et al., 2001) using specific primers IgM-Fw and IgM-Rev or IgM-PY-Rev respectively, to introduce Acc65I and BamHI sites and delete a stop codon in exon M2. After Acc65I and BamHI digest the resulting fragments were ligated into the same sites in pECFP-PTS1 generating the vectors pIgM-CFP-PTS1 and pIgM-PY-CFP-PTS. Intron extensions were generated by PCR amplification of fragments from the first intron of mouse RNA Pol II gene using genomic DNA from a murine erythroleukemia (MEL) cell line. The 24 MS2 repeat sequence was ligated into the blunted intronic BbvCI site in pIgM-CFP-PTS1, pIgM-PY-CFP-PTS1, pIgM-600-PY-CFP-PTS1 and into the intronic SwaI site in pIgM-1.7k-PY-CFP-PTS1 respectively. From the pECFP-N1 vector backbone, the final constructs were cut out with HindIII and HpaI and ligated into HindIII and Eco32I sites in pcDNA5/FRT/TO.

**Stable cell line construction**

As parental cell line to receive and express the *HBB* gene and IgM mini-gene constructs, we used Flp-In™ T-REx™-293 cells (purchased from Invitrogen Life Technologies). The cells were grown as monolayer in Dulbecco´s Modified Eagle Medium (D-MEM) supplemented with 10% fetal bovine serum and 2mM L-Glutamine (all cell culture reagents were from Invitrogen). Isogenic stable cell lines were generated through Flp recombinase-mediated integration by cotransfecting a plasmid that expresses the Flp recombinase (pOG44, Invitrogen) and the pcDNA5/FRT vector containing the *HBB* and IgM mini-gene constructs. After transfection, the Flp-In™ T-REx™-293 cells were maintained under selective pressure in the presence of 200 μg/ml

hygromycin B (Roche) and 15 µg/ml blasticidin (Invitrogen). Transgene expression was induced with either 1µg/ml tetracycline or 6µg/ml doxycycline (Clontech) for 4-18 hours.

**Transient transfection**

Cells were grown to approximately 70% confluency and transfected for 4-18 hours with plasmids encoding MS2 and λN fluorescent fusion proteins. Plasmid DNA was transfected using Lipofectamine$^{TM}$ 2000 (Invitrogen) according to the manufacturer's protocol.

**RNA and DNA Fluorescence In Situ hybridization**

Cells grown on glass coverslips were fixed in 3.7% paraformaldehyde in PBS for 10 min and permeabilized with 0.5% Triton X-100/ 2mM VRC in PBS for 10 min. For RNA FISH, cells were incubated for 30 min in blocking solution (1% BSA/ 1 µg/ml tRNA/ 2mM VRC/ 2x SSC/ 0.02% Triton X-100) and then hybridized with probes for 18 hours at 37ºC. Post-hybridization washes were in 52,5% formamide/ 1x SSC/ 0.01% Triton X100, for 3 times, 15min each, at 37ºC, followed by 2 washes in 2x SSC/ 0.02% Triton X100, 5 min each, at 37ºC. For DNA FISH, cells were denatured in 50% formamide/ 1x SSC/ 0.01% Triton X100 for 30 min at 80ºC. As a control, cells were treated with 20 µg/ml RNase A for 30 min at 37ºC before hybridization with the DNA probe. Probes were produced by nick translation using Cy3- or Cy5-dUTP (GE Healthcare UK Limited). The MS2 probe was obtained by *Bgl*II / *BamH*I digestion of pCMV5-MS2, and the beta-globin probe by *Nco*I / *Acc65*I/ *Pvu*I digestion of a *HBB* cDNA clone (Martins et al., 2011). The DNA probe was generated by *Rsa*I digestion of

the pcDNA5/FRT/TO plasmid. Probes were diluted in 50% formamide/ 2x SSC/ 50mM phosphate buffer/ 10% dextran sulphate/ 0.7 mg/ml sonicated herring sperm DNA/ 0.3 mg/ml human Cot1 DNA (Invitrogen), pH 7.0, and denatured before hybridization for 5 min at 75ºC. Samples were mounted with Vectashield (Vector Laboratories) and imaged in an Andor Revolution XD spinning-disk confocal microscope (Andor Technology) equipped with a Nikon Eclipse Ti-E microscope, a Yokogawa CSU-X1 confocal scanning head, an Andor iXon 897 EMCCD camera and a Prior piezo for fast z displacement. We used a DPSS 561 nm laser in conjugation with a 579-631 nm BP filter for Cy3 detection and a DPSS 640 nm laser line in conjunction with a 665-705 nm BP filter for Cy5 detection. Images were acquired using a Plan Apo VC 60x/1.4 oil objective.

**RNA purification and RT-PCR**

RNA was extracted using TRIzol® (Invitrogen) and treated with RNase-free DNaseI (Roche). Production of cDNA was carried out using Transcriptor High Fidelity cDNA Synthesis Kit (Roche) using a primer complementary to a region past the poly(A) site, pARr for *HBB* and BGH-pA-Rev for IgM construct. All primers are presented in Supplementary Table 1. Polymerase chain reaction (PCR) products amplified by either NZYLong DNA polymerase (NZYTech) or Bio-X-Act Short Taq or BioTaq polymerase were separated by agarose gel electrophoresis, detected by GelRed (Biotium, Inc.), and imaged on the AlphaImager HP Imaging System (Alpha Innotech) and Typhoon 9210 (Molecular Dynamics/GE Healthcare). Semiquantitative analysis of PCR product abundance was carried out by integrating the signal intensities from each band of non-saturated gel images from the typhoon system using ImageJ software.

**Spinning-disk confocal live-cell imaging**

Cells were plated on 25 mm diameter glass coverslips coated with 0.01% Poly-L-Lysine. Before imaging, the medium was changed to α-MEM without phenol red (Invitrogen) supplemented with 20 mM HEPES, pH 7.4 and 10% FBS. Each coverslip was mounted into a perfusion chamber and placed in a heated sample holder (20/20 Technology, Inc.; Wilmington, NC) mounted on the stage of a Marianas$^{TM}$ imaging system (Intelligent Imaging Innovations, Denver, CO) based on an Axio Observer inverted microscope (Carl Zeiss, Inc.; Thornwood, NY) equipped with a spinning-disk confocal head (Yokogawa Electric, Tokyo, Japan). The microscope stage and objective lenses were maintained inside an environmental chamber (Okolab) set at 37ºC with 5% CO2 and 100% humidity. The axial position of the sample was controlled with a piezo-driven stage (Applied Scientific Instrumentation, Eugene, OR). Samples were illuminated using 100 mW solid state lasers (λ=488 nm for GFP and λ =561 nm for mCherry, Coherent, Inc.; Santa Clara, CA) coupled to an acoustic-optical tunable filter (AOTF). Images were acquired using either 63x or 100x (Plan-Apo, 1.4 NA) oil immersion objectives (Carl Zeiss, Inc.) under control of Slidebook 5.0 software (Intelligent Imaging Innovations, Denver, CO). Digital images (16-bit) were obtained using a cooled CCD camera (QuantEM, 512SC, Photometrics, Tucson, AZ) with acquisition times between 15 and 30 ms (for transcription sites in the cell nucleus), or 1 s (for isolated GFP molecules).

**Image Analysis**

The following sequential steps were performed in the analysis of each single transcription site in a time lapse sequence. (1) The XY position of the transcription site

in the first frame of the sequence was determined as a local maximum, given approximate initial coordinates provided by the user (by clicking on the location of the fluorescent dot in a maximum intensity projection image). A volume of interest (a XYZ subset of the original z-stack, typically $5 \times 5 \times Z_N$, where $Z_N$ is the number of planes in the z-stack) was extracted from the original z-stack with the transcription site at its center. The Z plane corresponding to the highest fluorescence intensity value at the XY coordinates of the transcription site was recorded as its Z position. (2) Unless otherwise specified by the user (by clicking on a different location) the XYZ coordinates of the transcription site for subsequent time points were automatically determined by re-centering the volume of interest at the site for each time point by searching for a local intensity maximum, provided the signal-to-noise ratio of fluorescence intensity in the volume of interest was higher than a given threshold (typically 6). This condition ensured that the search for a local maximum did not occur when no labeled introns were detected in the image, thus avoiding wayward tracking results that could move the volume of interest away from the transcription site. (3) The total fluorescence intensity of the transcription site was calculated for each time point by performing a 2D Gaussian fit at the Z plane corresponding to the highest intensity value, with a modified implementation of the Gaussian fitting function developed by David Kolin

(http://www.cellmigration-gateway.com/resource/imaging/icsmatlab/ICSTutorial.html) using the formula

$$I(x,y) = A + I_0 e^{-\frac{\left(\frac{x-x_0}{a}\right)^2 + \left(\frac{y-y_0}{b}\right)^2}{2}}$$

where $A$ is the background nuclear intensity, $I_0$ is the peak intensity, $x_0$ and $y_0$ are the coordinates of the peak, $a$ is the Gaussian width in $x$ and $b$ the Gaussian width in $y$. All

these parameters were estimated by the fitting function. The total fluorescence intensity of the transcription site was calculated as the integral of the Gaussian curve: $TFI = 2\pi abI_0$ (Rust et al., 2006). The TFI for transcription sites in which the difference between the maximum intensity value $I_0$ and the background $A$ was smaller than the standard deviation of the background $\sigma_A$ estimated at the edges of the volume of interest (i.e. $|I_0 - A| < \sigma_A$) was set to zero. Also, transcription sites for which the Z plane of highest fluorescence intensity corresponded either to the first or last planes of the Z-stack were discarded from the analysis.

To determine the average TFI for a single intron, we used the same software to analyze images of SSA treated cells acquired with an exposure time of 30 ms. Because in the presence of SSA unspliced pre-mRNA molecules are no longer immobilized at the transcription site but diffuse throughout the nucleoplasm, we performed quantification in single optical planes. We used MATLAB to plot histograms of TFI and perform Gaussian fitting on the observed distributions to determine the TFI values which are within one standard deviation (68%) and two standard deviations (95%) away from the mean. Since the EMCCD camera output is linearly dependent on the exposure time, the average TFI corresponding to a single intron for a different exposure time can be obtained by multiplying the TFI at 30 ms by the ratio $T/30$, where $T$ is the new exposure time in ms. The number of fluorescent introns at a given transcription site is then estimated by dividing the TFI of the transcription site by the average TFI of a single intron for the same exposure time.

**Quantification of number of GFP particles**

Single GFP molecules (eGFP-6His) were expressed in *E. coli* and purified by passage over TALON Metal Affinity Resin (Clontech) as previously described (Cureton et al., 2012). Triple-GFP particles were obtained by transient transfection of HEK-293 cells with a plasmid coding for 3 GFP molecules synthesized in tandem. The cells were lysed, and the cytosolic fraction was sequentially diluted. Single and triple-GFP particles were adsorbed to the surface of glass coverslips and imaged continuously with exposure time of 1000 ms. Typically, 100 to 300 particles were visible per field of view (86.5 × 86.5 µm). Images were analyzed with software written in MATLAB. Single diffraction-limited fluorescent particles were detected as local maxima using a particle-detection algorithm in a maximum intensity projection image of the time lapse sequence. Fluorescent traces were analyzed by calculating their first derivative to determine the time points at which photobleaching events occurred and subsequently fitting the intensity profile over time *f(t)* to sigmoid shaped functions of the form

$$\sum_{i=1}^{N_B} a_i \left( 1 - erf\left( \frac{f(t) - b_i}{s} \right) \right) + c$$

where *erf* corresponds to the error function (also called Gauss error function), $N_B$ is the number of bleaching events (1 for single GFP and up to 3 for 3xGFP), $b_i$ is the time point corresponding to photobleaching event *i*, $a_i$ is the average fluorescence intensity prior to photobleaching event *i*, *s* is the photobleaching slope and *c* the background intensity. The total fluorescence intensity ΔTFI corresponding to each individual bleaching step (each individual GFP) corresponds to the difference between average values of TFI between bleaching events for 3xGFP and to the average pre-bleach TFI in the case of single GFP. The same spinning-disk confocal microscope, objective and

acquisition settings (Gain, Intensification, binning and ND value) were used in acquiring images of transcription sites in live cells, single GFP molecules and 3xGFP particles. The number of GFP molecules that correspond to a given *TFI* acquired with exposure time *T* was then calculated with the formula

$$N_{GFP} = \frac{TFI}{TFI_{GFP}} \frac{T_{GFP}}{T}$$

where $TFI_{GFP}$ is the average total fluorescence intensity value for a single GFP particle, and $T_{GFP}$ is the exposure time for GFP (1000 ms).

**Statistical tests**

The statistical significance for the differences between intron lifetimes was determined using a Kruskal-Wallis-ANOVA test.

**SUPPLEMENTAL FIGURES TITLES AND LEGENDS**

**Figure S1 – Imaging beta-globin introns in living cells, Related to Figure 1. (A)** Illustration of the structure and size (bp) of the *HBB* transgenes β-λM and β-Mλ with binding sites for the coat protein of bacteriophage MS2 and the antiterminator protein N of bacteriophage λ inserted at the indicated positions. Both transgenes were generated from a modified *HBB* construct (β-WTΔ), with the second intron shortened by an internal deletion of 593 bp. The transgene includes a region of approximately 1800 bp past the poly(A) site from a genomic clone of the human beta-globin gene (Antoniou et al., 1998). Tetracycline-inducible expression is under the control of a minimal human

cytomegalovirus promoter (Pmin CMV). **(B)** Representative image showing two fluorescent dots in the nucleus spaced less than 1 μm apart. **(C)** The two dots were tracked in 4D and the centroid coordinates of each dot were estimated by fitting a two-dimensional Gaussian in each xy plane and a one-dimensional Gaussian in z. The mean inter-centroid distance is plotted in red (0.7 μm). The mean inter-centroid distance for the two dots is also plotted as a function of time. **(D)** Representative image showing two fluorescent dots in the nucleus spaced more than 3 μm apart. **(E)** 3D centroid positions measured for each dot over time with mean inter-centroid distance plotted in red (3.9 μm) and mean inter-centroid distances plotted as a function of time.

**Figure S2 – FISH shows introns labeled with MS2 co-localizing with full-length pre-mRNA, Related to Figure 1.** After transcriptional induction, cells expressing β-λM transcripts were fixed and hybridized with a Cy5-labeled probe complementary to the MS2 repeats (pseudocolored green) and a Cy3-labeled probe complementary to beta-globin exons (pseudocolored red). Double hybridization merged images are shown and enlarged insets depict the transcription site with the corresponding intensity line scans.

**Figure S3 – An image analysis application for 3D tracking and TFI quantification, Related to Figure 2. (A)** The initialization user interface window provides two options for analysis: single time lapse sequence processing and multiple time lapse sequences batch processing. In the single time lapse processing mode, TFI calculations can be performed for z-stack or single plane sequences (used in SSA experiments) and for a single dot at each time point or multiple dots in each time point (for SSA experiments). (**B**) The main user interface window allows for interactive visualization of time lapse

sequence(s), specification of XY coordinates for the dot (by clicking on the image) and parameterization of tracking and fitting variables, such as the search radius and sensitivity (which corresponds to the signal-to-noise threshold value, see Methods) and the fit radius. The graphs on the right are automatically refreshed after each fitting operation and plot the Z position, the peak intensity of the dot (the pixel with highest intensity value) and the TFI for each time point. Given initial coordinates, the dot position can be tracked automatically for the whole sequence using the option 'Auto Track'. The complete procedure of tracking and Gaussian fitting with specified parameters can be automated for every time point in the sequence (and for all sequences in the batch mode) using the option 'AUTO'. The XY coordinates for each individual timelapse can be recorded in a Microsoft Excel file for batch processing using the option 'Save XY Data' and the TFI, Z and I values from previously analyzed timelapses can be loaded for verification or re-analysis using the 'Load Data' option (**C**) A 2D Gaussian fit window is displayed at every fitting step for quality assessment. The grid represents the fitted 2D Gaussian surface plotted over the pseudo-color coded pixel intensity data. (**D**) TFI profiles obtained for each sequence (blue dots) are exported as a Microsoft Excel file and as a JPG image file.

**Figure S4 - Cycles of fluorescence gain and loss are visible for introns labeled with either MS2-GFP or λN-GFP, Related to Figure 4.** Graphs depicting time-course fluctuations in total fluorescence intensity (TFI). (**A**) Cells expressing beta-globin transcripts tagged with MS2-GFP in the first intron. (**B**) Cells expressing beta-globin transcripts tagged with MS2-GFP in the second intron. (**C**) Cells expressing beta-globin transcripts tagged with λN-GFP in the first intron. (**D**) Cells expressing beta-globin transcripts tagged with λN-GFP in the second intron.

**Figure S5 - Imaging the intron located between IgM exons M1 and M2, Related to Figure 5. (A)** Schematics of the structure and size (bp) of the IgM constructs with MS2 binding sites inserted at the indicated positions. Tetracycline-inducible expression is under the control of a minimal human cytomegalovirus promoter (Pmin CMV). **(B)** After transcriptional induction by tetracycline, the transcription site (arrow) is detected by mCherry-MS2 fluorescence and cyan fluorescence is seen in cytoplasmic peroxisomes.

**Figure S6 - Cycles of fluorescence gain and loss for IgM introns, Related to Figure 6. (A)** Graphs depicting discrete fluorescence fluctuations around background in cells expressing the IgM-weakPy construct. **(B)** Graphs depicting discrete fluorescence fluctuations around background in cells expressing the IgM-strongPy construct. **(C)** Graphs depicting discrete fluorescence fluctuations around background in cells expressing the IgM-strongPy+600 construct. **(D)** Graphs depicting discrete fluorescence fluctuations around background in cells expressing the IgM-strongPy+1700 construct. Total fluorescence intensity values were converted into number of GFP molecules.

**SUPPLEMENTAL MOVIES TITLES AND LEGENDS**

**Movie S1, In vivo time-lapse imaging and analysis of a cell where the second intron is labeled with MS2-GFP, Related to Figure 2 –** Time lapse maximum-intensity projection movie depicting fluctuations in fluorescence intensity at the site of transcription of a cell where the second intron is labeled with MS2-GFP (β-λM cell line). The inset at the top left shows a magnified view of the transcription site color-

coded for total fluorescence intensity. The panels on the right show the image analysis software output plots for maximum intensity Z-plane (top) and TFI value (bottom).

**Movie S2, In vivo time-lapse imaging of cells where the second intron is labeled with MS2-GFP that were exposed to SSA for 6h, Related to Figure 3 –** Time lapse movie depicting cells with the second intron labeled with MS2-GFP (β-λM cell line) that were exposed to SSA for 6h and imaged at the indicated time points. Each frame is a single 2D section from a z-stack acquired through the sample.

**Table S1 - Sequence of primers used in this study, Related to Figures 1 and 5.**

| Gene Name (organism) | Gene Symbol GeneID | Primer Designation | Primer Sequence (5'– 3') |
|---|---|---|---|
| Hemoglobin, beta (human) | HBB 3043 | E1f: HBB ExI_nt59 fw | ACGTGGATGAAGTTGGTGGT |
| | | E2f: HBB ExII_nt2 fw | CTGCTGGTGGTCTACCCTTG |
| | | E2r: HBB ExII_nt179 rev | CACTCAGTGTGGCAAAGGTG |
| | | E3r: HBB ExIII_nt27 rev | CACACAGACCAGCACGTTG |
| | | I1f: HBB IVSI_nt71 fw | TCTTGGGTTTCTGATAGGCAC |
| | | I2 f: HBB IVSII_nt52 fw | AGTTCATGTCATAGGAAGGGGAGAAG |
| | | pARr: HBB post_ExIII_nt478 rev | CTTGAATCCTTTTCTGAGGGATG |
| | | GAPDH.Rev | GAAGATGGTGATGGGATTTC |
| (Enterobacteria phage MS2) | | *Bgl*II5MS2For | GCAGATCTGTCCAACTACCAAACTGGGTC |
| | | *BamH*I5MS2Rev | GTAGGATCCGTGACACTATAGAATAGGGC |
| (Enterobacteria phage Lambda) | | *Bgl*II5BoxFor | CTAGATCTAACTGGGGATTCCTGGGC |
| | | *BamH*I5BoxRev | CGAGGATCCAGATAATATCCTCGATAGG |
| acyl-CoA-oxidase (human) | ACOX | PTS1-Fw | CCCTGTACAAGGCCACAGGAGTGTCTTAC AAGCACCTGAAGTCA |
| | | PTS1-Rev | GCTCTAGAGCTTCAGAGCTTGGACTGCAG TGACTTCAGGTGCTT |
| Immuno- globulin µ heavy chain constant region (mouse) | IgM | IgM-Fw | TTCGGTACCACCATGGTCACCCTGTTCAA GGTAGTATG |
| | | IgM-Rev | TGCCGGATCCCATTTCACCTGCAGGTGAC AGAGA |
| | | IgM-PY-Rev | TGCCGGATCCCATTTCACCTGCGGAAAAA AAAGG |
| | | IgM-Ex1-Univ-Fw | GAATTCTGCAGTCGACGGTAC |
| | | IgM-In1-Fw | CACGGATCTGAGGACACAGG |
| Cyan fluorescent protein (Aequorea. victoria) | CFP | IgM-Ex2-CFP-Rev | CAGGTCAGGGTGGTCACG |
| Bovine growth hormone | BGH | BGH-pA-Rev | GGTTCTTTCCGCCTCAGAAG |
| RNA Pol II - intron I (mouse) | mPolr2a | mPolr2a-In1-Fw | GGTCAGTGATGCTGATGTTGTGC |
| | | mPol2a-In1-Rev | GGTTTCCCCTGGCCCCAGACA |
| | | Polr2a-In1-BbvCI-Fw | AAGATGCCTCAGCAGGTAAGAGAACTG |
| | | Polr2a-In1-BbvCI-Rev | TCAGCTGAGGGATTATGGGAGCACAC |

## SUPPLEMENTAL REFERENCES


Antoniou, M., Geraghty, F., Hurst, J., and Grosveld, F. (1998). Efficient 3'-end formation of human beta-globin mRNA in vivo requires sequences within the last intron but occurs independently of the splicing reaction. Nucleic acids research *26*, 721-729.

Audibert, A., Weil, D., and Dautry, F. (2002). In vivo kinetics of mRNA splicing and transport in mammalian cells. Molecular and cellular biology *22*, 6706-6718.

Baron-Benhamou, J., Gehring, N.H., Kulozik, A.E., and Hentze, M.W. (2004). Using the lambdaN peptide to tether proteins to RNAs. Methods Mol Biol *257*, 135-154.

Beach, D.L., Salmon, E.D., and Bloom, K. (1999). Localization and anchoring of mRNA in budding yeast. Current biology : CB *9*, 569-578.

Bertrand, E., Chartrand, P., Schaefer, M., Shenoy, S.M., Singer, R.H., and Long, R.M. (1998). Localization of ASH1 mRNA particles in living yeast. Molecular cell *2*, 437-445.

Boireau, S., Maiuri, P., Basyuk, E., de la Mata, M., Knezevich, A., Pradet-Balade, B., Backer, V., Kornblihtt, A., Marcello, A., and Bertrand, E. (2007). The transcriptional cycle of HIV-1 in real-time and live cells. The Journal of cell biology *179*, 291-304.

Cureton, D.K., Harbison, C.E., Cocucci, E., Parrish, C.R., and Kirchhausen, T. (2012). Limited transferrin receptor clustering allows rapid diffusion of canine parvovirus into clathrin endocytic structures. Journal of virology *86*, 5330-5340.

Fusco, D., Accornero, N., Lavoie, B., Shenoy, S.M., Blanchard, J.M., Singer, R.H., and Bertrand, E. (2003). Single mRNA molecules demonstrate probabilistic movement in living mammalian cells. Current biology : CB *13*, 161-167.

Gattoni, R., Keohavong, P., and Stevenin, J. (1986). Splicing of the E2A premessenger RNA of adenovirus serotype 2. Multiple pathways in spite of excision of the entire large intron. Journal of molecular biology *187*, 379-397.

Gehring, N.H., Neu-Yilik, G., Schell, T., Hentze, M.W., and Kulozik, A.E. (2003). Y14 and hUpf3b form an NMD-activating complex. Molecular cell *11*, 939-949.

Guth, S., Martinez, C., Gaur, R.K., and Valcarcel, J. (1999). Evidence for substrate-specific requirement of the splicing factor U2AF(35) and for its function after polypyrimidine tract recognition by U2AF(65). Molecular and cellular biology *19*, 8263-8271.

Guth, S., Tange, T.O., Kellenberger, E., and Valcarcel, J. (2001). Dual function for U2AF(35) in AG-dependent pre-mRNA splicing. Molecular and cellular biology *21*, 7673-7681.

Lykke-Andersen, J., Shu, M.D., and Steitz, J.A. (2000). Human Upf proteins target an mRNA for nonsense-mediated decay when bound downstream of a termination codon. Cell *103*, 1121-1131.

Rust, M.J., Bates, M., and Zhuang, X. (2006). Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). Nature methods *3*, 793-795.

Schluter, A., Real-Chicharro, A., Gabaldon, T., Sanchez-Jimenez, F., and Pujol, A. (2010). PeroxisomeDB 2.0: an integrative view of the global peroxisomal metabolome. Nucleic acids research *38*, D800-805.

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. The Journal of cell biology *193*, 819-829.

Tyagi, S. (2009). Imaging intracellular RNA distribution and dynamics in living cells. Nature methods *6*, 331-338.