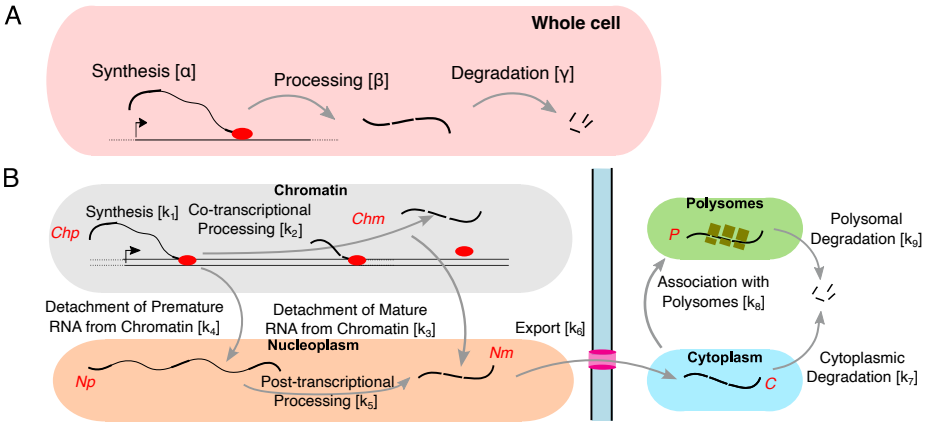
Nanodynamo quantifies subcellular RNA dynamics revealing extensive coupling between steps of the RNA life cycle

Supplemental information

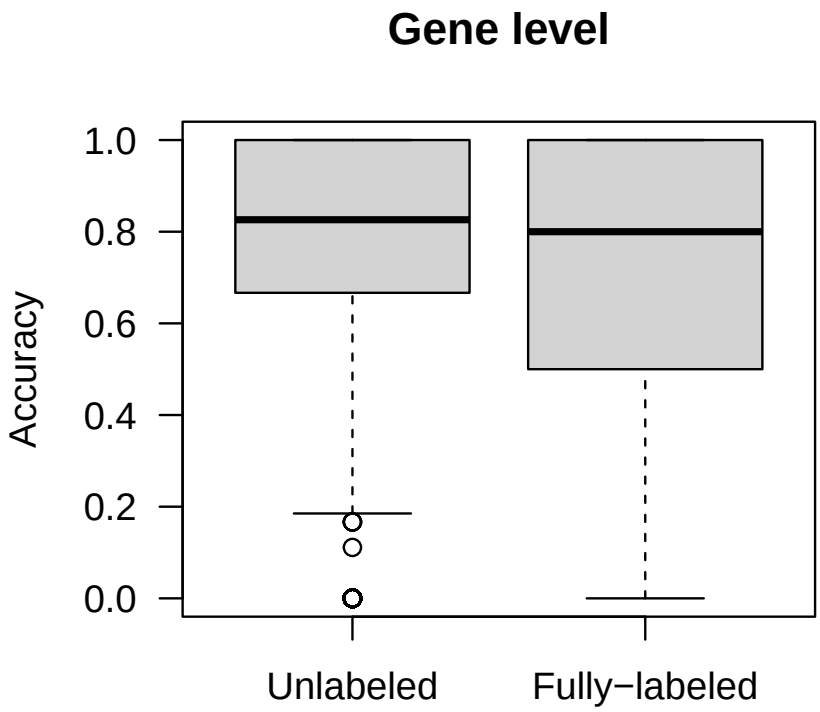
Lucia Coscujuela Tarrero, Valeria Famà, Giacomo D'Andrea, Simone Maestri, Anna de Polo, Stefano Biffo, Mattia Furlan, Mattia Pelizzola

Contents

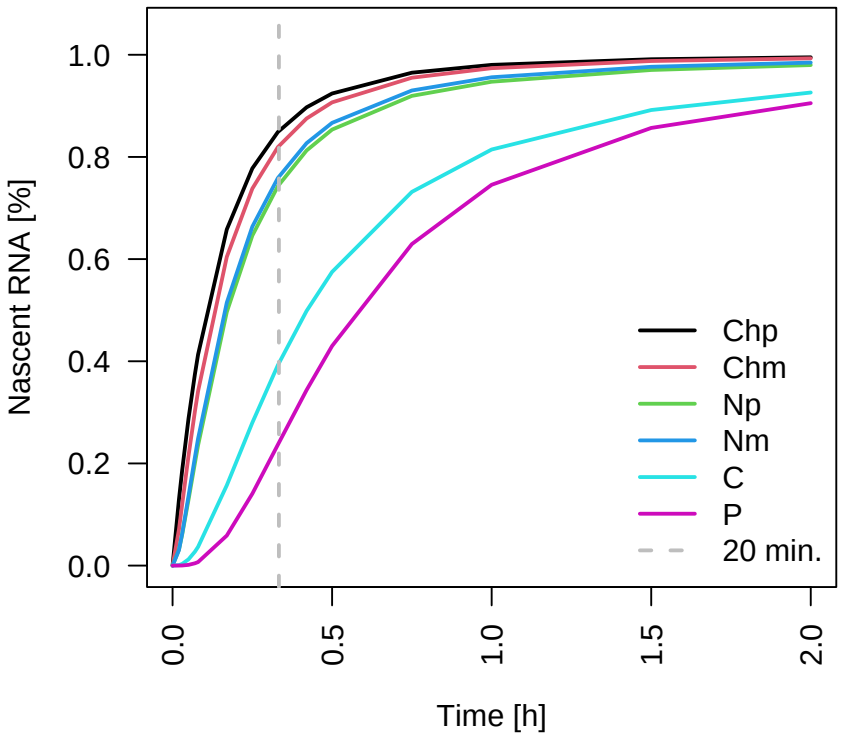
1 Supplementary figures



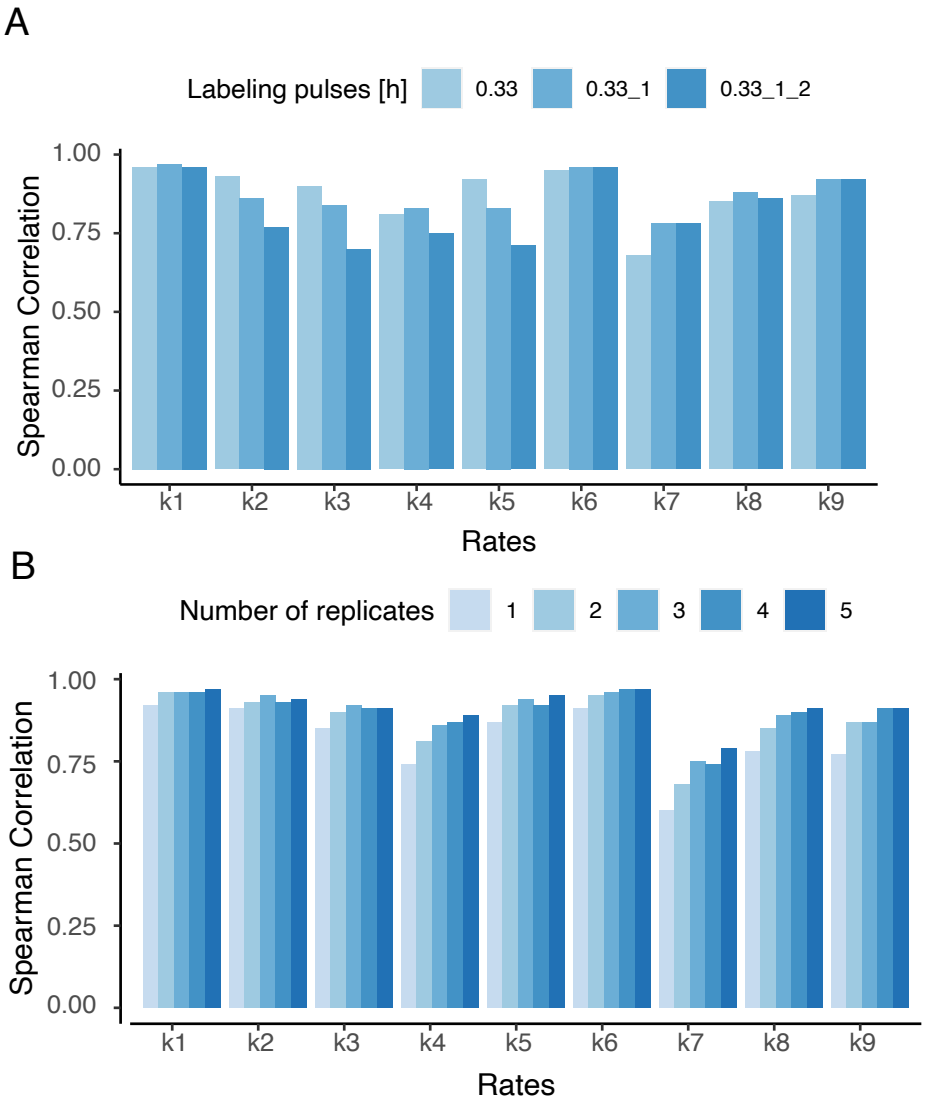
Supplementary Figure 1 - **The RNA life cycle.** (A) Scheme of the RNA life cycle implemented in INSPEcT accounting for RNA synthesis, processing, and degradation. (B) Scheme of the RNA life cycle implemented in Nanodynamo, including RNA synthesis, detachment, co- and post-transcriptional processing, export, association to polysomes, cytoplasmic and polysomal degradation.



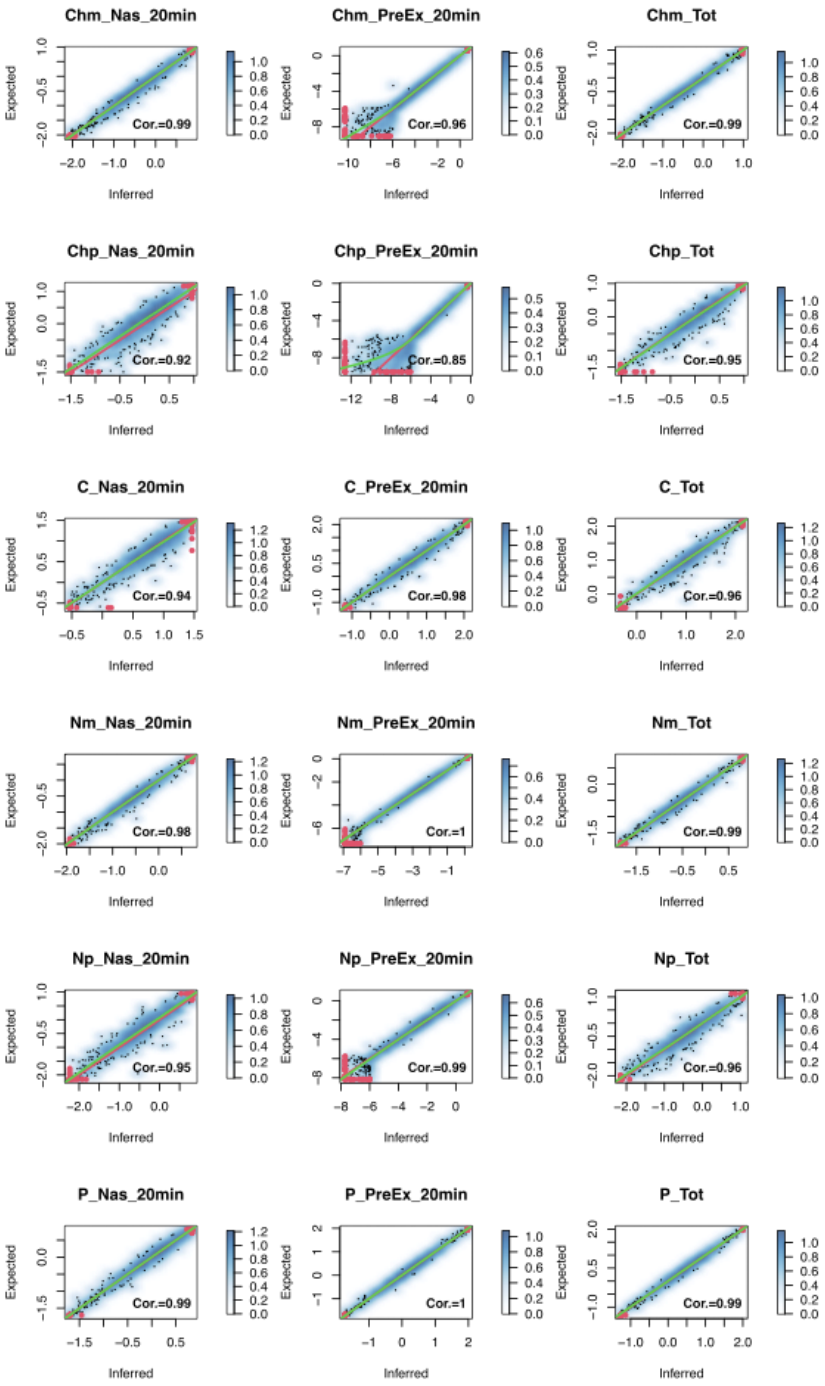
Supplementary Figure 2: **Nano-ID accuracy.** Accuracy distributions, at single gene resolution, for unlabelled and labelled reads; evaluation performed on the test set (9.5k genes, 455k unlabelled reads, 219k labelled reads). The horizontal line represents the median value, the box edges represent the 25th (Q1) and 75th (Q3) percentiles, and the whiskers show the range of data excluding outliers (observations lower that Q1 - 1.5 * interquartile range or larger than Q3 + 1.5 * interquartile range). Source data are provided as a Source Data file.



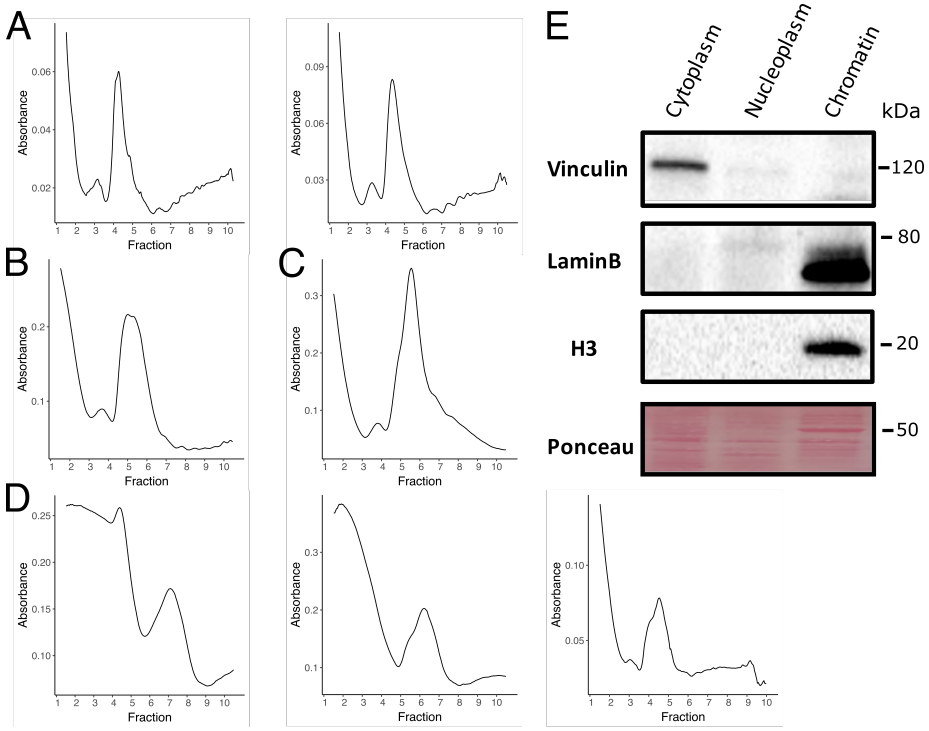
Supplementary Figure 3: **Nascent RNA saturation.** Saturation profile of nascent RNA for all RNA species involved in the model. The profiles were obtained as the mean of 1000 genes simulated without noise (CV=0). Ch, N, C and P refer to chromatin associated, nucleoplasmatic, cytoplasmic and polysomal RNA, respectively, p and m further specifying the premature or mature form. Source data are provided as a Source Data file.



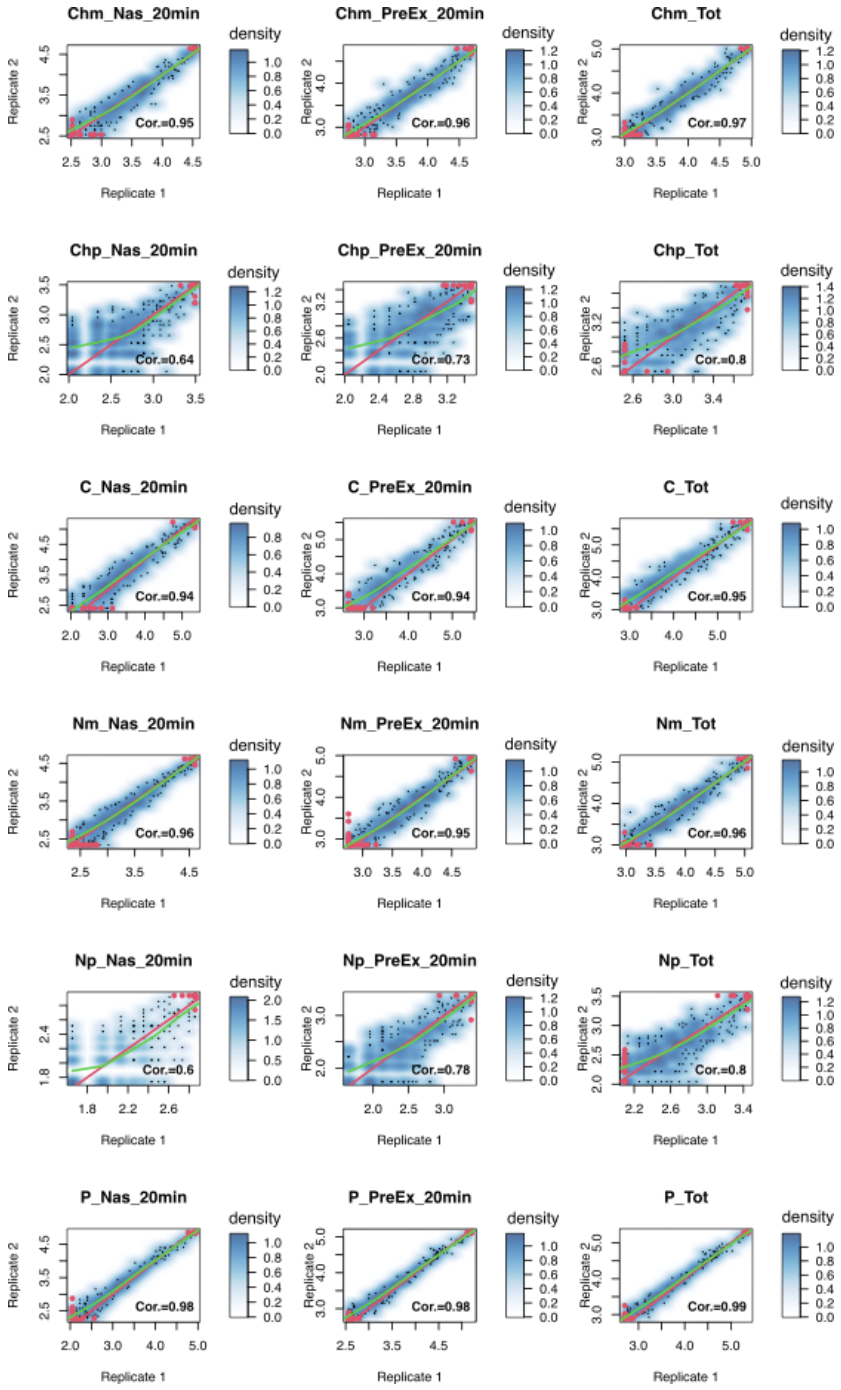
Supplementary Figure 4: **Labelling times set-up.** (A) Spearman correlation coefficients for three datasets simulated with 1000 genes, CVs assigned accordingly to each species, 2 replicates and increasing number of labelling pulses (20, 60, 120 minutes). (B) Spearman correlation coefficients for five datasets simulated with 1000 genes, CVs assigned accordingly to each species, a labelling pulse of 20 minute and an increasing number of replicates (from 1 to five). Source data are provided as a Source Data file.



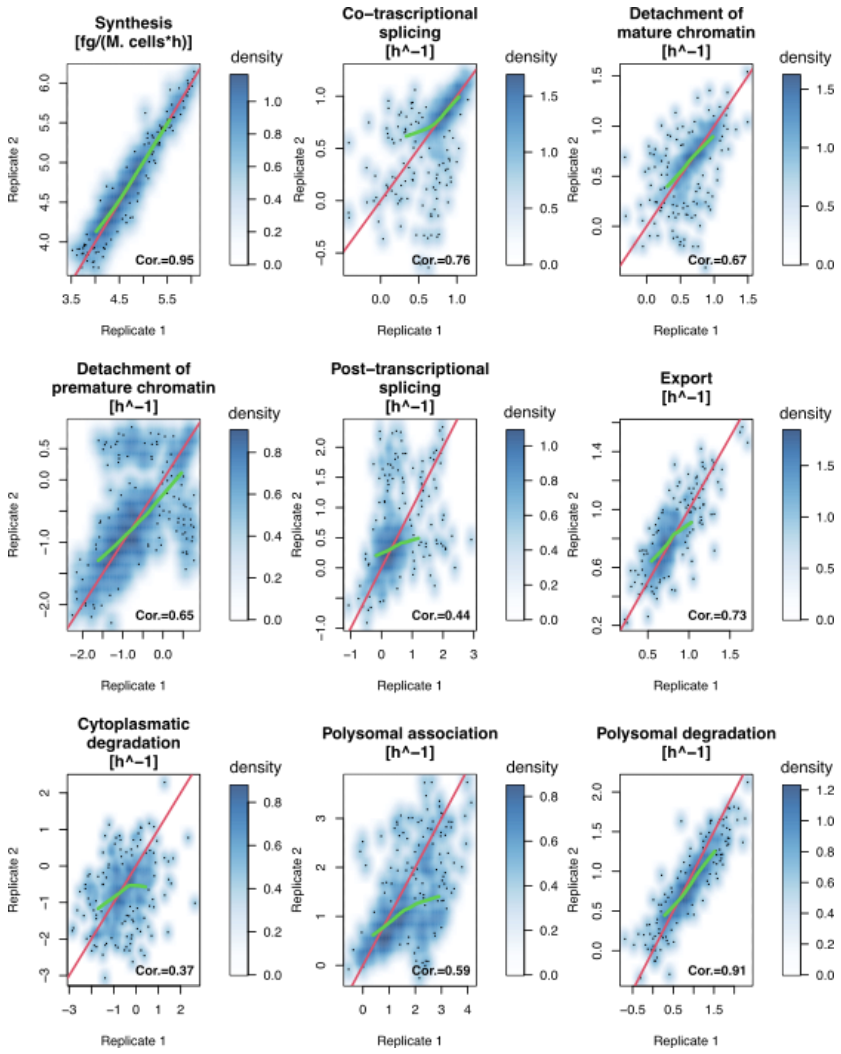
Supplementary Figure 5: **Goodness of fit on simulated data.** Smooth density scatterplots between modelled RNA expression levels and their expected counterparts for a simulated dataset of 1000 genes generated with CVs assigned accordingly to each species, 2 replicates, 1 labelling time of 20'. For each plot, we report the identity line (red), the loess line (green), and the Spearman correlation coefficient (Cor.). Red dots represent saturated points (i.e. data points in the bottom and top 2.5% of the distribution). Ch, N, C and P refer to chromatin associated, nucleoplasmatic, cytoplasmatic and polysomal RNA, respectively, p and m further specifying the premature or mature form. Source data are provided as a Source Data file.



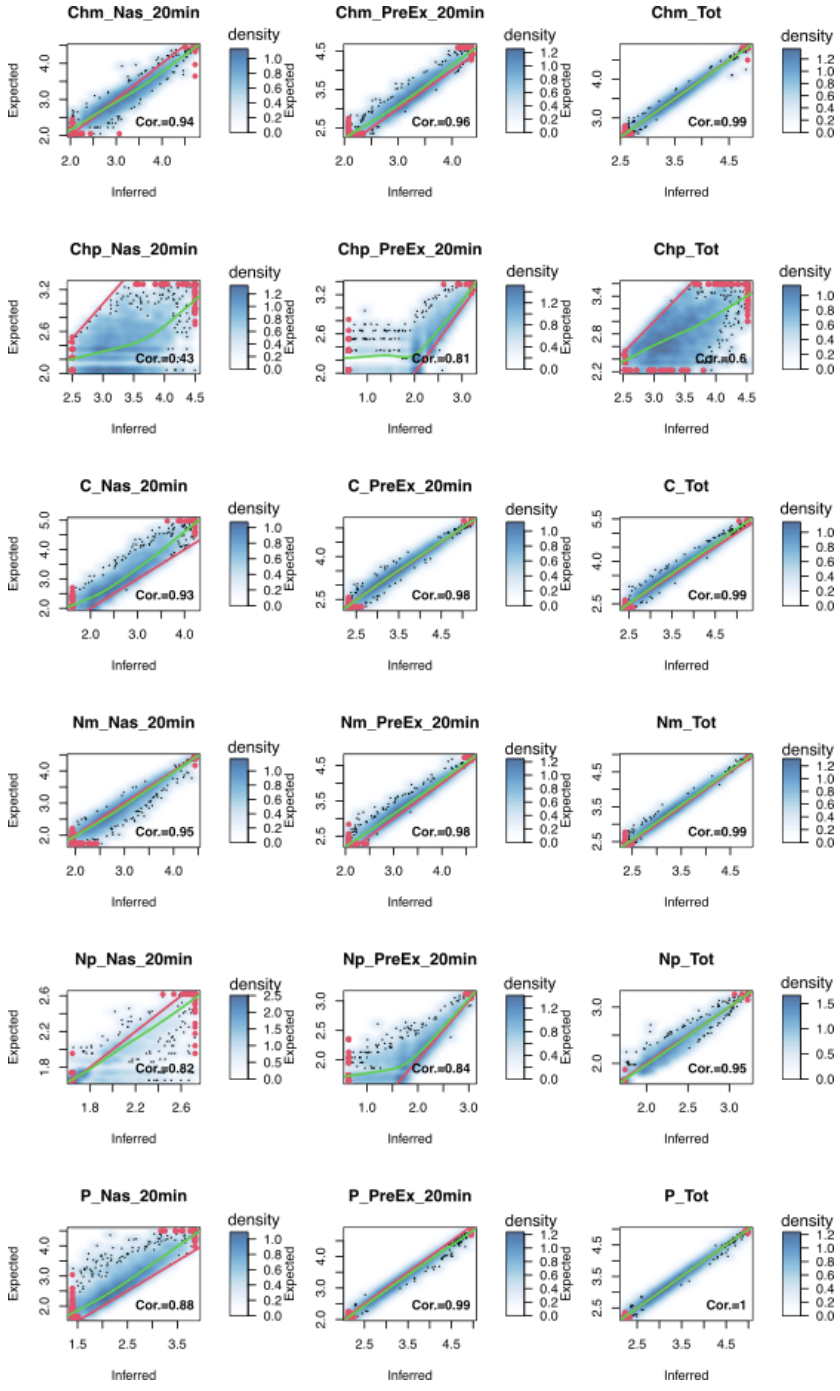
Supplementary Figure 6: **Polysomal tracing and validation of cellular fractionation.** (A) Polysomal tracing for untreated SUM159 cells. (B) Polysomal tracing for Pladienolide B treated cells. (C) Polysomal tracing for Harringtonine treated cells. (D) Polysomal tracing for Leptomycin B treated cells. (E) Western blot analysis of the three different fractions of SUM159 cells carried out with the indicated antibodies. Ponceau was performed as a loading control. The experiment was repeated in triplicate obtaining similar results. Source data are provided as a Source Data file, Western blot acquisitions are available as Supplementary Figure 57.



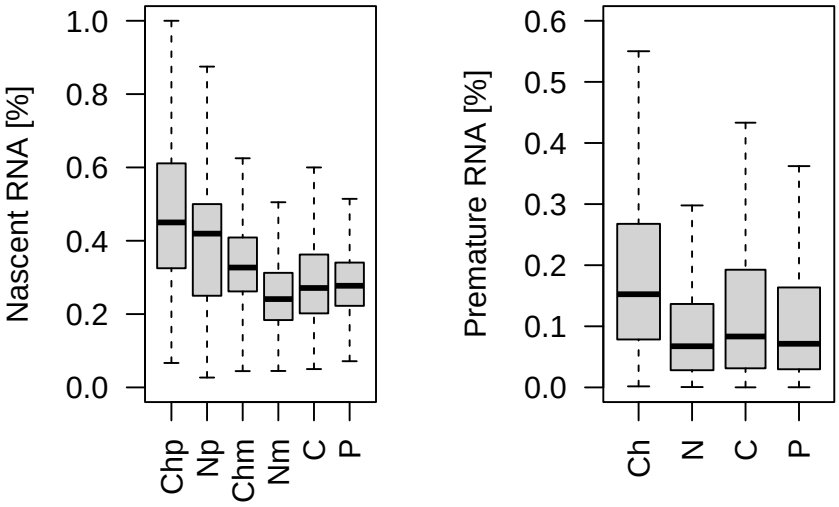
Supplementary Figure 7: **Untreated SUM159 reproducibility – Expression levels.** Smooth density scatterplots between the experimental expression levels profiled in 2 biological replicates. For each scatterplot, we report the identity line (red), the loess line (green), and the Spearman correlation coefficient (Cor.). Red dots represent saturated points (i.e. data points in the bottom and top 5% of the distribution). Ch, N, C and P refer to chromatin associated, nucleoplasmatic, cytoplasmatic and polysomal RNA, respectively, p and m further specifying the premature or mature form. Source data are provided as a Source Data file.



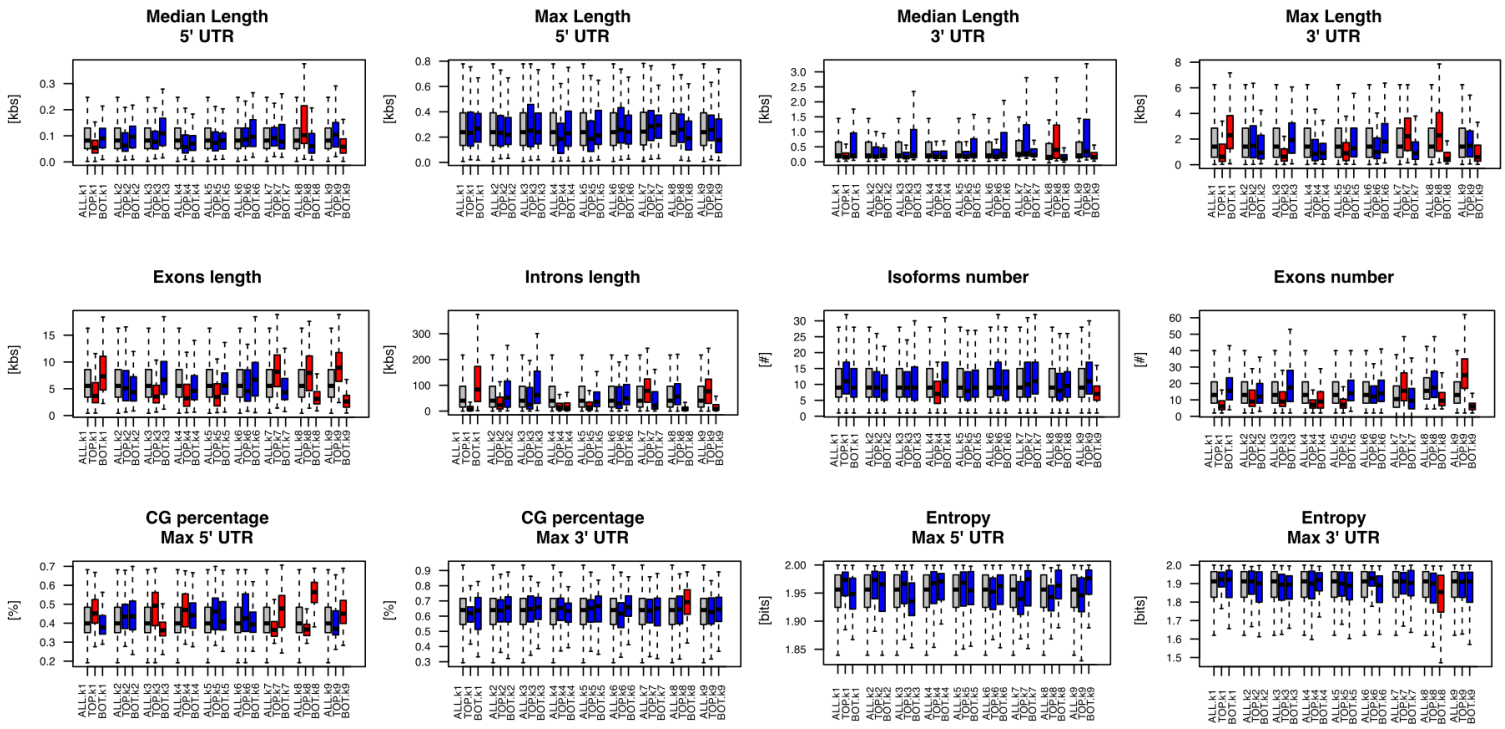
Supplementary Figure 8: **Untreated SUM159 reproducibility – Kinetic rates.** Smooth density scatterplots between the inferred kinetic rates profiled in 2 biological replicates. For each scatterplot, we report the identity line (red), the loess line (green), and the Spearman correlation coefficient (Cor.). Source data are provided as a Source Data file.



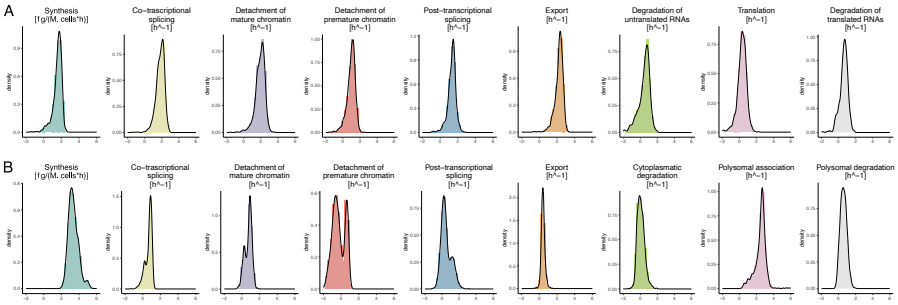
Supplementary Figure 9: **Untreated SUM159 goodness of fit.** Comparison between expected and inferred expression levels. For each smooth density scatterplot, we report the identity line (red), the loess line (green), and the Spearman correlation coefficient (Cor.). Red dots represent saturated points (i.e. data points in the bottom and top 2.5% of the distribution). Ch, N, C and P refer to chromatin associated, nucleoplasmatic, cytoplasmatic and polysomal RNA, respectively, p and m further specifying the premature or mature form. Source data are provided as a Source Data file.



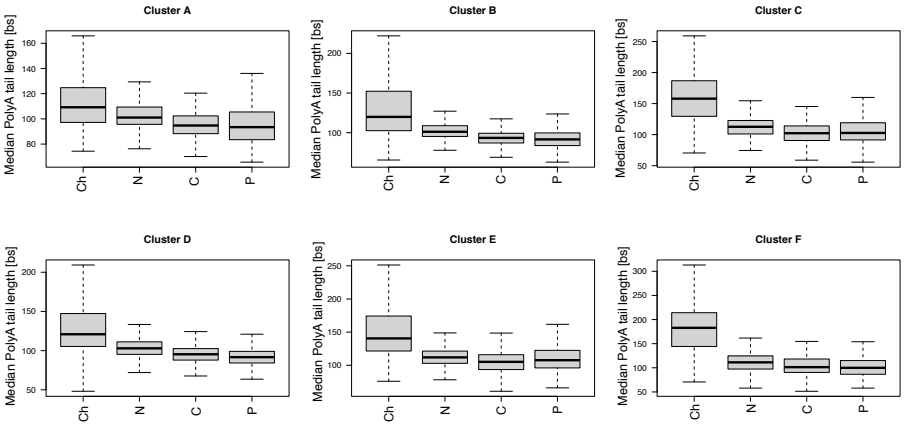
Supplementary Figure 10: **Nascent and Premature RNA proportions in untreated SUM159 cells.** Distributions of nascent (left) and premature (right) reads proportions, at single-gene resolution, for all the RNA species involved in the model; replicates mean values. Ch, N, C and P refer to chromatin associated, nucleoplasmatic, cytoplasmatic and polysomal RNA, respectively, p and m further specifying the premature or mature form. The horizontal line represents the median value, the box edges represent the 25th (Q1) and 75th (Q3) percentiles, and the whiskers show the range of data excluding outliers (observations lower that Q1 - 1.5 * interquartile range or larger than Q3 + 1.5 * interquartile range). Source data are provided as a Source Data file.



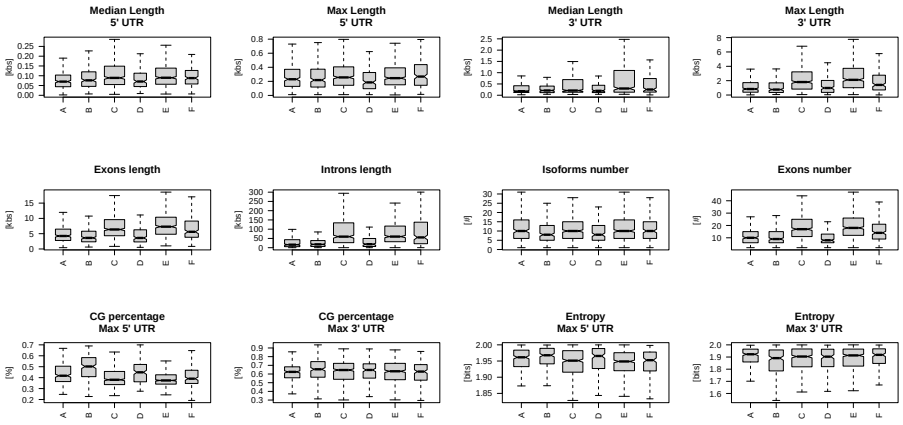
Supplementary Figure 11: **Sequence features of genes with fast and slow kinetics in untreated SUM159.** Characterization in terms of sequence features of the top and bottom 100 genes for each kinetic rate. In grey, we reported feature distributions for all the analysed genes. In red, we reported distributions for the top/ bottom genes which statistically differ from the counterpart estimated on all the genes (P < 1e-4). In blue, we reported distributions for the top/ bottom genes which do not statistically differ from the counterpart estimated on all the genes. The horizontal line represents the median value, the box edges represent the 25th (Q1) and 75th (Q3) percentiles, and the whiskers show the range of data excluding outliers (observations lower that Q1 - 1.5 * interquartile range or larger than Q3 + 1.5 * interquartile range). Source data are provided as a Source Data file.



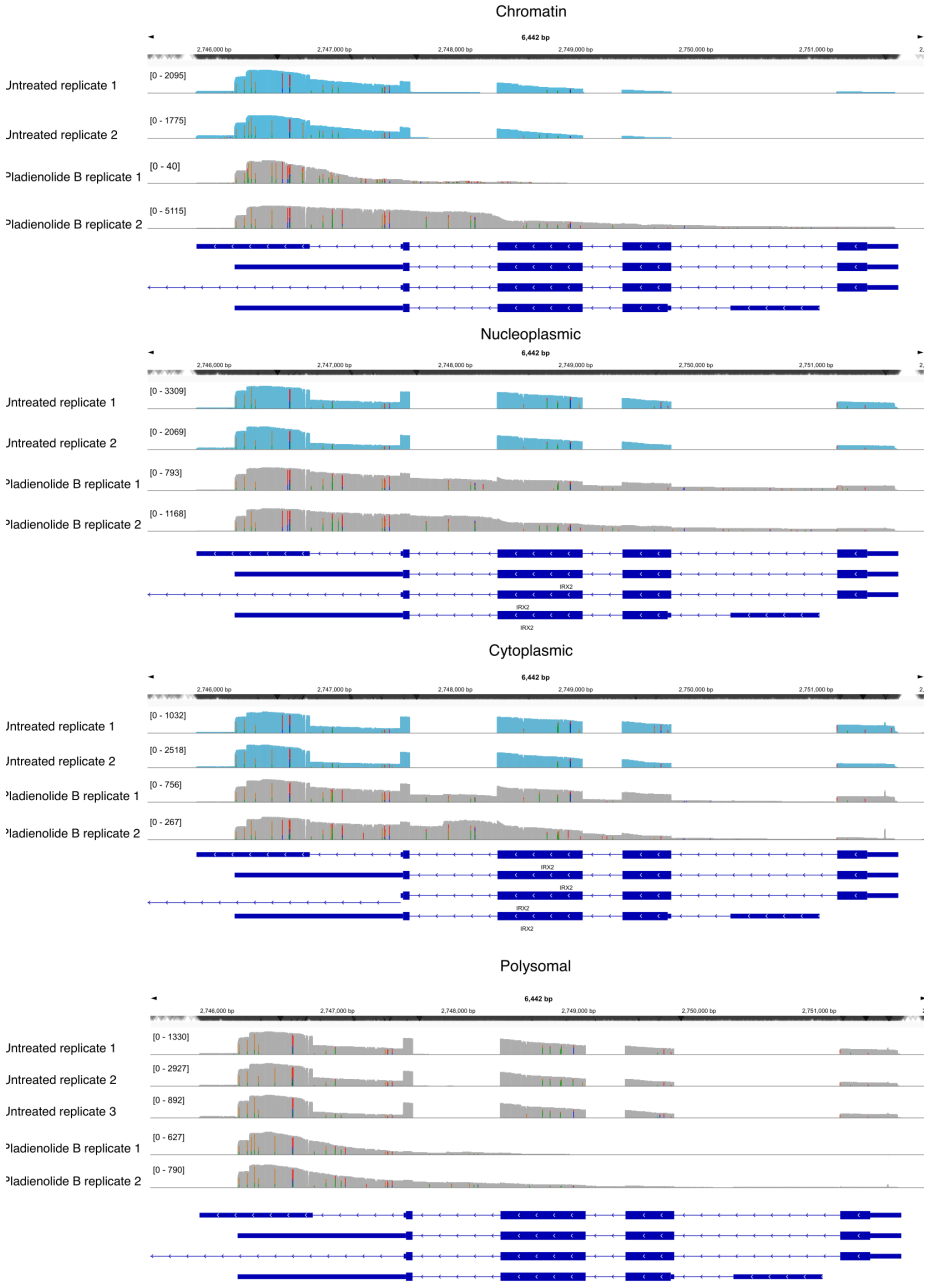
Supplementary Figure 12: **Rates distributions – Simulated dataset and untreated SUM159 cells. (A)** Distributions of inferred kinetic rates, in log_{10} scale, for a simulated dataset of 1000 genes generated with CV assigned accordingly to each species, 2 replicates, 1 labelling time of 20'. **(B)** Distributions of inferred kinetic rates, in Log10 scale, for untreated SUM159 cells; one initial condition instead of three was used to perform this inference. Source data are provided as a Source Data file.



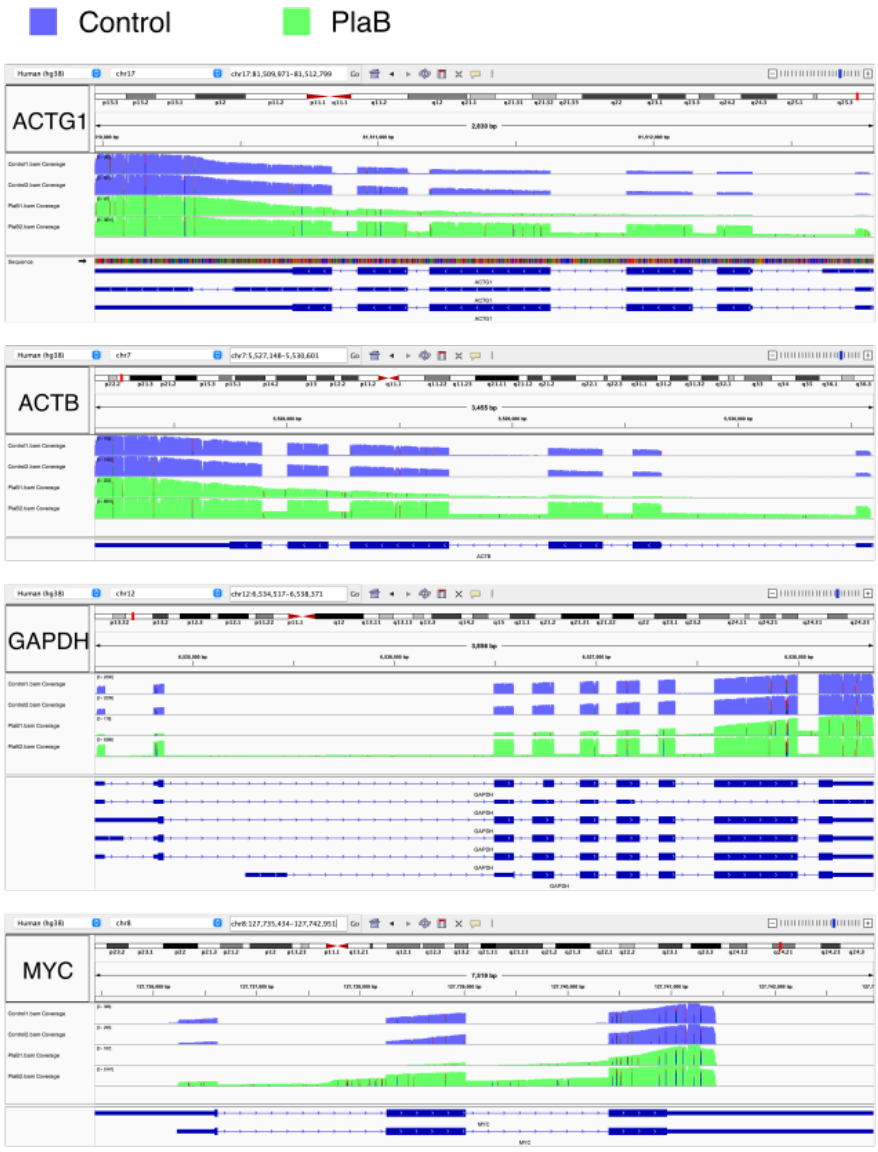
Supplementary Figure 13: **Characterization of polyA tails in genes clusters for untreated SUM159.** Gene-level median polyA tail length distribution, over different RNA pools, for groups of genes identified by k-means clustering according to both abundance of RNA species and kinetic rates as reported in Figure 2E. Ch, N, C and P refer to chromatin associated, nucleoplasmatic, cytoplasmatic and polysomal RNA, respectively, p and m further specifying the premature or mature form. The horizontal line represents the median value, the box edges represent the 25th (Q1) and 75th (Q3) percentiles, and the whiskers show the range of data excluding outliers (observations lower that Q1 - 1.5 * interquartile range or larger than Q3 + 1.5 * interquartile range). Source data are provided as a Source Data file.



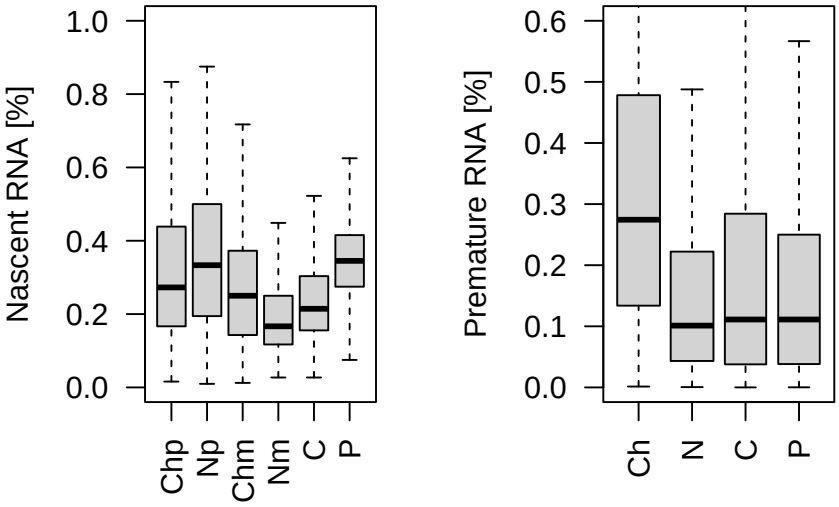
Supplementary Figure 14: **Sequence features in genes clusters for untreated SUM159.** Box plot for various sequence features for the genes in the clusters identified according to gene expression levels and inferred kinetic rates in untreated SUM159 cells as reported in Figure 2E. The horizontal line represents the median value, the box edges represent the 25th (Q1) and 75th (Q3) percentiles, and the whiskers show the range of data excluding outliers (observations lower that Q1 - 1.5 * interquartile range or larger than Q3 + 1.5 * interquartile range). Source data are provided as a Source Data file.



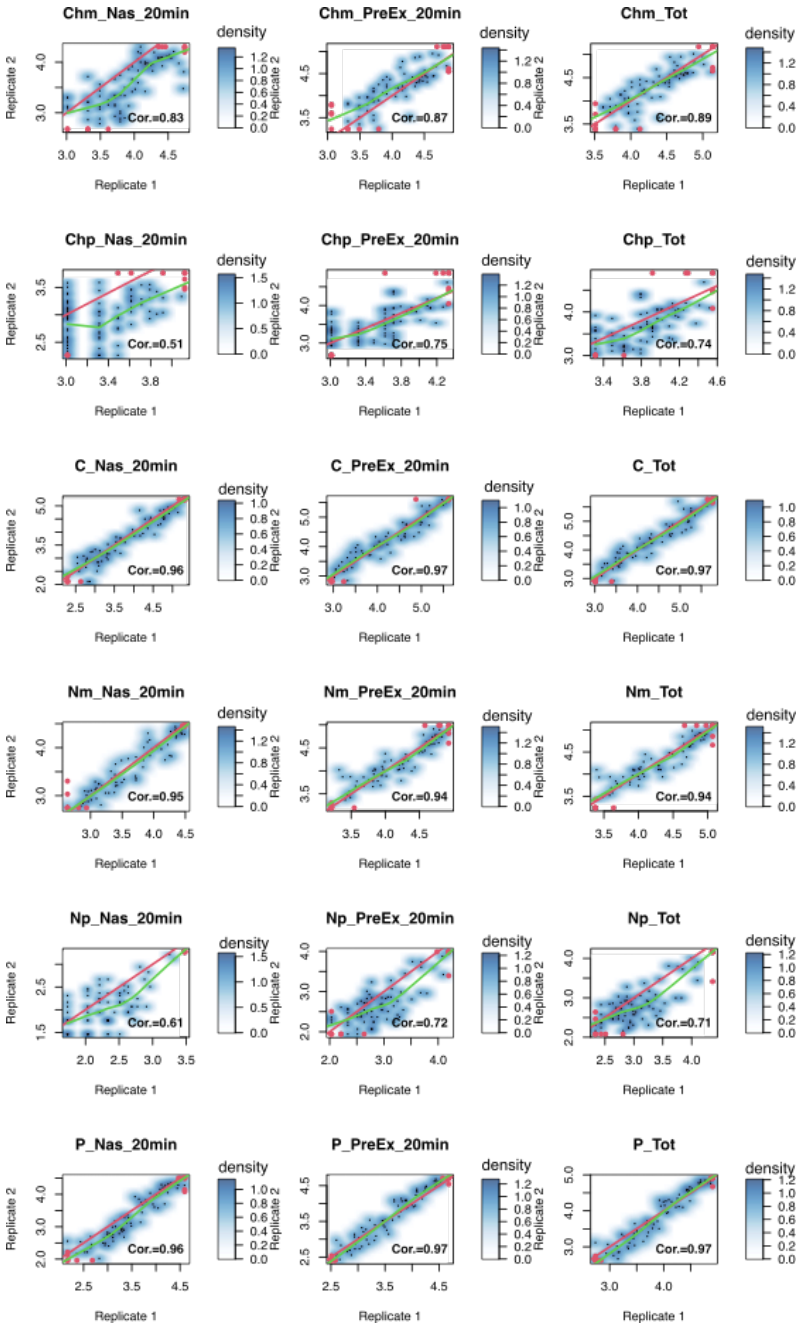
Supplementary Figure 15: ***IRX2.*** Genome browser snapshot of the coverage of IRX2 gene in untreated (light blue) and Pladienolide B treated (grey) cells for every fraction; intronic regions are depicted as thin blue lines.



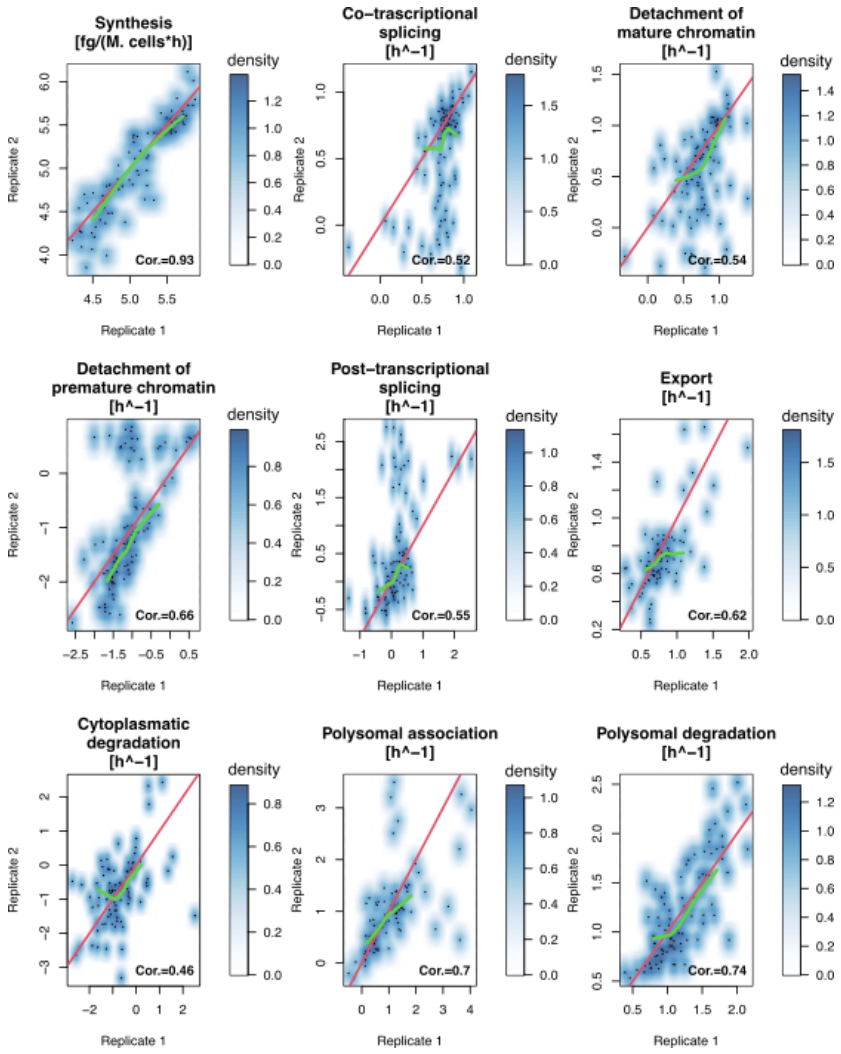
Supplementary Figure 16: **Validation of Pladienolide B treatment.** Genome browser snapshots of the coverage of several genes in untreated (blue) and Pladienolide B treated (green) cells; intronic regions are depicted as thin blue lines.



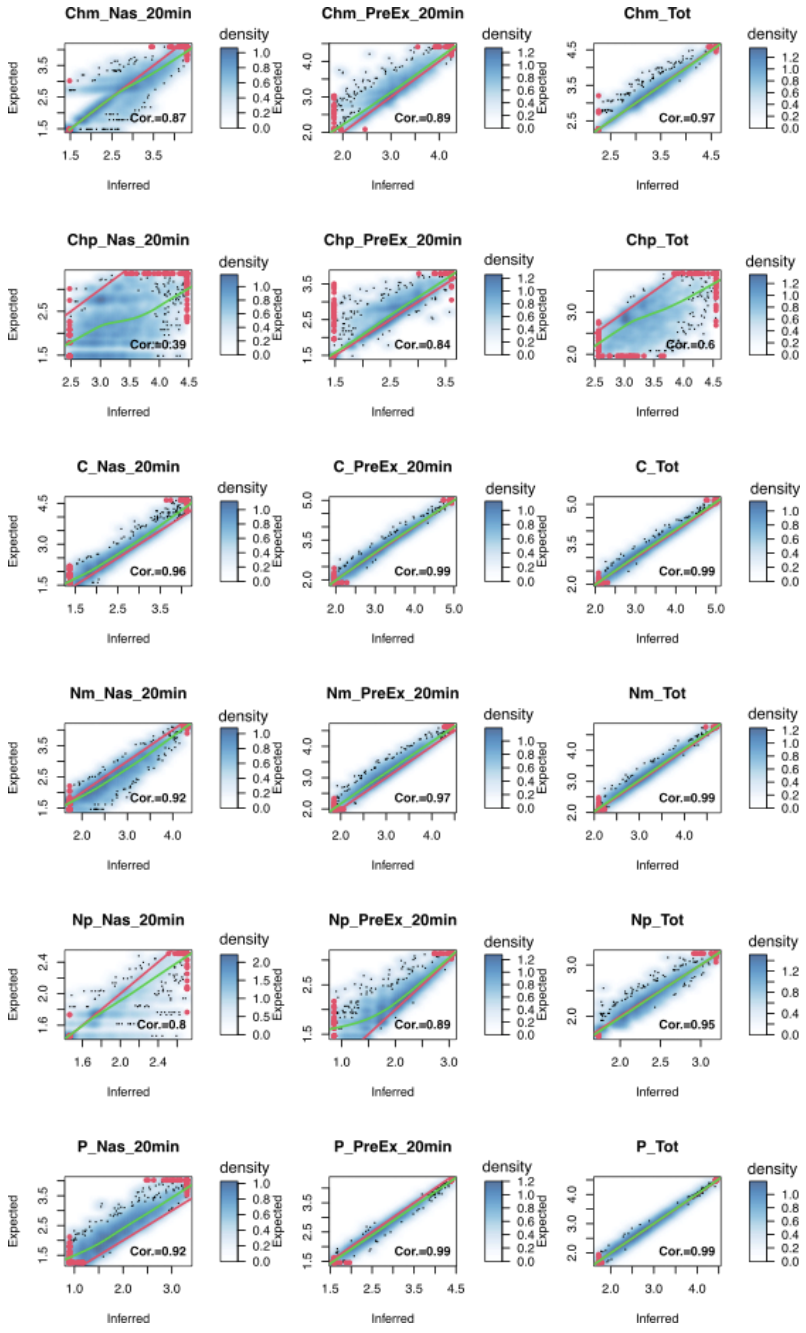
Supplementary Figure 17: **Nascent and premature RNA proportions in Pladienolide B treated SUM159 cells.** Distributions of nascent (left) and premature (right) reads proportions, at single gene resolution, for all the RNA species involved in the model. Ch, N, C and P refer to chromatin associated, nucleoplasmatic, cytoplasmatic and polysomal RNA, respectively, p and m further specifying the premature or mature form. The horizontal line represents the median value, the box edges represent the 25th (Q1) and 75th (Q3) percentiles, and the whiskers show the range of data excluding outliers (observations lower that Q1 - 1.5 * interquartile range or larger than Q3 + 1.5 * interquartile range). Source data are provided as a Source Data file.



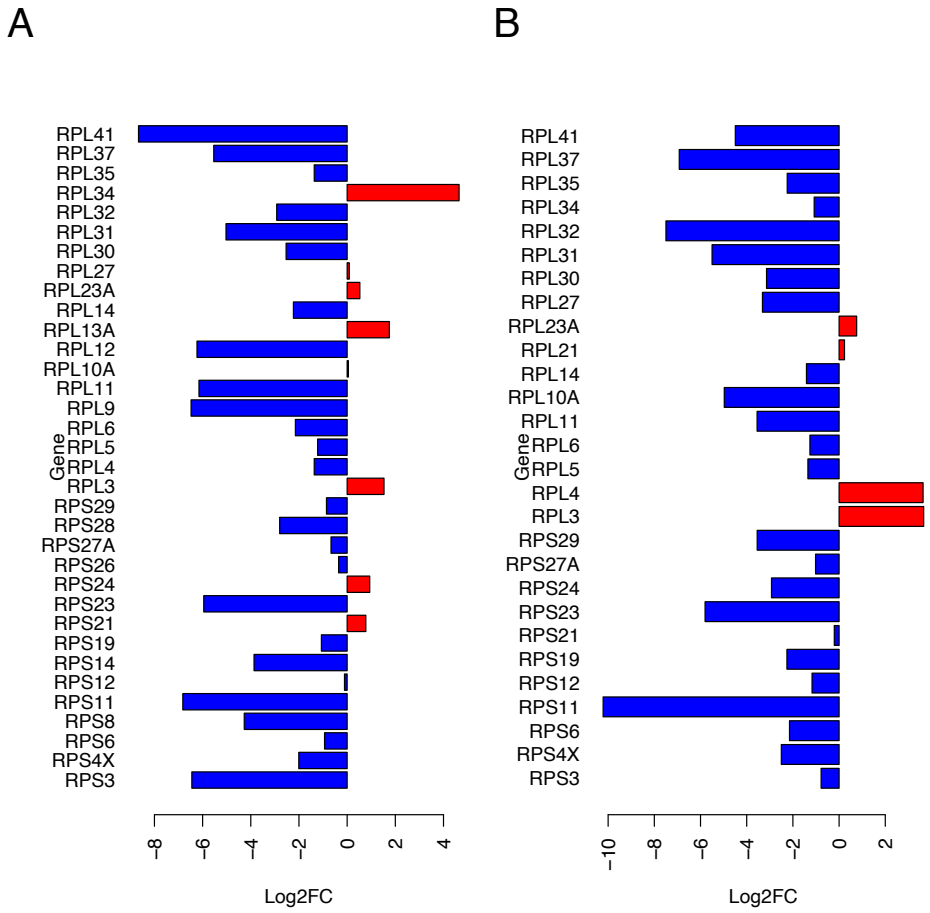
Supplementary Figure 18: **Pladienolide B treated SUM159 reproducibility – Expression levels.** Smooth density scatterplots between the experimental expression levels profiled in 2 biological replicates. For each scatterplot, we report the identity line (red), the loess line (green), and the Spearman correlation coefficient (Cor.). Red dots represent saturated points (i.e. data points in the bottom and top 5% of the distribution). Ch, N, C and P refer to chromatin associated, nucleoplasmatic, cytoplasmatic and polysomal RNA, respectively, p and m further specifying the premature or mature form. Source data are provided as a Source Data file.



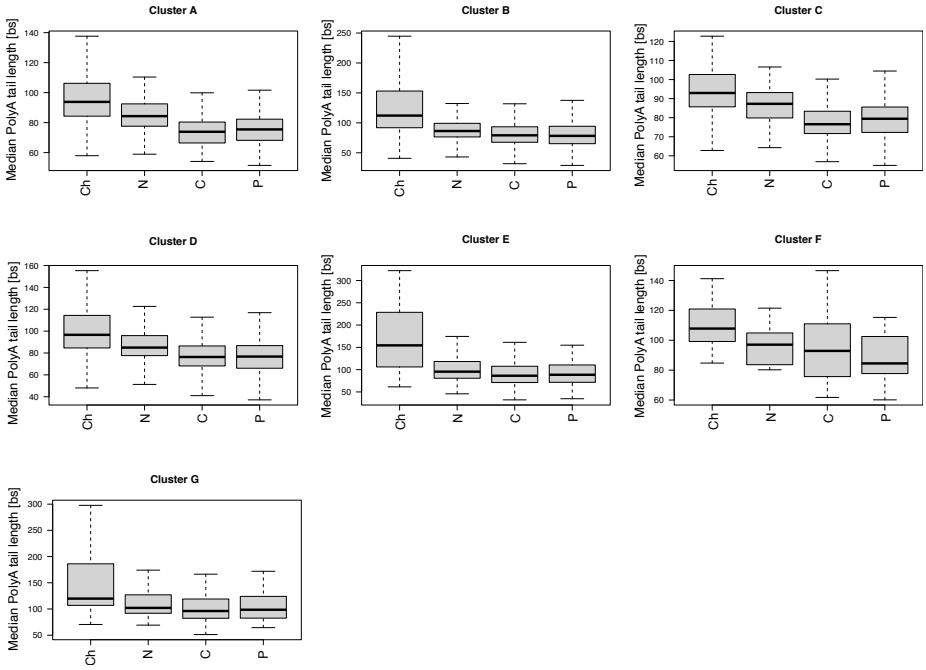
Supplementary Figure 19: **Pladienolide B treated SUM159 reproducibility – Kinetic rates.** Smooth density scatterplots between the inferred kinetic rates profiled in 2 biological replicates. For each scatterplot, we report the identity line (red), the loess line (green), and the Spearman correlation coefficient (Cor.). Source data are provided as a Source Data file.



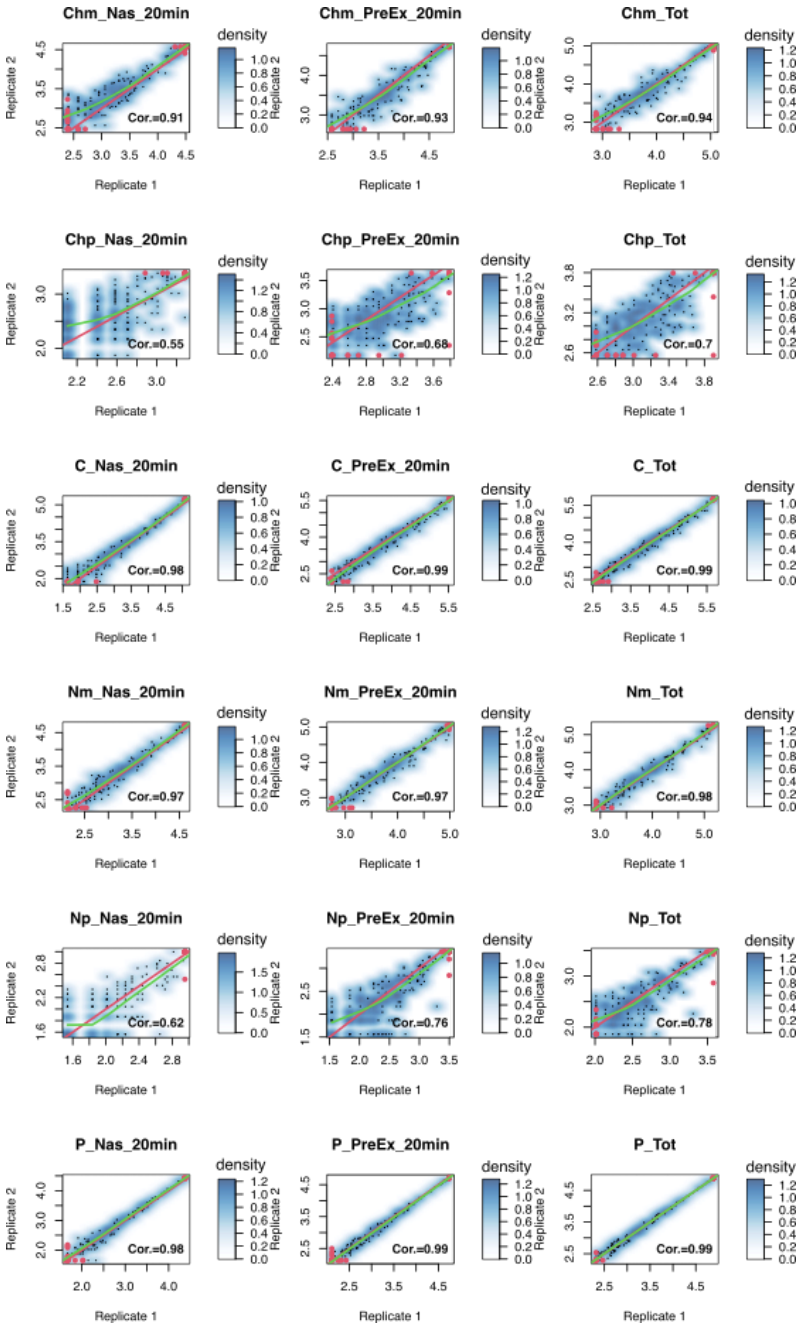
Supplementary Figure 20: **Pladienolide B treated SUM159 goodness of fit.** Comparison between expected and inferred expression levels. For each smooth density scatterplot, we report the identity line (red), the loess line (green), and the Spearman correlation coefficient (Cor.). Red dots represent saturated points (i.e. data points in the bottom and top 2.5% of the distribution). Ch, N, C and P refer to chromatin associated, nucleoplasmatic, cytoplasmatic and polysomal RNA, respectively, p and m further specifying the premature or mature form. Source data are provided as a Source Data file.



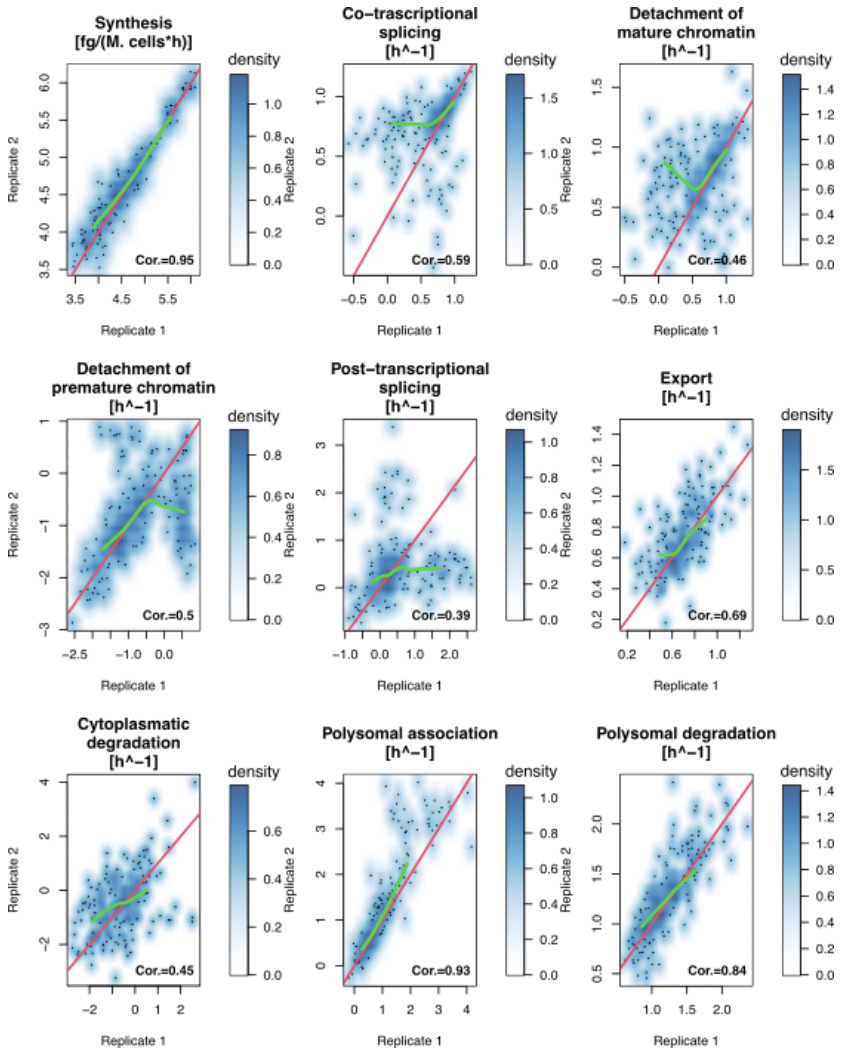
Supplementary Figure 21: **Modulation of the rate of association with polysomes for the 5'TOP factors in response to treatments.** Log_2 ratio of the rate of polysomal association for a set of 5'TOP factors in response to Pladienolide B (A) and Leptomycin B (B) treatments compared to the untreated condition; up- and down-regulations are depicted in red and blue respectively. Source data are provided as a Source Data file.



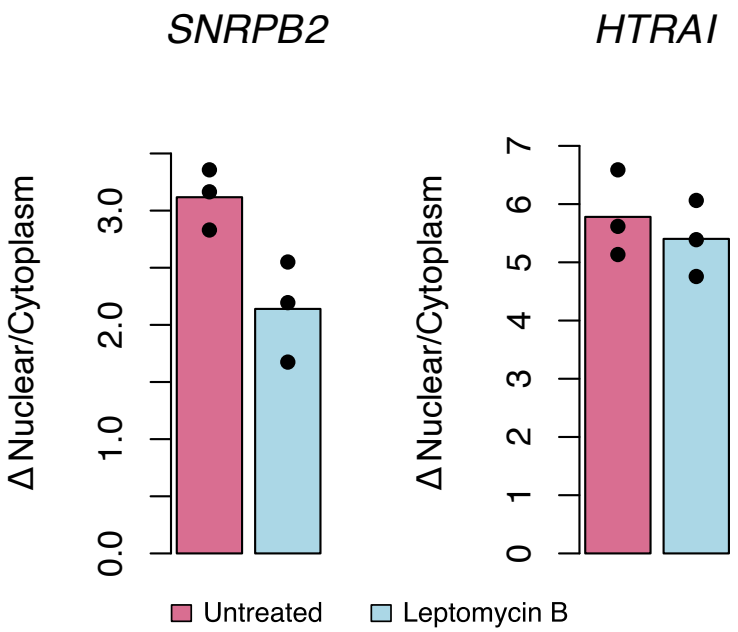
Supplementary Figure 22: **Characterization of PolyA tails in genes clusters for Pladienolide B treated SUM159 cells.** Gene-level median polyA tail length distribution, over different RNA pools, for groups of genes identified by k-means clustering according to the modulation of both their RNA abundances and kinetic rates compared to the untreated condition, as reported in Figure 3F. Ch, N, C and P refer to chromatin associated, nucleoplasmatic, cytoplasmatic and polysomal RNA, respectively, p and m further specifying the premature or mature form. The horizontal line represents the median value, the box edges represent the 25th (Q1) and 75th (Q3) percentiles, and the whiskers show the range of data excluding outliers (observations lower that Q1 - 1.5 * interquartile range or larger than Q3 + 1.5 * interquartile range). Source data are provided as a Source Data file.



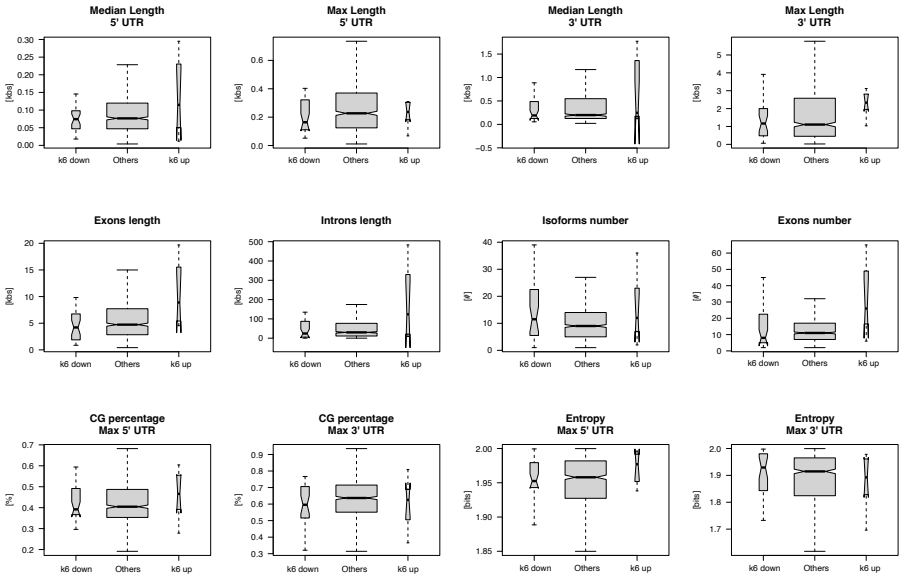
Supplementary Figure 23: **Leptomycin B treated SUM159 cells reproducibility – Expression Levels.** Smooth density scatterplots between the experimental expression levels profiled in 2 biological replicates. For each scatterplot, we report the identity line (red), the loess line (green), and the Spearman correlation coefficient (Cor.). Red dots represent saturated points (i.e. data points in the bottom and top 5% of the distribution). Ch, N, C and P refer to chromatin associated, nucleoplasmatic, cytoplasmatic and polysomal RNA, respectively, p and m further specifying the premature or mature form. Source data are provided as a Source Data file.



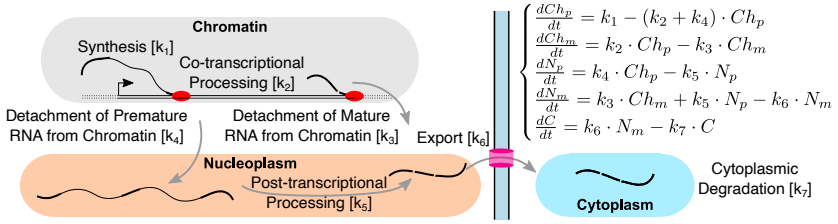
Supplementary Figure 24: **Leptomycin B treated SUM159 cells reproducibility – Kinetic rates.** Smooth density scatterplots between the inferred kinetic rates profiled in 2 biological replicates. For each scatterplot, we report the identity line (red), the loess line (green), and the Spearman correlation coefficient (Cor.). Source data are provided as a Source Data file.



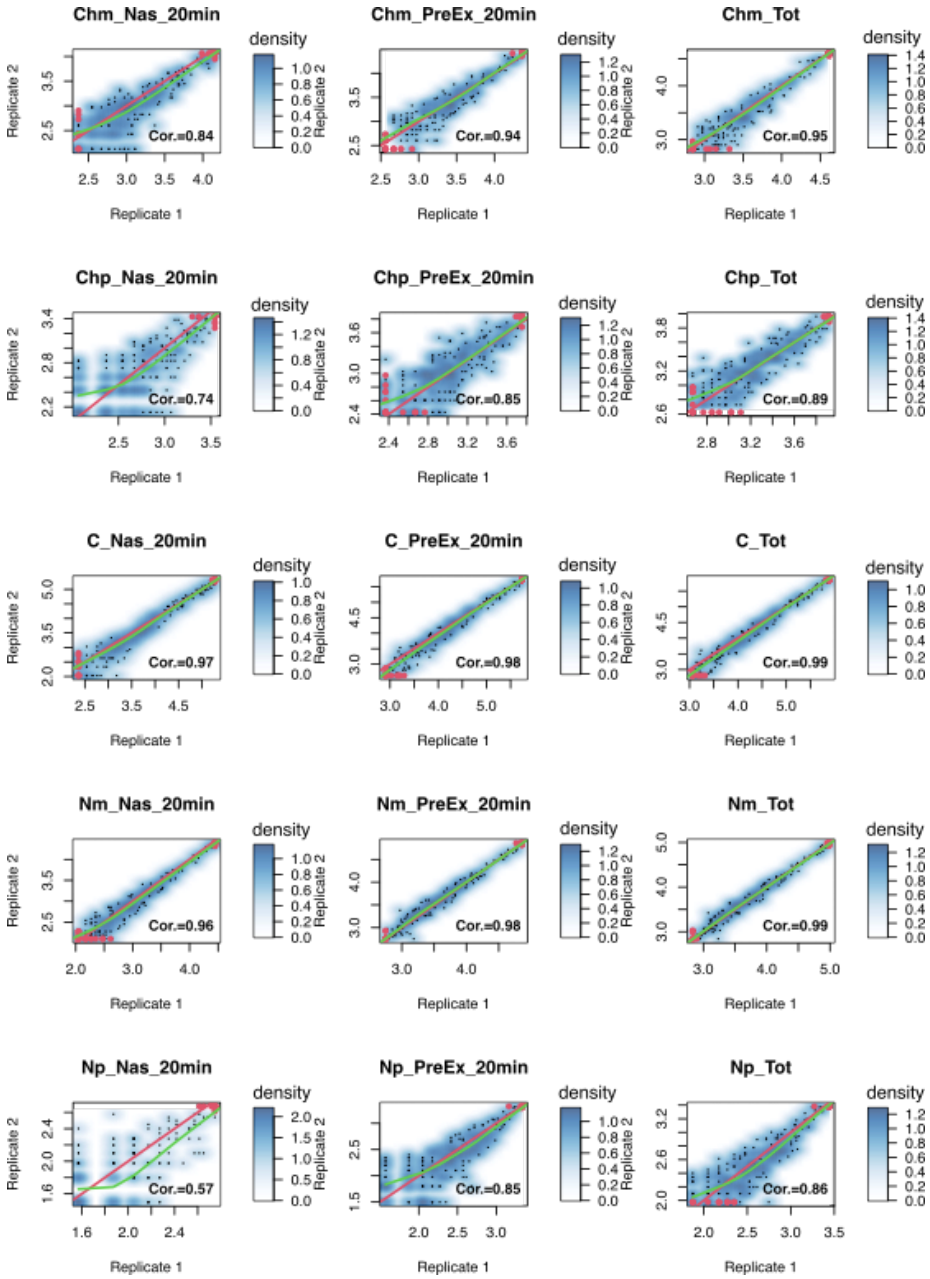
Supplementary Figure 25: **Validation of genes not variable in export in response to the Leptomycin B treatment.** Nuclear vs cytoplasmic RNA expression for the indicated genes in untreated cells (purple) after 16 hours of Leptomycin B treatment (light blue). Expression levels are relative to *RPLPO*, 3 technical replicates and their mean values are reported. Source data are provided as a Source Data file.



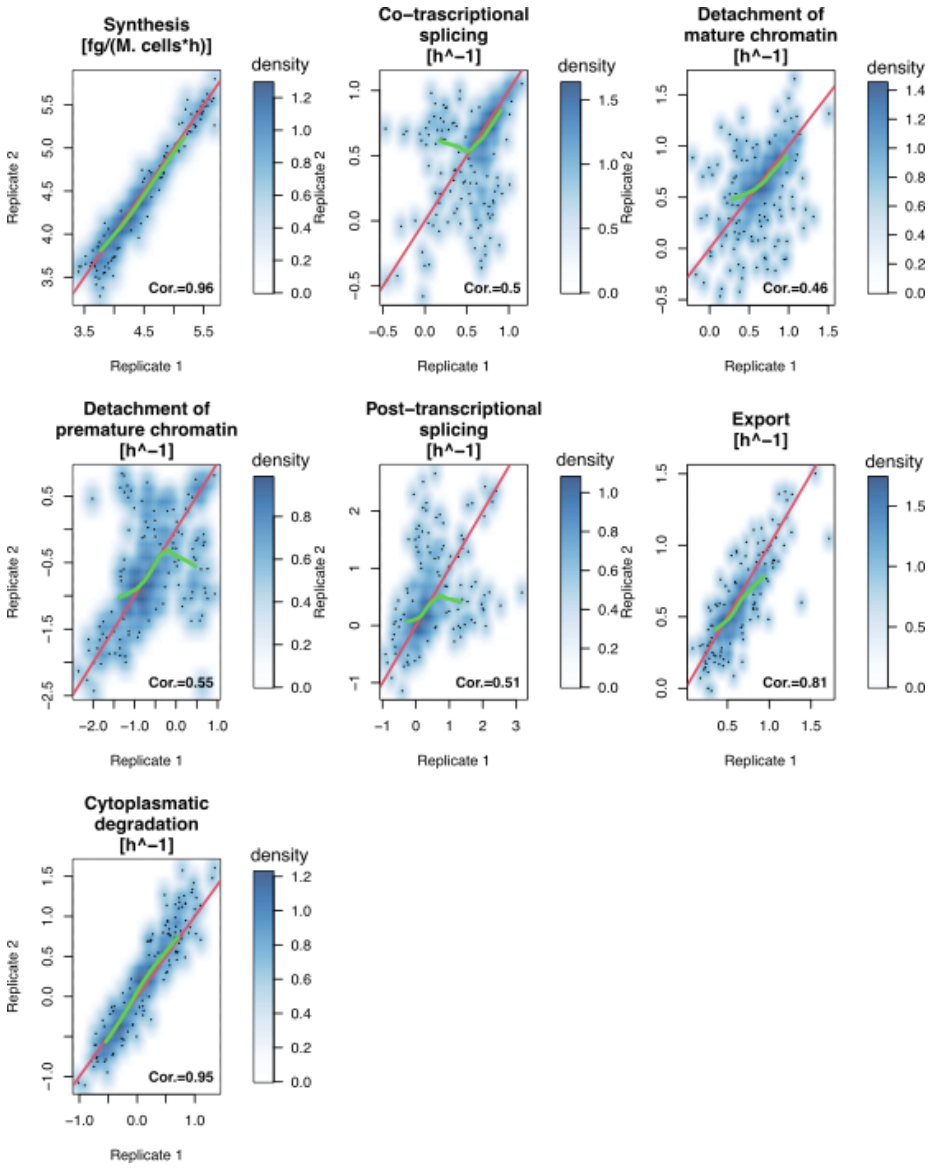
Supplementary Figure 26: **Sequence features in genes clusters for Leptomycin B treated SUM159 cells.** Sequence features characterization of the clusters identified according to gene expression levels and inferred kinetic rates Log2 Fold Changes in Leptomycin B treated SUM159 cells, as reported in Figure 4F. The horizontal line represents the median value, the box edges represent the 25th (Q1) and 75th (Q3) percentiles, and the whiskers show the range of data excluding outliers (observations lower that Q1 - 1.5 * interquartile range or larger than Q3 + 1.5 * interquartile range). Source data are provided as a Source Data file.



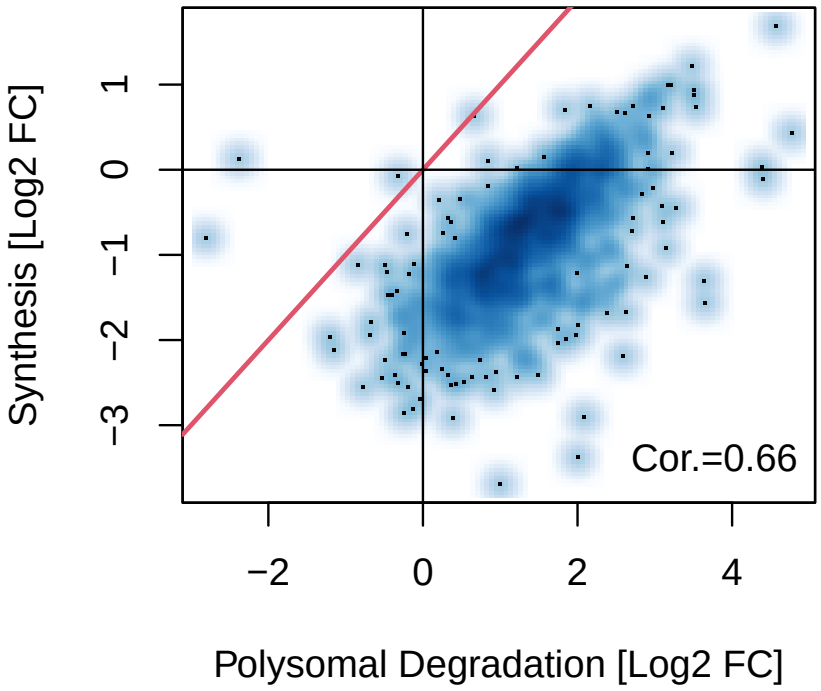
Supplementary Figure 27: **Model lacking association to polysomes.** The nuclear and cytoplasmic steps of the RNA life cycle, with the corresponding kinetic rates (squared brackets), for a simplified model lacking association to polysomes and polysomal degradation. Coloured ellipses represent cellular fractions. Mathematical formulation in terms of Ordinary Differential Equations of the model is included. Ch, N, C represent chromatin associated, nucleoplasmatic, cytoplasmatic and polysomal RNA, respectively, p and m further specifying the premature or mature form.



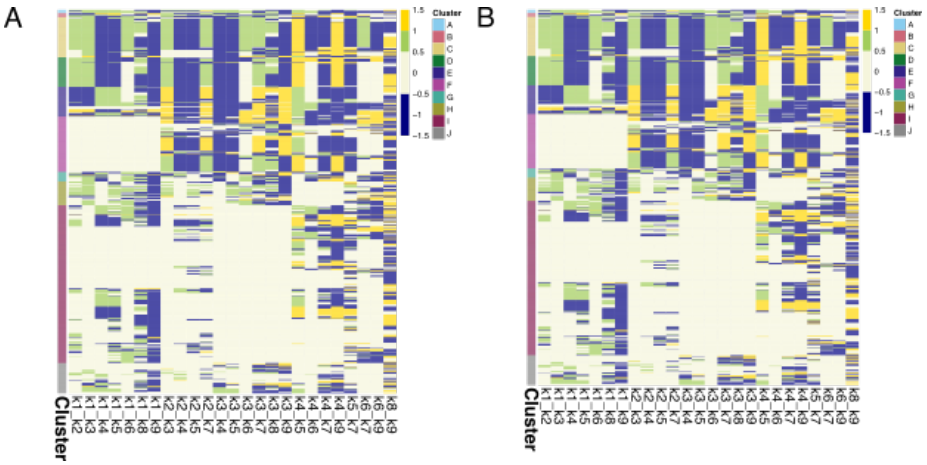
Supplementary Figure 28: **Harringtonine treated SUM159 cells reproducibility – Expression Levels.** Smooth density scatterplots between the experimental expression levels profiled in 2 biological replicates. For each scatterplot, we report the identity line (red), the loess line (green), and the Spearman correlation coefficient (Cor.). Red dots represent saturated points (i.e. data points in the bottom and top 5% of the distribution). Ch, N, C and P refer to chromatin associated, nucleoplasmatic, cytoplasmatic and polysomal RNA, respectively, p and m further specifying the premature or mature form. Source data are provided as a Source Data file.



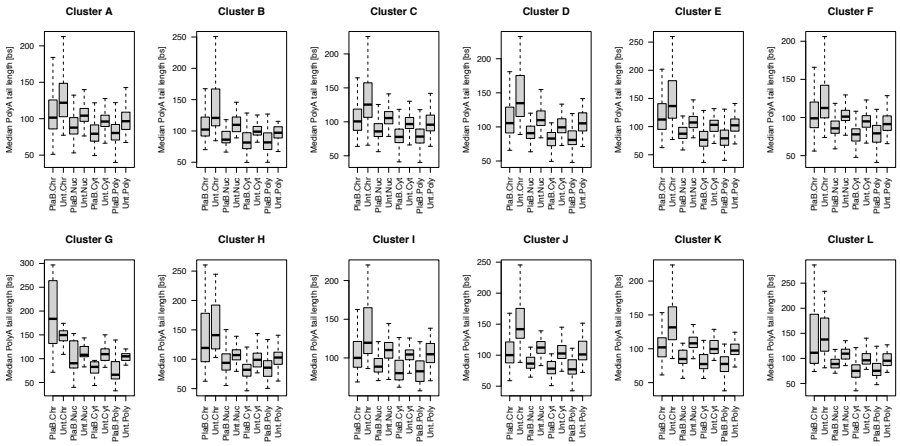
Supplementary Figure 29: **Harringtonine treated SUM159 cells reproducibility – Kinetic rates.** Smooth density scatterplots between the inferred kinetic rates profiled in 2 biological replicates. For each scatterplot, we report the identity line (red), the loess line (green), and the Spearman correlation coefficient (Cor.). Source data are provided as a Source Data file.



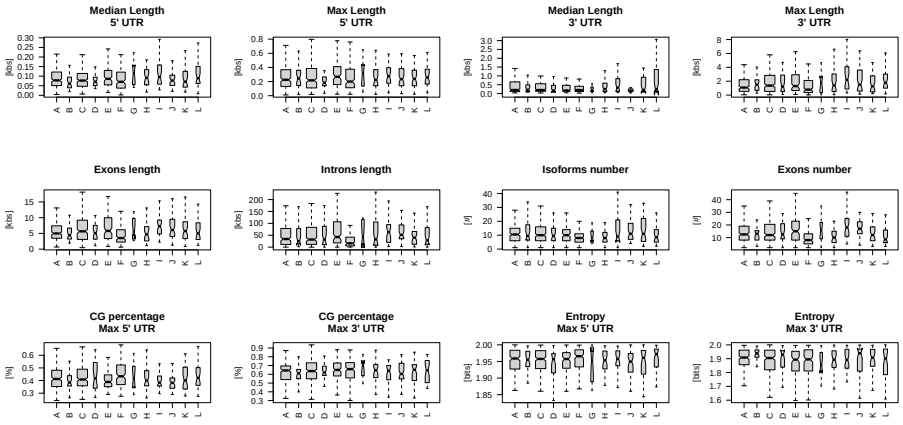
Supplementary Figure 30: **Synthesis and Polysomal degradation modulations in response to Pladienolide B.** Smooth density scatterplot between the log_2 Fold Changes of synthesis (y-axis) and polysomal degradation (x-axis) rates in Pladienolide B treated SUM159 cells compared to the untreated condition. Spearman correlation coefficient (Cor.) is reported. Source data are provided as a Source Data file.



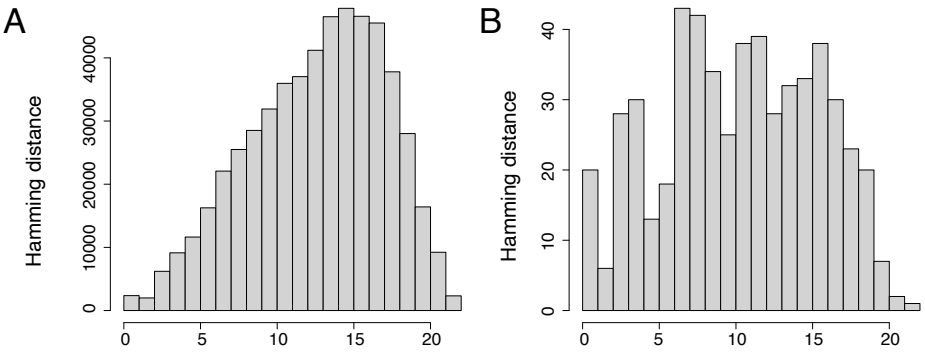
Supplementary Figure 31: **Couplings heatmaps in Leptomycin B and Harringtonine treated SUM159 cells.**Heatmaps depicting, for Leptomycin B (A) and Harringtonine (B) treated SUM159 cells, the possible coupling mechanisms (see the corresponding networks in Figure 6A) (columns) exploited by each gene (rows). The colour code represents the specific regulations: beige indicates that at least one of the two rates defining a given coupling is not regulated, yellow denotes coordinated up-regulation of the two rates, green denotes coordinated down-regulation of the two rates, and blue indicates opposite regulation of the two rates. A rate is considered modulated given an absolute log_2 fold change larger than 0.5. Source data are provided as a Source Data file.



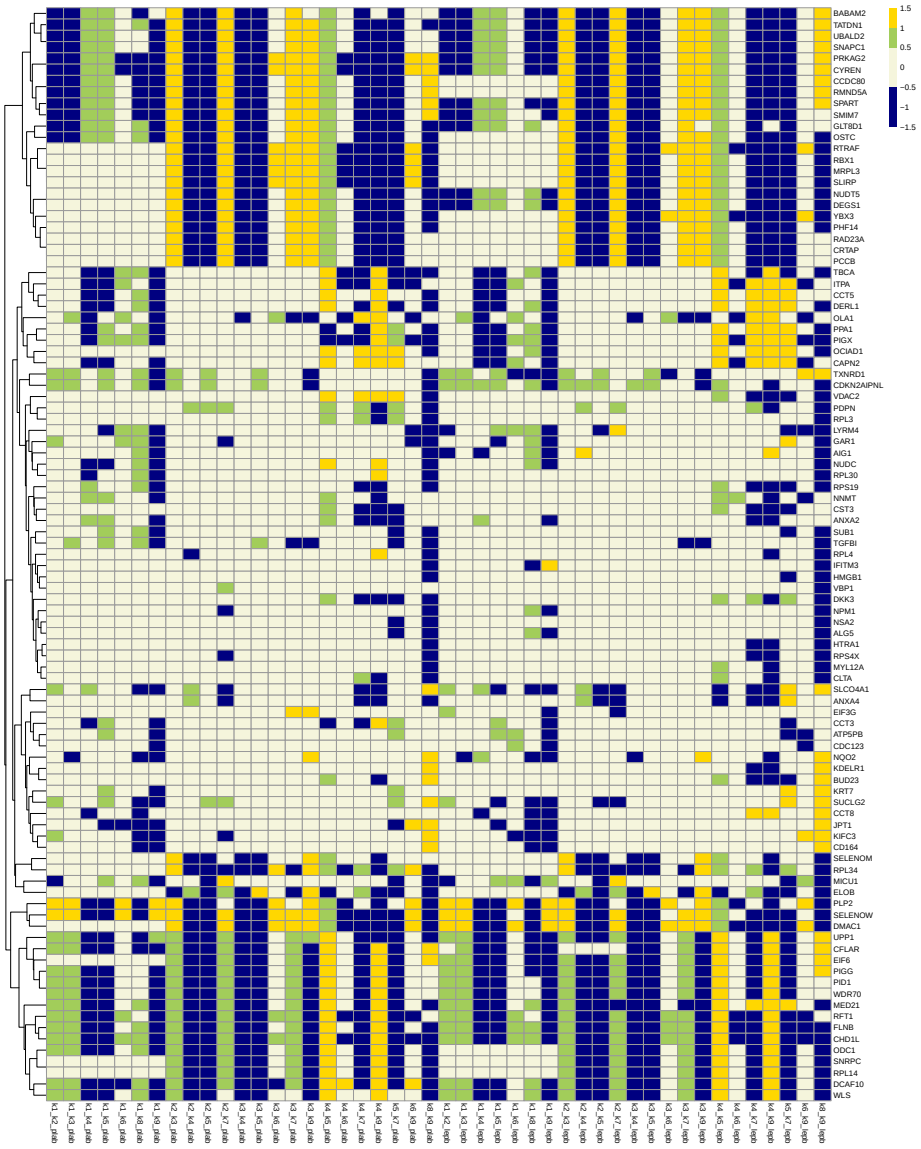
Supplementary Figure 32: **Characterization of polyA tails of gene clusters defined according to coupling mechanisms in response to Pladienolide B.** Gene-level median polyA tail length distributions, over different RNA pools and treatment conditions (untreated=*Unt*, Pladienolide B=*PlaB*), for groups of genes identified by k-means clustering according to the couplings exploited in response to the Pladienolide B treatment as reported in Figure 6C. The horizontal line represents the median value, the box edges represent the 25th (Q1) and 75th (Q3) percentiles, and the whiskers show the range of data excluding outliers (observations lower that Q1 - 1.5 * interquartile range or larger than Q3 + 1.5 * interquartile range). Source data are provided as a Source Data file.



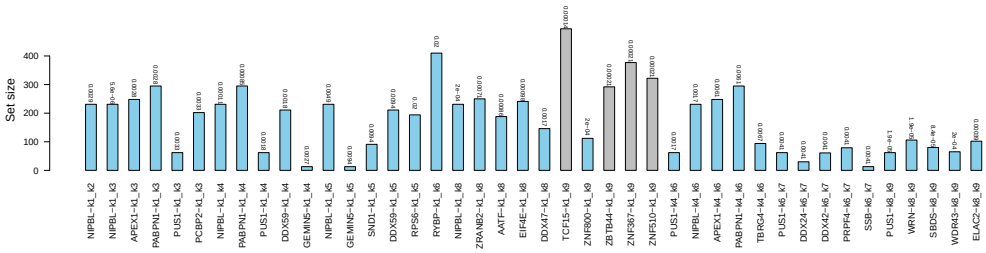
Supplementary Figure 33: **Sequence features for gene clusters defined according to coupling mechanisms in response to Pladienolide B.** Sequence features characterization of the clusters identified according to the couplings exploited in response to the Pladienolide B treatment, as reported in Figure 6C. The horizontal line represents the median value, the box edges represent the 25th (Q1) and 75th (Q3) percentiles, and the whiskers show the range of data excluding outliers (observations lower that Q1 - 1.5 * interquartile range or larger than Q3 + 1.5 * interquartile range). Source data are provided as a Source Data file.



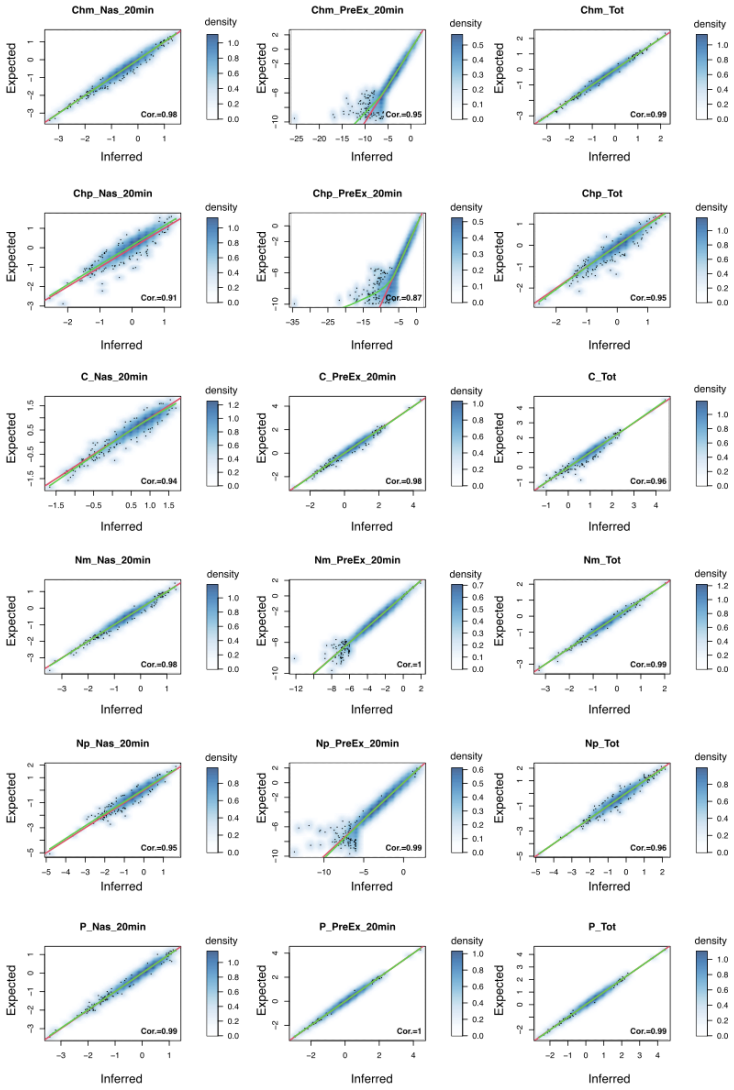
Supplementary Figure 34: **Hamming distance distribution.** Hamming distance distributions for the null model (A) and Pladienolide B treated cells compared to Leptomycin B treatment (B). Source data are provided as a Source Data file.



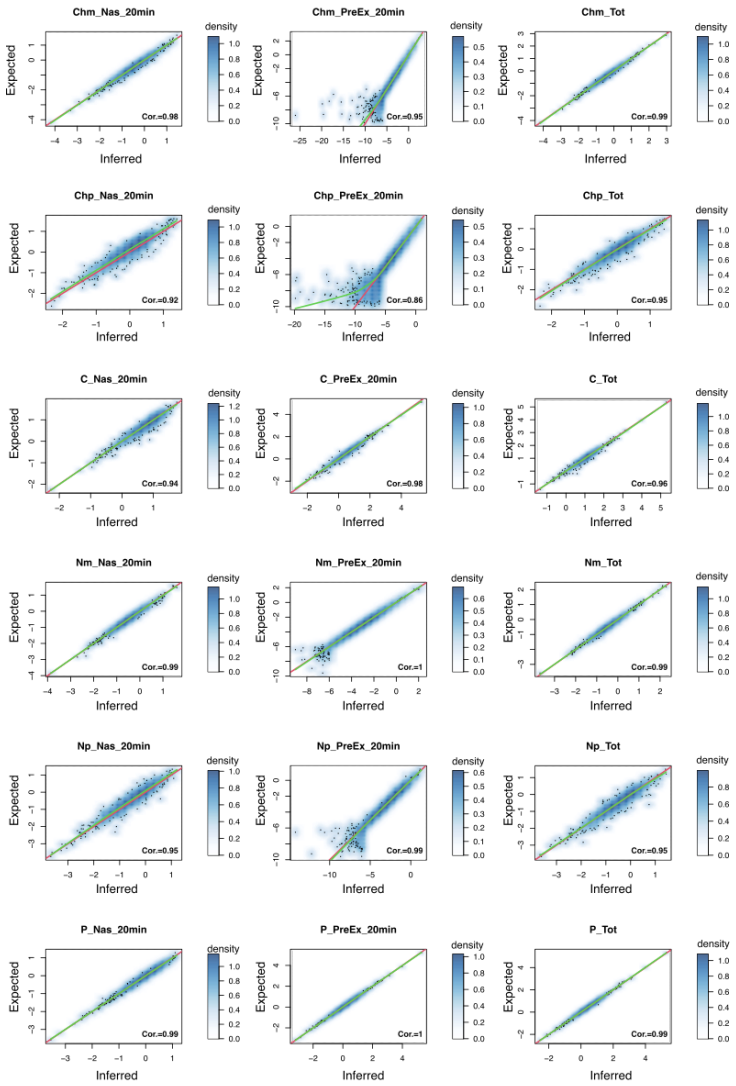
Supplementary Figure 35: **Coupling conservation.** Heatmap of genes with the most conserved couplings across Pladienolide B (*plab*) and Leptomycin B (*lepb*) treatments. The colour code represents the specific regulations: beige indicates that at least one of the two rates defining a given coupling is not regulated, yellow denotes coordinated up-regulation of the two rates, green denotes coordinated down-regulation of the two rates, and blue indicates opposite regulation of the two rates. A rate is considered modulated given an absolute log_2 fold change larger than 0.5. Only shared couplings were reported. Source data are provided as a Source Data file.



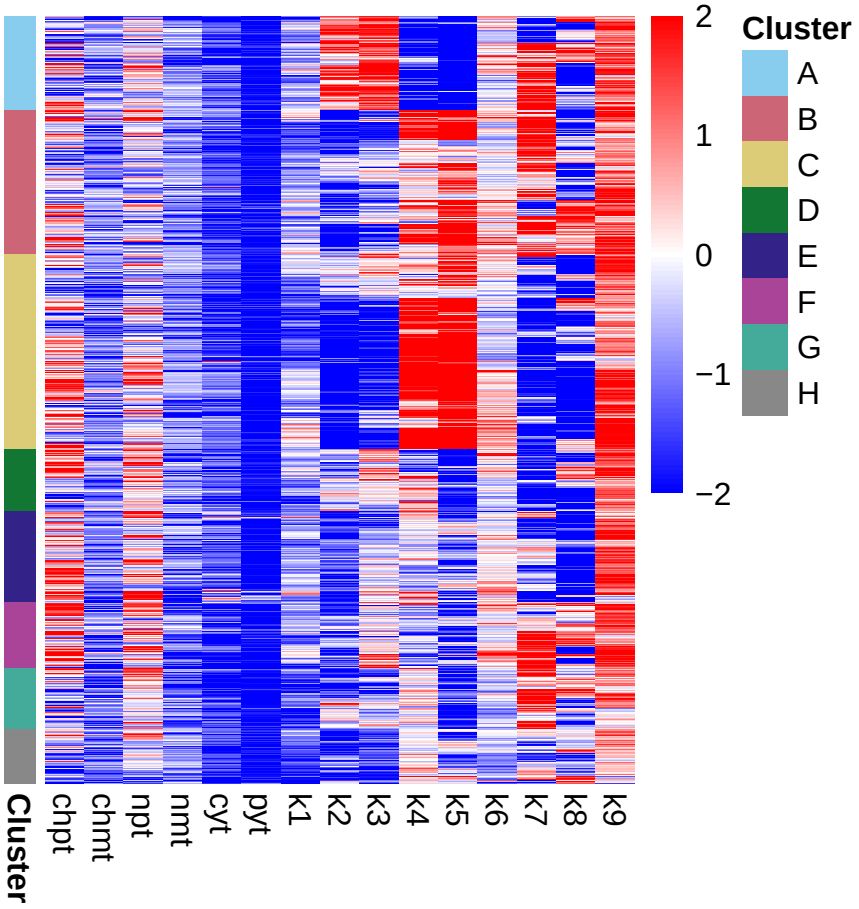
Supplementary Figure 36: **Coupling factors.** Barplot reporting the 5 most significant proteins (RBPs in light blue and TFs in grey), according to a GSEA analysis, for each coupling emerged from the Leptomycin B analysis (see the corresponding network in Figure 6A). For a pair of rates defining a coupling, the GSEA ranking was defined according to the product of rates modulations in Leptomycin B treated versus untreated cells (log_2 fold changes) times the sign of their Spearman correlation. GSEA p-values are reported above each bar. Source data are provided as a Source Data file.



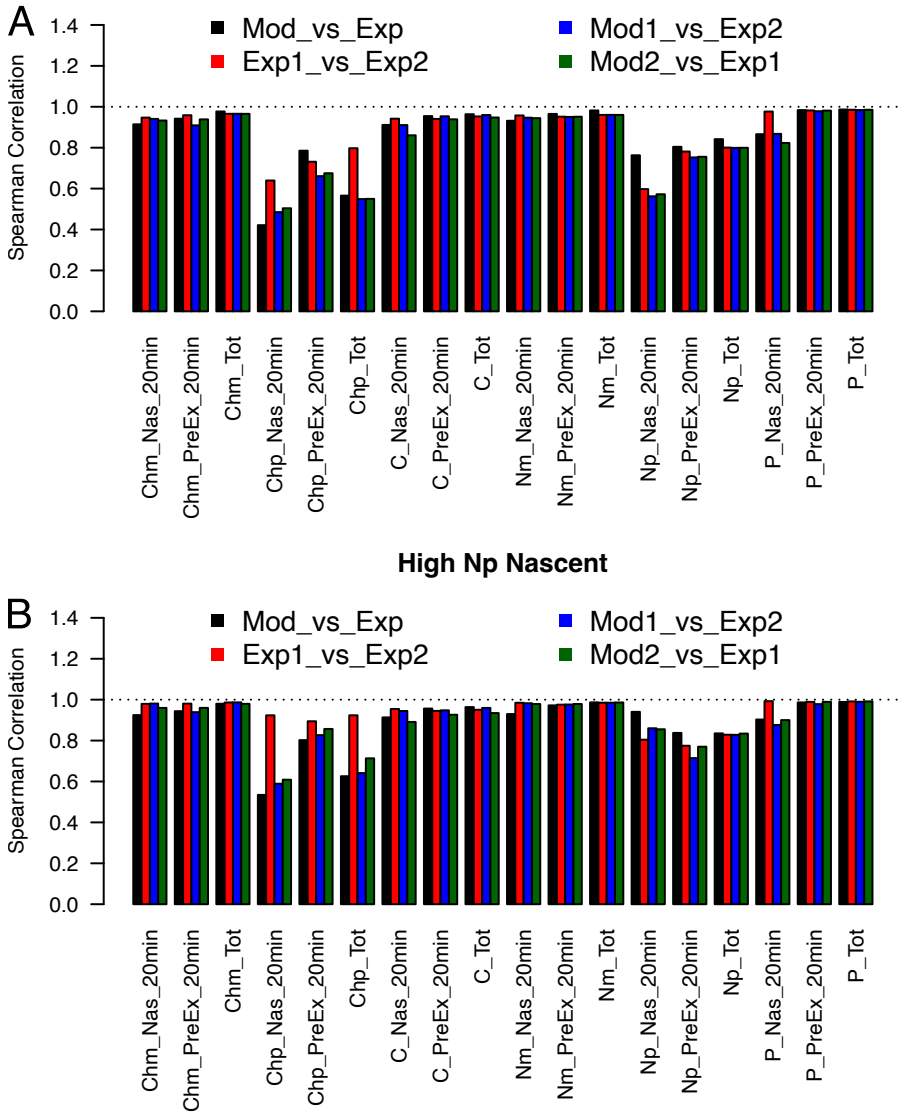
Supplementary Figure 37: **Nucleoplasmic mature RNA degradation model.** Scatterplots between inferred kinetic rates and their expected counterparts for a simulated generated accounting for nucleoplasmic mature RNA degradation (1000 genes, CVs assigned accordingly to each species, 2 replicates, and 1 labelling time of 20 minutes). For each plot, we report the identity line (red), the loess line (green), and the Spearman correlation coefficient (Cor.). Source data are provided as a Source Data file.



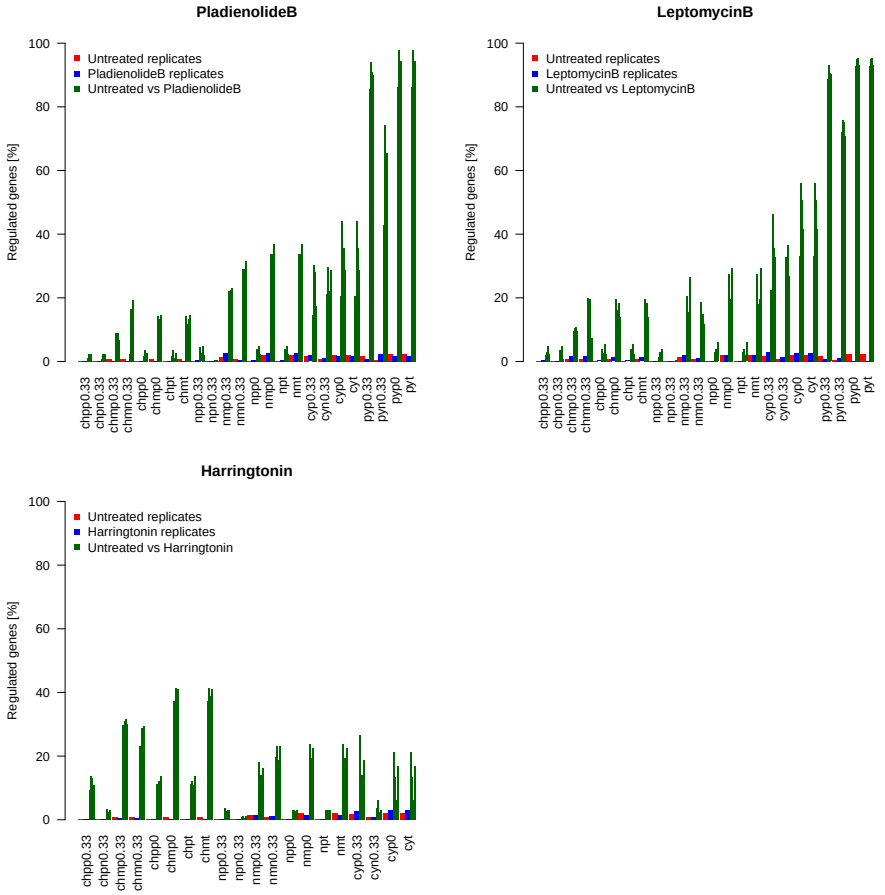
Supplementary Figure 38: **Nucleoplasmic premature RNA degradation model.** Scatterplots between inferred kinetic rates and their expected counterparts for a simulated generated accounting for nucleoplasmic premature RNA degradation (1000 genes, CVs assigned accordingly to each species, 2 replicates, and 1 labelling time of 20 minutes). For each plot, we report the identity line (red), the loess line (green), and the Spearman correlation coefficient (Cor.). Source data are provided as a Source Data file.



Supplementary Figure 39: **Premature RNA export - Differential heatmap for Pladienolide B treated SUM159 cells.** Heatmap of log_2 ratios in changes of abundance of RNA species and kinetic rates compared to untreated cells; values are saturated between -2 and +2. The left colour bar indicates groups of genes identified by k-means clustering according to changes in both abundance of RNA species and kinetic rates. The modulations reported in this figure were obtained analysing both untreated and Pladienolide B treated conditions with a model including the export, cytoplasmic degradation, polysomal association and degradation of premature RNA. Ch, N, C and P refer to chromatin associated, nucleoplasmatic, cytoplasmatic and polysomal RNA, respectively, p and m further specifying the premature or mature form. Source data are provided as a Source Data file.



Supplementary Figure 40: **Overfit quantification.** (A) Expression levels Spearman correlations between: (i) experimental and inferred data following joint replicates analysis (Mod vs Exp - black), (ii) experimental data replicate 1 vs 2 (Exp1 vs Exp2 - red), (iii) inferred data for replicate 1 vs experimental data for replicate 2 (Mod1 vs Exp2 - blue), and (iv) inferred data for replicate 2 vs experimental data for replicate 1 (Mod2 vs Exp1 - green). (B) As (A) for genes with a high level of nascent nucleoplasmic premature RNA (top 10%). Source data are provided as a Source Data file.

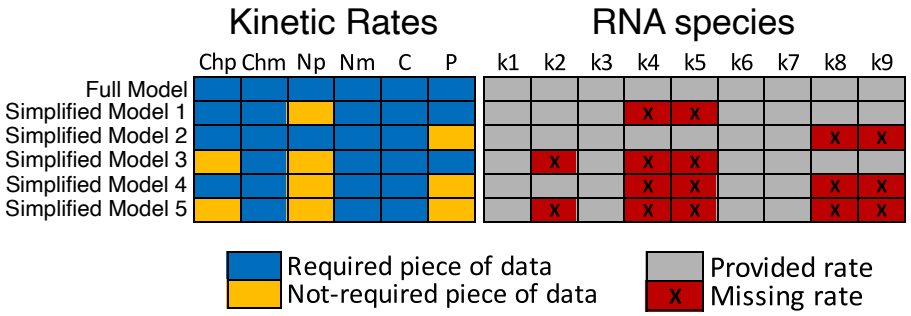


Supplementary Figure 41: **Differential RNA species.** (Green) For each RNA species, percentage of differential genes in response to each treatment with respect to the untreated condition. Each bar corresponds to one of the four possible confronts given two replicates per condition. The same analyses performed on the two replicates of either the treatment or the control conditions are reported in blue and red respectively. Source data are provided as a Source Data file.

2 Alternative simplified models of the RNA life cycle

For each simplified model (scheme in Supplmenetary figure 42), we checked parameters global identifiability, and we validated the goodness of fit through simulated data (1000 genes, CV chosen as describe in Methods section of the main text, 2 replicates, 20 minutes labeling); the resulting correlations between

expected and inferred rates are globally as good as those presented for the complete model in Main Figure 1D (Supplementary figures 43-47).



Supplementary Figure 42: RNA species and kinetic rates for the simplified models (rows) implemented in Nanodynamo, compared to the complete model (1st row); input data required for each model are indicated in blue, and the subset of kinetic rates that is returned is indicated in green.

$$\begin{cases} \frac{dCh_p}{dt} = k_1 - (k_2 + k_4) \cdot Ch_p \\ \frac{dCh_m}{dt} = k_2 \cdot Ch_p - k_3 \cdot Ch_m \\ \frac{dN_p}{dt} = k_4 \cdot Ch_p - k_5 \cdot N_p \\ \frac{dN_m}{dt} = k_3 \cdot Ch_m + k_5 \cdot N_p - k_6 \cdot N_m \\ \frac{dC}{dt} = k_6 \cdot N_m - (k_7 + k_8) \cdot C \\ \frac{dP}{dt} = k_8 \cdot C - k_9 \cdot P \end{cases} \qquad \begin{cases} Ch_p = \frac{k_1}{k_2+k_4} \\ Ch_m = \frac{k_2}{k_3} \cdot \frac{k_1}{k_2+k_4} \\ N_p = \frac{k_4}{k_5} \cdot \frac{k_1}{k_2+k_4} \\ N_m = \frac{k_1}{k_6} \\ C = \frac{k_1}{k_7+k_8} \\ P = \frac{k_8}{k_9} \cdot \frac{k_1}{k_7+k_8} \end{cases} \tag{1}$$

1: System of equation describing the full model (left) with its initial conditions (right).

$$\begin{cases} \frac{dCh_p}{dt} = k_1 - (k_2 + k_4) \cdot Ch_p \\ \frac{dCh_m}{dt} = k_2 \cdot Ch_p - k_3 \cdot Ch_m \\ \frac{dN_p}{dt} = k_4 \cdot Ch_p - k_5 \cdot N_p \\ \frac{dN_m}{dt} = k_3 \cdot Ch_m + k_5 \cdot N_p - k_6 \cdot N_m \\ \frac{dC}{dt} = k_6 \cdot N_m - k_7 \cdot C \end{cases} \qquad \begin{cases} Ch_p = \frac{k_1}{k_2+k_4} \\ Ch_m = \frac{k_2}{k_3} \cdot \frac{k_1}{k_2+k_4} \\ N_p = \frac{k_4}{k_5} \cdot \frac{k_1}{k_2+k_4} \\ N_m = \frac{k_1}{k_6} \\ C = \frac{k_1}{k_7} \end{cases} \tag{2}$$

2: System of equation describing the model lacking polysomal RNA (left) with its initial conditions (right).

$$\begin{cases} \frac{dCh_p}{dt} = k_1 - k_2 \cdot Ch_p \\ \frac{dCh_m}{dt} = k_2 \cdot Ch_p - k_3 \cdot Ch_m \\ \frac{dN_m}{dt} = k_3 \cdot Ch_m - k_6 \cdot N_m \\ \frac{dC}{dt} = k_6 \cdot N_m - (k_7 + k_8) \cdot C \\ \frac{dP}{dt} = k_8 \cdot C - k_9 \cdot P \end{cases} \quad \begin{cases} Ch_p = \frac{k_1}{k_2} \\ Ch_m = \frac{k_1}{k_3} \\ N_m = \frac{k_1}{k_6} \\ C = \frac{k_1}{k_7 + k_8} \\ P = \frac{k_8}{k_9} \cdot \frac{k_1}{k_7 + k_8} \end{cases} \tag{3}$$

3: System of equation describing the model lacking nucleoplasmic premature RNA (left) with its initial conditions (right).

$$\begin{cases} \frac{dCh_p}{dt} = k_1 - k_2 \cdot Ch_p \\ \frac{dCh_m}{dt} = k_2 \cdot Ch_p - k_3 \cdot Ch_m \\ \frac{dN_m}{dt} = k_3 \cdot Ch_m - k_6 \cdot N_m \\ \frac{dC}{dt} = k_6 \cdot N_m - k_7 \cdot C \end{cases} \quad \begin{cases} Ch_p = \frac{k_1}{k_2} \\ Ch_m = \frac{k_1}{k_3} \\ N_m = \frac{k_1}{k_6} \\ C = \frac{k_1}{k_7} \end{cases} \tag{4}$$

4: System of equation describing the model lacking both nucleoplasmic premature RNA and polysomal RNA (left) with its initial conditions (right).

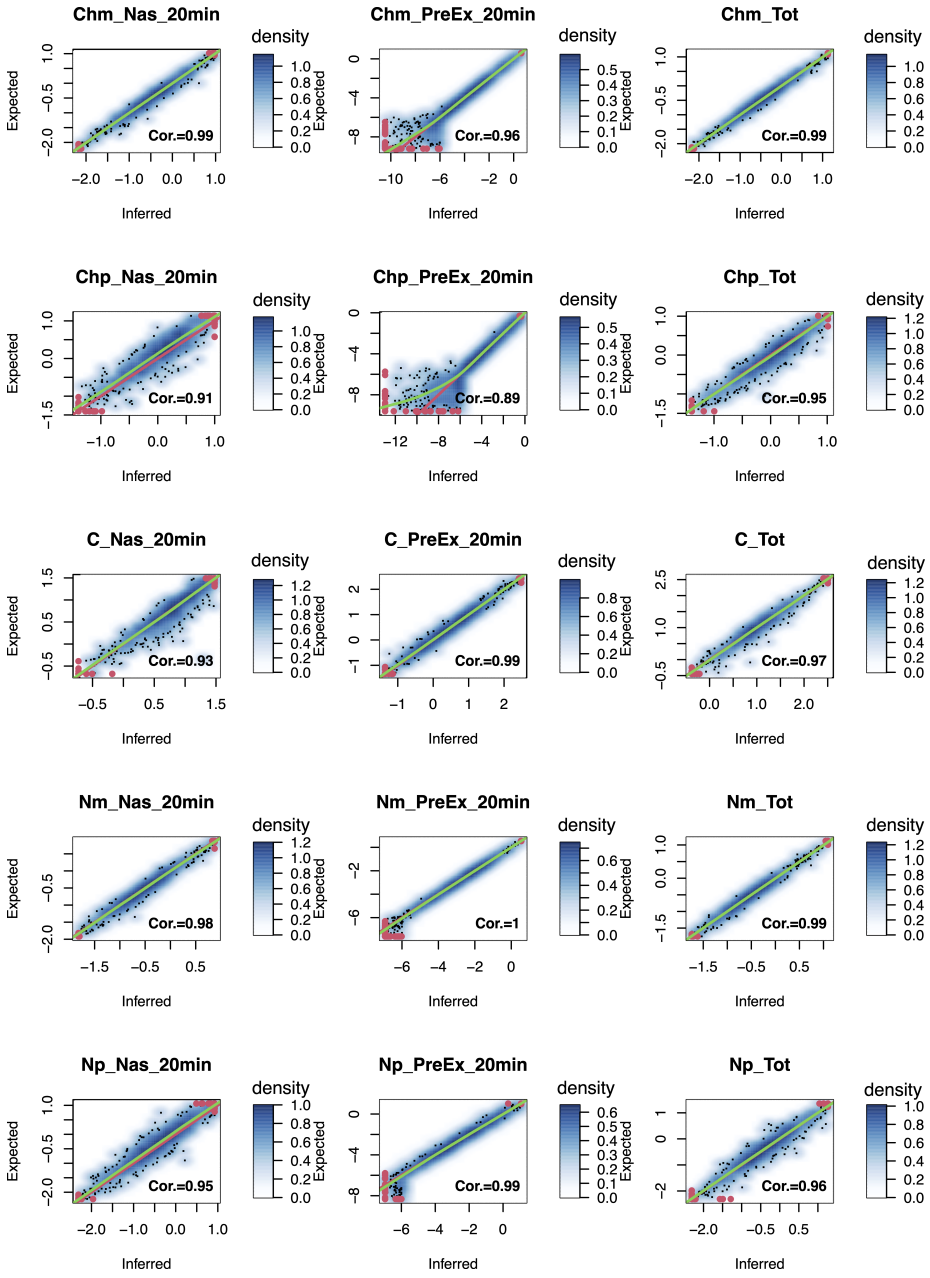
$$\begin{cases} \frac{dCh_m}{dt} = k_1 - k_3 \cdot Ch_m \\ \frac{dN_m}{dt} = k_3 \cdot Ch_m - k_6 \cdot N_m \\ \frac{dC}{dt} = k_6 \cdot N_m - (k_7 + k_8) \cdot C \\ \frac{dP}{dt} = k_8 \cdot C - k_9 \cdot P \end{cases} \quad \begin{cases} Ch_m = \frac{k_1}{k_3} \\ N_m = \frac{k_1}{k_6} \\ C = \frac{k_1}{k_7 + k_8} \\ P = \frac{k_8}{k_9} \cdot \frac{k_1}{k_7 + k_8} \end{cases} \tag{5}$$

5: System of equation describing the model lacking premature RNA (left) with its initial conditions (right).

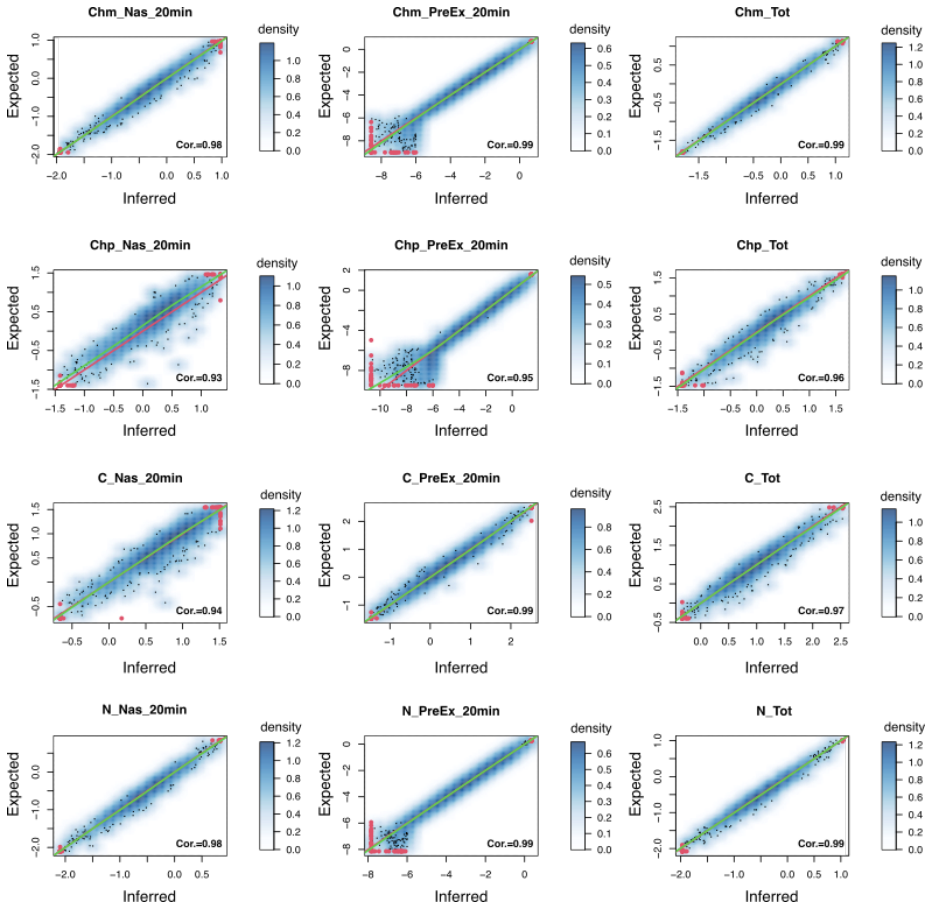
$$\begin{cases} \frac{dCh_m}{dt} = k_1 - k_3 \cdot Ch_m \\ \frac{dN_m}{dt} = k_3 \cdot Ch_m - k_6 \cdot N_m \\ \frac{dC}{dt} = k_6 \cdot N_m - k_7 \cdot C \end{cases} \quad \begin{cases} Ch_m = \frac{k_1}{k_3} \\ N_m = \frac{k_1}{k_6} \\ C = \frac{k_1}{k_7} \end{cases} \tag{6}$$

6: System of equation describing the model lacking premature RNA and polysomal RNA (left) with its initial conditions (right).

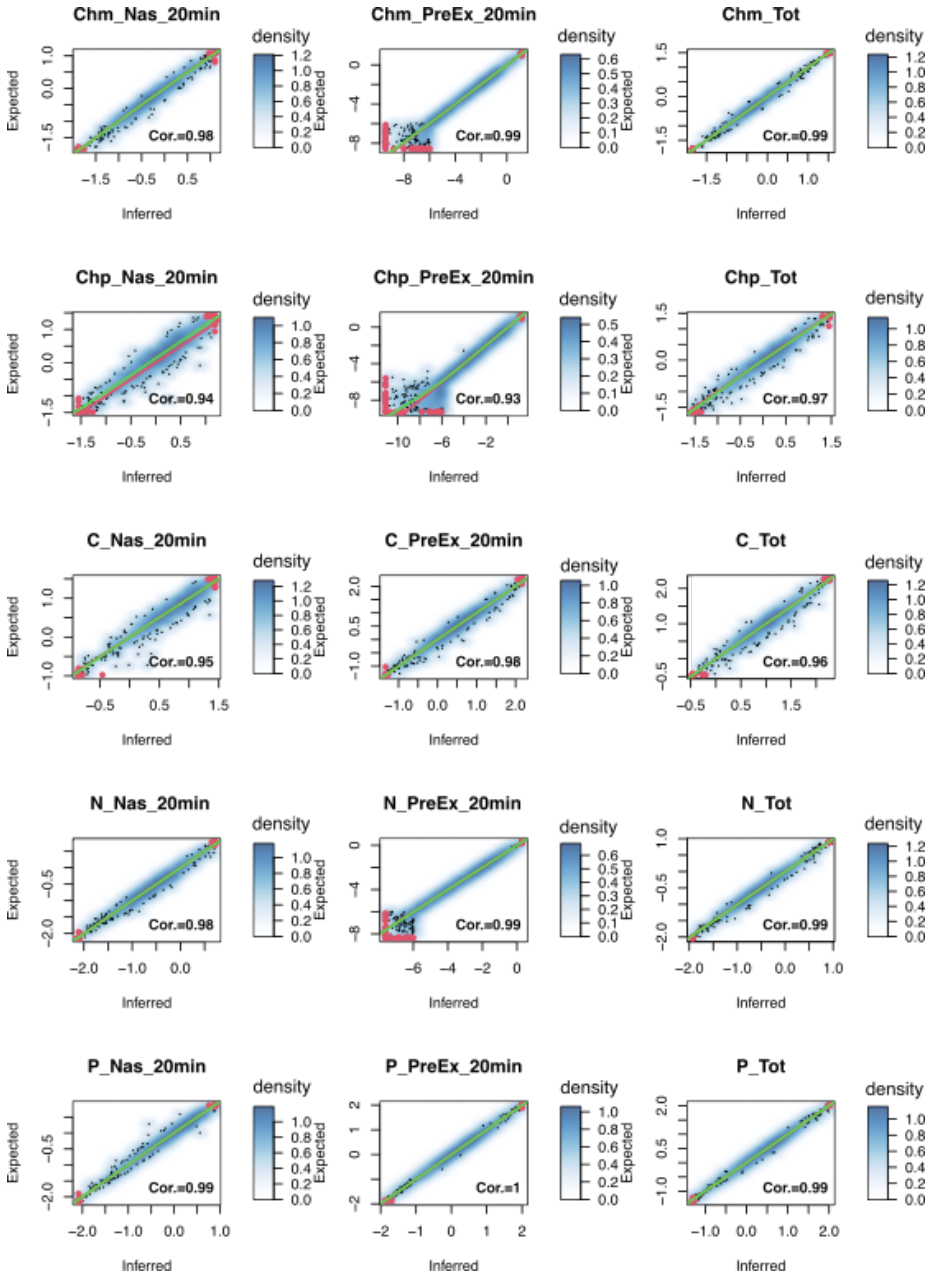
Then, we applied the simpler frameworks to profile the RNA life cycle kinetics rates for all those genes we could not process with the full model in the SUM159 untreated dataset. Specifically, we recovered: 11 genes just missing polysomal RNA (system 2), 3126 genes without nucleoplasmic premature RNA (22 of them were also missing polysomal RNA - systems 3 and 4 respectively), and 5297 genes lacking any intronic signal (59 of them were also missing polysomal RNA - systems 5 and 6 respectively). The small set of genes just missing polysomal RNA is characterised by low transcriptional activity (median synthesis rate of 3194 $fg \cdot Mcells^{-1}h^{-1}$) and, in agreement with our observations from the complete model, this is accompanied by a fast post-transcriptional processing (median k_4 and k_5 of 1.17 h^{-1} and 4.59 h^{-1} respectively). These transcriptional units also show high degradation rates (median of 5.4 h^{-1}) suggesting that they might be actively repressed (Supplementary figure 48). The set of genes lacking premature nucleoplasmic RNA largely recapitulates what we observed for the full model in terms of rates magnitude. However, the synthesis rate distribution lacks the fastest component ($k_1 > 1e5$), , the cytoplasmic decay rate is lower (median of 0.12 h^{-1}, andwhile the co-transcriptional processing rates do not show the characteristic bimodality; this is in agreement with the observed exclusivity of the two processing pathways (Figure Supplementary figure 49). Again, transcriptional units lacking also polysomal RNA are inefficiently transcribed (2085 $fg \cdot Mcells^{-1}h^{-1}$) and unstable (median of 5.41 h^{-1} - Supplementary figure 50). Genes without any premature signal confirm rates magnitudes except for the synthesis , cytoplasmic degradation and polysomal association rates which are slower compared to the full model (5613 $fg \cdot Mcells^{-1}h^{-1}$, 0.38 h^{-1} and 0.42 h^{-1} respectively). Interestingly, the detachment step linking chromatin RNA synthesis to nucleoplasmic RNA has a median rate of 6.56 h^{-1} in agreement with the co-transcriptional processing rates of the full model (Supplementary figure 51). This is even more evident for the transcriptional units processed with the simplest model which are also characterized by the slowest synthesis rates (1272 $fg \cdot Mcells^{-1}h^{-1}$ - Supplementary figure 52).



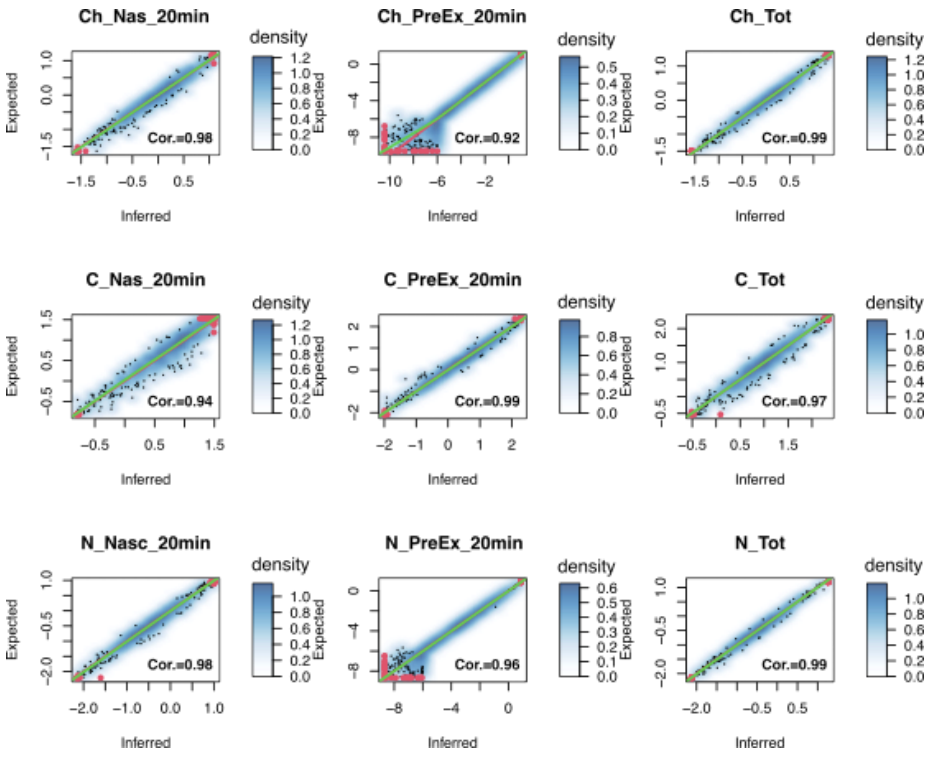
Supplementary figure 43: **Simplified model 2** describing genes lacking polysomal RNA. Scatterplots between modelled RNA expression levels and their expected counterparts for a simulated dataset of 1000 genes generated with CVs assigned accordingly to each species, 2 replicates, 1 labelling time of 20'. For each plot, we report the identity line (red), the loess line (green), and the Spearman correlation coefficient (Cor.). Red dots represent saturated points (i.e. data points in the bottom and top 2.5% of the distribution). Source data are provided as a Source Data file.



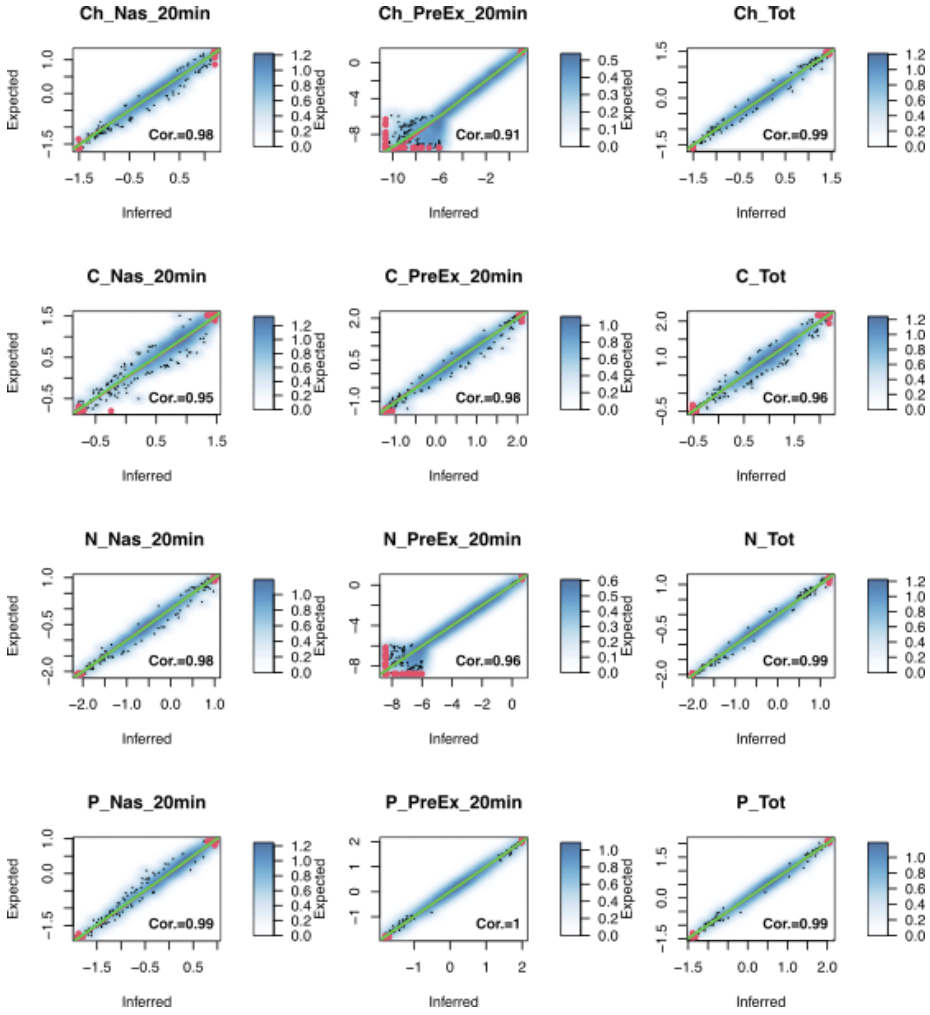
Supplementary figure 44: **Simplified model 4** describing genes lacking premature nucleoplasmic RNA and polysomal RNA. Scatterplots between modelled RNA expression levels and their expected counterparts for a simulated dataset of 1000 genes generated with CVs assigned accordingly to each species, 2 replicates, 1 labelling time of 20'. For each plot, we report the identity line (red), the loess line (green), and the Spearman correlation coefficient (Cor.). Red dots represent saturated points i.e. data points in the bottom and top 2.5% of the distribution). Source data are provided as a Source Data file.



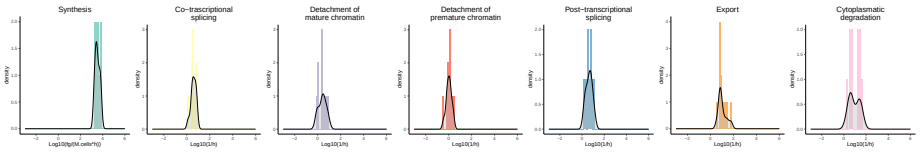
Supplementary figure 45: **Simplified model 1** describing genes lacking premature nucleoplasmic RNA. Scatterplots between modelled RNA expression levels and their expected counterparts for a simulated dataset of 1000 genes generated with CVs assigned accordingly to each species, 2 replicates, 1 labelling time of 20'. For each plot, we report the identity line (red), the loess line (green), and the Spearman correlation coefficient (Cor.). Red dots represent saturated points (i.e. data points in the bottom and top 2.5% of the distribution). Source data are provided as a Source Data file.



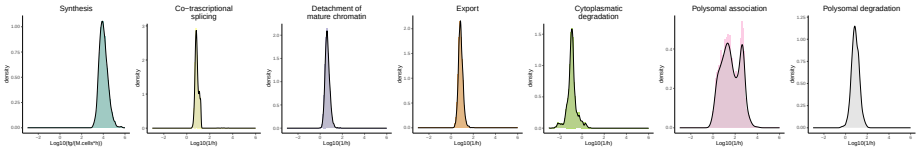
Supplementary figure 46: **Simplified model 5** describing genes lacking premature RNA and polysomal RNA. Scatterplots between modelled RNA expression levels and their expected counterparts for a simulated dataset of 1000 genes generated with CVs assigned accordingly to each species, 2 replicates, 1 labelling time of 20'. For each plot, we report the identity line (red), the loess line (green), and the Spearman correlation coefficient (Cor.). Red dots represent saturated points (i.e. data points in the bottom and top 2.5% of the distribution). Source data are provided as a Source Data file.



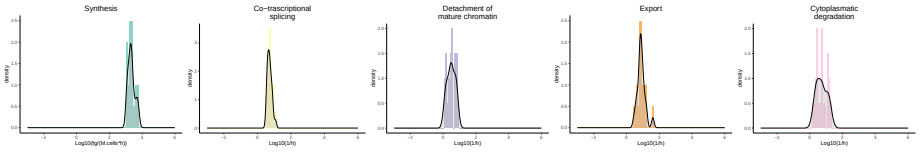
Supplementary figure 47: **Simplified model 3** describing genes lacking premature RNA. Scatterplots between modelled RNA expression levels and their expected counterparts for a simulated dataset of 1000 genes generated with CVs assigned accordingly to each species, 2 replicates, 1 labelling time of 20'. For each plot, we report the identity line (red), the loess line (green), and the Spearman correlation coefficient (Cor.). Red dots represent saturated points (i.e. data points in the bottom and top 2.5% of the distribution).



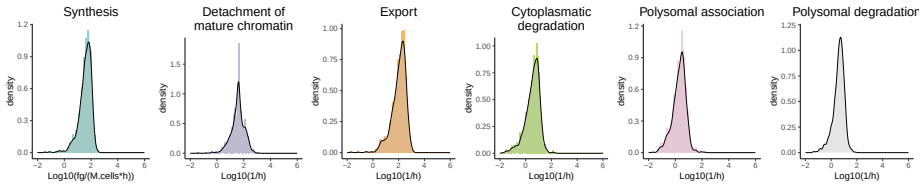
Supplementary figure 48: **Simplified model 2**. Rates distributions of genes lacking polysomal RNA. Source data are provided as a Source Data file.



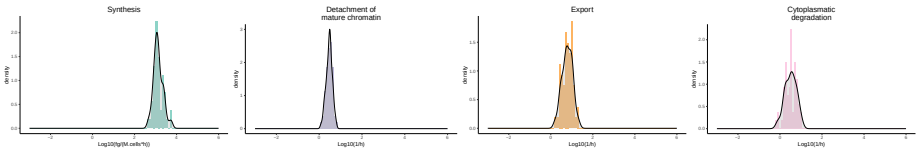
Supplementary figure 49: **Simplified model 1**. Rates distributions of genes lacking premature nucleoplasmic RNA. Source data are provided as a Source Data file.



Supplementary figure 50: **Simplified model 4**. Rates distributions of genes lacking premature nucleoplasmic RNA and polysomal RNA. Source data are provided as a Source Data file.

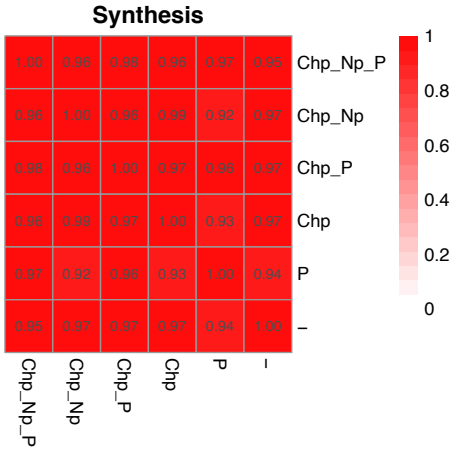


Supplementary figure 51: **Simplified model 3**. Rates distributions of genes lacking premature RNA. Source data are provided as a Source Data file.

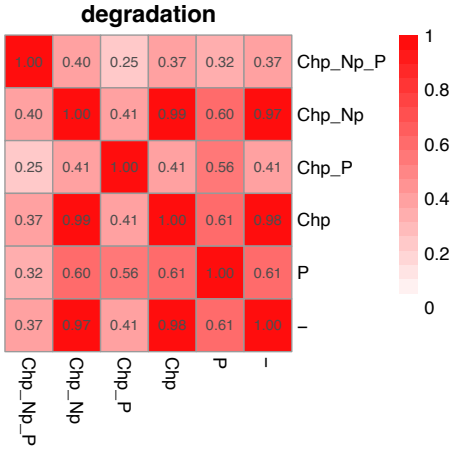


Supplementary figure 52: **Simplified model 5**. Rates distributions of genes lacking premature RNA and polysomal RNA. Source data are provided as a Source Data file.

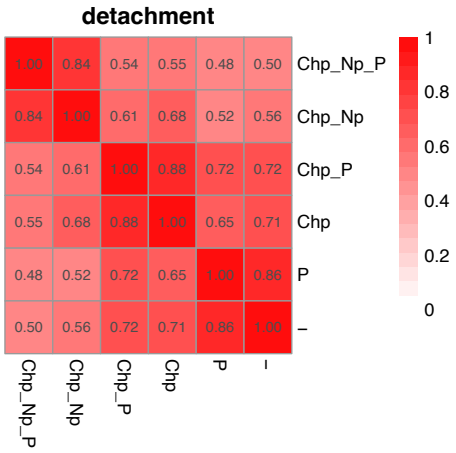
Finally, we processed with each simplified framework the set of 1914 genes previously analysed with the complete model and we performed a correlative analysis for the four rates shared by all the configurations: synthesis, chromatin mature RNA detachment, export, and cytoplasmic degradation. For RNA synthesis the agreement is remarkable with minimum Spearman correlation of 0.92 (Supplementary figure 53). Regarding the cytoplasmic degradation rate, the full model shows a good agreement with its counterparts lacking polysomal RNA while the models with polysomal are remarkably correlated to each other (Supplementary figure 54). For the rate of chromatin mature RNA detachment (Supplementary figure 55), we see minimum impact due to the presence of polysomal RNA, while the exclusion of nucleoplasmic premature RNA seems to significantly alter the rate estimation. Finally, the estimation of the export rate is mainly impacted by the lack of both premature RNA species (Supplementary figure 56).



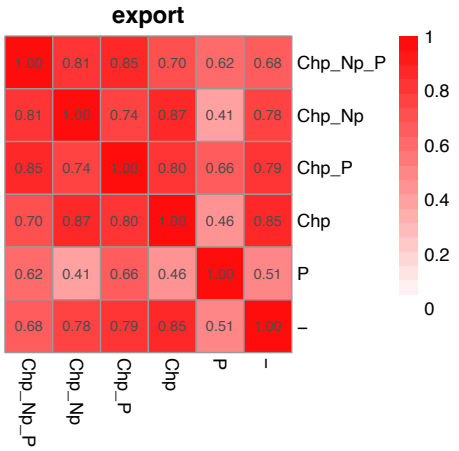
Supplementary figure 53: Synthesis rate spearman correlation for a set of 879 genes analysed with models of variable complexity: complete model (Chp Np P), model lacking polysomal RNA (Chp Np), model lacking nucleoplasmic premature RNA (Chp P), model lacking polysomal and nucleoplasmic premature RNA (Chp), model lacking chromatin and nucleoplasmic premature RNA (P), model lacking polysomal and premature RNA (-). Ch, N and P refer to chromatin associated, nucleoplasmatic and polysomal RNA, respectively, p further specifying the premature form. Source data are provided as a Source Data file.



Supplementary figure 54: Degradation rate spearman correlation for a set of 879 genes analysed with models of variable complexity: complete model (Chp Np P), model lacking polysomal RNA (Chp Np), model lacking nucleoplasmic premature RNA (Chp P), model lacking polysomal and nucleoplasmic premature RNA (Chp), model lacking chromatin and nucleoplasmic premature RNA (P), model lacking polysomal and premature RNA (-). Ch, N and P refer to chromatin associated, nucleoplasmatic and polysomal RNA, respectively, p further specifying the premature form. Source data are provided as a Source Data file.



Supplementary figure 55: Rate of mature chromatin RNA detachment spearman correlation for a set of 879 genes analysed with models of variable complexity: complete model (Chp Np P), model lacking polysomal RNA (Chp Np), model lacking nucleoplasmic premature RNA (Chp P), model lacking polysomal and nucleoplasmic premature RNA (Chp), model lacking chromatin and nucleoplasmic premature RNA (P), model lacking polysomal and premature RNA (-). Ch, N and P refer to chromatin associated, nucleoplasmatic and polysomal RNA, respectively, p further specifying the premature form. Source data are provided as a Source Data file.



Supplementary figure 56: Export rate spearman correlation for a set of 879 genes analysed with models of variable complexity: complete model (Chp Np P), model lacking polysomal RNA (Chp Np), model lacking nucleoplasmic premature RNA (Chp P), model lacking polysomal and nucleoplasmic premature RNA (Chp), model lacking chromatin and nucleoplasmic premature RNA (P), model lacking polysomal and premature RNA (-). Ch, N and P refer to chromatin associated, nucleoplasmatic and polysomal RNA, respectively, p further specifying the premature form. Source data are provided as a Source Data file.

3 Models with nuclear decay

Extensions of the model presented in Main Figure 1 (1) including degradation of nucleoplasmic premature (7) or mature (8) RNA.

$$\begin{cases} \frac{dCh_p}{dt} = k_1 - (k_2 + k_4) \cdot Ch_p \\ \frac{dCh_m}{dt} = k_2 \cdot Ch_p - k_3 \cdot Ch_m \\ \frac{dN_p}{dt} = k_4 \cdot Ch_p - (k_5 + k_{11}) \cdot N_p \\ \frac{dN_m}{dt} = k_3 \cdot Ch_m + k_5 \cdot N_p - k_6 \cdot N_m \\ \frac{dC}{dt} = k_6 \cdot N_m - (k_7 + k_8) \cdot C \\ \frac{dP}{dt} = k_8 \cdot C - k_9 \cdot P \end{cases} \quad \begin{cases} Ch_p = \frac{k_1}{k_2+k_4} \\ Ch_m = \frac{k_2}{k_3} \cdot \frac{k_1}{k_2+k_4} \\ N_p = \frac{k_4}{k_5} \cdot \frac{k_1}{k_2+k_4} \\ N_m = \frac{k_1}{k_6 \cdot (k_2+k_4)} \cdot \left(k_2 + \frac{k_4 \cdot k_5}{k_5+k_{11}}\right) \\ C = \frac{k_1}{(k_2+k_4) \cdot (k_7+k_8)} \cdot \left(k_2 + \frac{k_4 \cdot k_5}{k_5+k_{11}}\right) \\ P = \frac{k_7}{k_8} \cdot \frac{k_1}{(k_2+k_4) \cdot (k_7+k_8)} \cdot \left(k_2 + \frac{k_4 \cdot k_5}{k_5+k_{11}}\right) \end{cases} \tag{7}$$

7: System of equation describing the full model including nucleoplasmic premature RNA decay (left) with its initial conditions.

$$\begin{cases} \frac{dCh_p}{dt} = k_1 - (k_2 + k_4) \cdot Ch_p \\ \frac{dCh_m}{dt} = k_2 \cdot Ch_p - k_3 \cdot Ch_m \\ \frac{dN_p}{dt} = k_4 \cdot Ch_p - k_5 \cdot N_p \\ \frac{dN_m}{dt} = k_3 \cdot Ch_m + k_5 \cdot N_p - (k_6 + k_{10}) \cdot N_m \\ \frac{dC}{dt} = k_6 \cdot N_m - (k_7 + k_8) \cdot C \\ \frac{dP}{dt} = k_8 \cdot C - k_9 \cdot P \end{cases} \quad \begin{cases} Ch_p = \frac{k_1}{k_2+k_4} \\ Ch_m = \frac{k_2}{k_3} \cdot \frac{k_1}{k_2+k_4} \\ N_p = \frac{k_4}{k_5} \cdot \frac{k_1}{k_2+k_4} \\ N_m = \frac{k_1}{k_6+k_9} \\ C = \frac{k_1 \cdot k_6}{(k_6+k_{10}) \cdot (k_7+k_8)} \\ P = \frac{k_8}{k_9} \cdot \frac{k_1 \cdot k_6}{(k_6+k_{10}) \cdot (k_7+k_8)} \end{cases} \tag{8}$$

8: System of equation describing the full model including nucleoplasmic mature RNA decay (left) with its initial conditions.

4 Supplementary tables

GENE	FORWARD PRIMER	REVERSE PRIMER
LGALS1	AGCAGCGGGAGGCTGTCTTTC	ATCCATCTGGCAGCTTGACGGT
PRKG2	CTGAAAGCACGGAGCCTTCCAT	CAGCCTGATAGCTCATCTGGAG
POLE4	CTTGGTGAAGGCAGATCCCGAC	TTTCCCTGCTGAGCGCAACAGT
ADARB1	TACATCAGCACCTCTCCCTGTG	CAGTGAGCCATGATTGCACCAC
CYREN	CCAAGAGGATGAGAATGGCAGC	GAGCAACATCAACTATCTCAGCC
DOTIL	GTTCCTGGCATACACAAAGACCC	GCTGAAACAGCCTCCTGATCTC
HTRA1	CAGACGTGATCTCAGGAGCGTA	TCGCTGACATCATTGGCGGAGA
SNRPB2	CTGTGGAACAGACTGCAACAACC	ACCTGAGGATTTGGTGTTGAATTTC
CNBP Spliced	AAACAGCCTCTACCTTGCGA	GCTTCTCATTCCACGACCAC
CNBP Unspliced	CCAAACACACAAACACTGGC	TGAGAACTTGAGGCCACCAA
RPLPO	TTCATTGTGGGAGCAGAC	CAGCAGTTTCTCCAGAGC

Supplementary table 1: RT-qPCR primers.

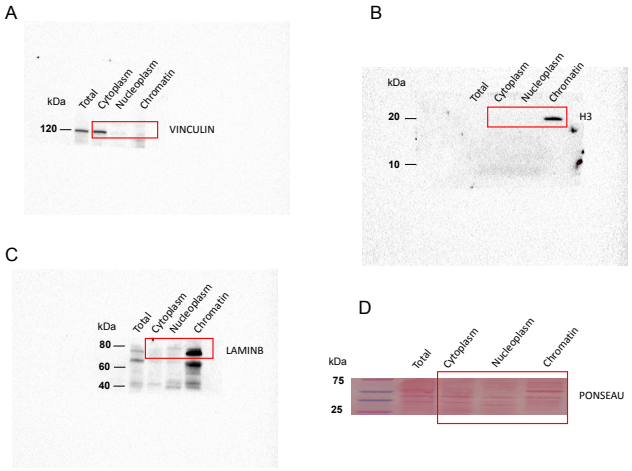
Cell line	Condition	Fraction	Rep	#cells [10^6]	PolyA [ng]				Cell line	Condition	Fraction	Rep	#cells [10^6]	PolyA [ng]
SUM159	Untreated	Chromatin	1	3.0	128.4				SUM159	Leptomycin B	Chromatin	1	3	165.0
SUM159	Untreated	Chromatin	2	2.8	109.2				SUM159	Leptomycin B	Chromatin	2	3	243.0
SUM159	Untreated	Nucleoplasm	1	3.0	122.0				SUM159	Leptomycin B	Nucleoplasm	1	3	160.0
SUM159	Untreated	Nucleoplasm	2	2.8	88.5				SUM159	Leptomycin B	Nucleoplasm	2	3	190.0
SUM159	Untreated	Cytoplasm	1	3.0	335.0				SUM159	Leptomycin B	Cytoplasm	1	3	325.0
SUM159	Untreated	Cytoplasm	2	2.8	308.0				SUM159	Leptomycin B	Cytoplasm	2	3	320.0
SUM159	Untreated	Polysomes	1	29	1298.5				SUM159	Leptomycin B	Polysomes	1	15	117.9
SUM159	Untreated	Polysomes	2	13	516.0				SUM159	Leptomycin B	Polysomes	2	27	164.1
SUM159	Pladienolide B	Chromatin	1	3.3	78.8				SUM159	Leptomycin B	Polysomes	3	15	520.0
SUM159	Pladienolide B	Chromatin	2	3.0	77.9				SUM159	Leptomycin B	Polysomes	4	27	117.8
SUM159	Pladienolide B	Nucleoplasm	1	3.3	70.6				SUM159	Harringtonine	Chromatin	1	3.1	58.1
SUM159	Pladienolide B	Nucleoplasm	2	3.0	67.2				SUM159	Harringtonine	Chromatin	2	2.5	82.8
SUM159	Pladienolide B	Cytoplasm	1	3.3	152.0				SUM159	Harringtonine	Nucleoplasm	1	3.1	57.0
SUM159	Pladienolide B	Cytoplasm	2	3.0	148.2				SUM159	Harringtonine	Nucleoplasm	2	2.5	81.0
SUM159	Pladienolide B	Polysomes	1	19	222.0				SUM159	Harringtonine	Cytoplasm	1	3.1	414.0
SUM159	Pladienolide B	Polysomes	2	14	120.0				SUM159	Harringtonine	Cytoplasm	2	2.5	380.0

Supplementary table 2: Fractions yields.

Library ID SRA	Cell line	Condition	Fraction	Rep	#cells [10^6]	Sequenced Reads	Filtered Reads	ENO2 Reads	ERCC Reads	Median length	IQR length
FAW53197	SUM159	Untreated	Chromatin	1	16	624562	458463	127234	2267	877	901
FAW53157	SUM159	Untreated	Nucleoplasm	1	16	1283483	1130347	54340	1295	919	962
FAW39129	SUM159	Untreated	Cytoplasm	1	16	1619781	1311751	73272	1637	685	528
FAW52964	SUM159	Untreated	Polysomes	1	11	1040230	905645	26715	1150	666	551
FAW76121	SUM159	Untreated	Chromatin	2	21	583189	464126	17227	1675	860	804
FAW72364	SUM159	Untreated	Nucleoplasm	2	21	1205776	988909	16542	2427	787	765
FAW73347	SUM159	Untreated	Cytoplasm	2	21	1414955	1136804	13136	2134	712	601
FAW65788	SUM159	Untreated	Polysomes	2	13	2123150	1837222	140483	1340	685	618
FAX17069	SUM159	Untreated	Polysomes	3	18	1334666	1131305	56549	3524	643	504
FAW87860	SUM159	Pladienolide B	Chromatin	1	16	185925	176153	145110	1179	667	722
FAW88194	SUM159	Pladienolide B	Nucleoplasm	1	16	838799	773917	47376	842	916	902
FAW73342	SUM159	Pladienolide B	Cytoplasm	1	16	1919882	1621256	68536	1180	691	559
FAW76128	SUM159	Pladienolide B	Polysomes	1	19	883590	629265	16902	596	518	362
FAW71080 FAW71176	SUM159	Pladienolide B	Chromatin	2	60	693963 682922	664288 653902	161222 149489	3858 4423	1361 1361	1093 1129
FAW66855	SUM159	Pladienolide B	Nucleoplasm	2	60	1561310	1434238	47475	2171	864	920
FAW65841	SUM159	Pladienolide B	Cytoplasm	2	60	1150186	984680	33570	1472	726	641
FAW89001	SUM159	Pladienolide B	Polysomes	2	14	1308264	1146479	253433	1682	708	744
FAX28954	SUM159	Leptomycin B	Chromatin	1	45	445089	414075	62076	1282	1224	1120
FAX25321	SUM159	Leptomycin B	Nucleoplasm	1	45	1693518	1551779	93019	1855	828	848
FAX27288	SUM159	Leptomycin B	Cytoplasm	1	45	1970524	1670255	75117	1506	653	529
FAX25378	SUM159	Leptomycin B	Polysomes	1	27	1318907	1106207	60759	1487	594	462
FAX25016 FAX25447	SUM159	Leptomycin B	Chromatin	2	46	449346 485219	416796 450591	142801 154108	4507 4867	1175 1167	720 724
FAX27057	SUM159	Leptomycin B	Nucleoplasm	2	46	1601751	1430324	75275	3158	722	611
FAX25494	SUM159	Leptomycin B	Cytoplasm	2	46	2351045	1913355	50976	2244	567	373
FAX17033	SUM159	Leptomycin B	Polysomes	2	27	1027902	861495	47038	1248	595	463
FAX07279	SUM159	Harringtonine	Chromatin	1	21	449159	429253	145509	5494	1331	743
FAX03031	SUM159	Harringtonine	Nucleoplasm	1	21	1204434	1100092	78759	1766	927	937
FAX03159	SUM159	Harringtonine	Cytoplasm	1	21	1902861	1645277	86214	1800	731	721
FAX27070	SUM159	Harringtonine	Chromatin	2	28	357306	342898	99310	3194	1352	973
FAX25063	SUM159	Harringtonine	Nucleoplasm	2	28	1432127	1327587	52952	1157	987	959
FAX27271	SUM159	Harringtonine	Cytoplasm	2	28	1974244	1738768	49226	1029	796	790
FAL77120	K562	Untreated	Ribo0	1	-	3025431	1012758	-	-	378	234
FAL77745	K562	Untreated	PolyA	1	-	1285947	995966	-	-	630	460

Supplementary table 3: Direct RNA-seq runs statistics.

5 Western blots acquisitions



Supplementary figure 57: Original acquisitions for the Western blots presented in Supplementary Figure 6: (A) Vinculin, (B) H3, (C) LaminB, (D) Ponceau. The experiment was repeated in triplicate obtaining similar results.