Supplementary Information for:

Cell-to-cell variability of alternative RNA splicing

Zeev Waks, Allon M. Klein, and Pamela Silver

Supplementary Figures

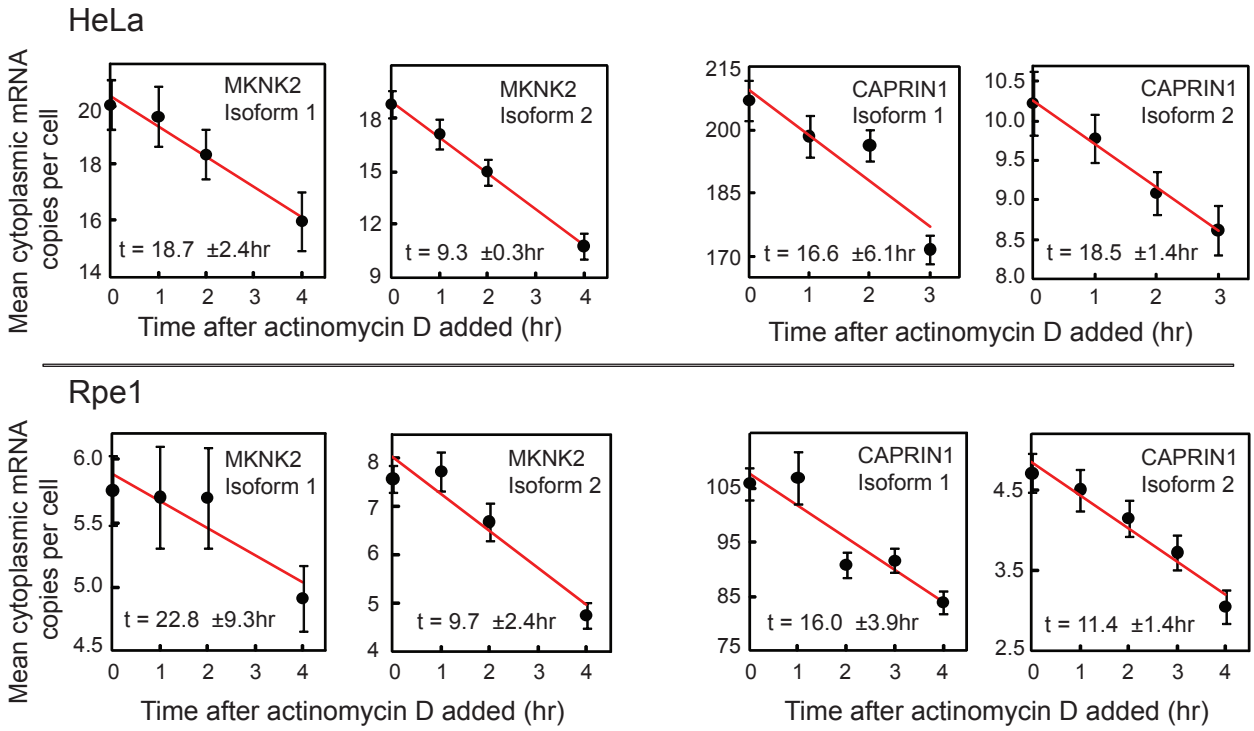

Fig. S1. Isoform lifetime measurements. Isoform lifetimes were measured by quantifying the decrease in mean cytoplasmic isoform copies per cell following actinomycin D treatment. Isoform lifetime estimates are given beneath each mRNA decay curve. Data for each time point represents mean ±s.em. between all cells (See table S2). Errors in isoform lifetimes are ±s.em. of the fitted curve.

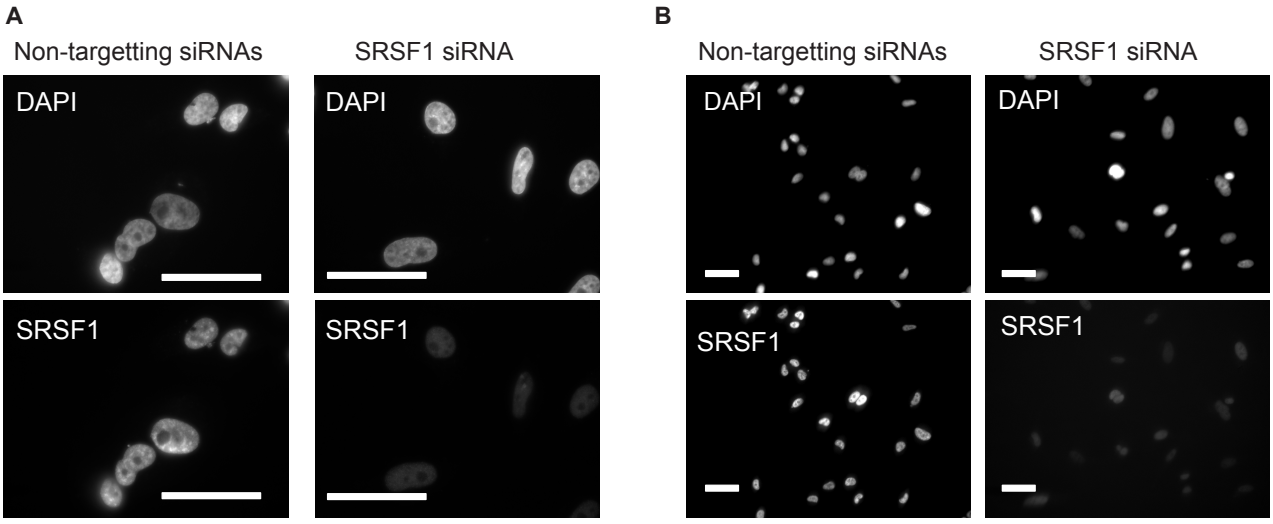

Fig. S2. SRSF1 siRNA transfection results in efficient knockdown across all cells. (A) High magnification (60X) immunofluorescence images of SRSF1 knockdown in HeLa cells demonstrate efficient knockdown of SRSF1. **(B)** Low magnification (20X) images of SRSF1 siRNA transfection in HeLa cells show efficient knockdown among all cells. Scale bars, 50μM.

Supplementary Tables

Table S1. qRT-PCR validation of candidate genes imaging alternative splicing using smFISH. Three primer sets for each isoform were used to estimate mRNA levels relative to GAPDH in HeLa cells. The lengths of the two mRNA sequences available for in situ hybridization (Alexa555 probes and Cy5 probes) are given for each gene.

Gene	Isoform 1 Transcription level (Threshold cycle normalized to Gapdh)			Isoform 2 Transcription level (Threshold cycle normalized to Gapdh)			Sequence length for FISH (nucleotides)	
	Primers A	Primers B	Primers C	Primers A	Primers B	Primers C	Sequence 1	Sequence 2
CAPRIN1	**9.5**	**8.3**	**5.4**	**8.9**	**8.2**	**11.4**	**3308**	**1299**
MKNK2	**7.7**	**7.1**	**7.6**	**9.5**	**7.3**	**9.5**	**2027**	**1778**
DFFA	6.9	7.1	6.8	8.1	9.6	undetected	1355	2053
SYNM	9.5	10.3	7.3	7.5	6.8	7.1	936	6419
TACC3	10.4	11.3	10.8	7.8	5.9	undetected	1080	1700
TRIM11	7.1	7.1	13.8	8.9	13.4	11.1	1573	1413
ING3	8.3	8.6	8.3	8.3	9.3	8.8	1794	821
SF3B1	4.4	4.5	4.6	10.2	9.6	10.7	3827	1664
PFKFB2	13.5	13.2	12.8	10	14.4	15.1	2070	5614
POFUT1	7.7	6.8	7.7	undetected	11.6	undetected	915	4625

Table S2. Number of cells in actinomycin D and ASF/SF2 knockdown experiments.

Experiment	Cell line	Gene or condition	Cells per experiment
Actinomycin D	HeLa	CAPRIN1	62 (0hr), 89 (1hr), 108 (2hr) and 104 cells (3hr)
		MKNK2	185 (0hr), 139 (1hr), 164 (2hr), and 131 cells (4hr)
	Rpe1	CAPRIN1	126 (0hr), 129 (1hr), 154 (2hr), 137 (3hr), and 121 cells (4hr)
		MKNK2	345 (0hr), 168 (1hr), 165 (2hr), and 175 cells (4hr)
ASF/SF2 knockdown	HeLa	ASF/SF2 siRNA	147 cells, 127 cells, and 160 cells
		Control siRNA	143 cells, 145 cells

Table S3. List of Cy5-coupled probes sequences for MKNK2. Probes are ordered from 5' to 3' end on the mRNA.

Probe #	GC%	Sequence	Probe #	GC%	Sequence
1	65	accatcttctgtccgggccc	31	55	ttgagtttgatgccgctgcc
2	45	aagttcggctggtttcttct	32	60	tggagataggggagcagtcc
3	50	ttgaacgaacggtggaaacc	33	60	tacctccggggccatgtact
4	65	cagctcgaaggggttctgcc	34	65	agcctcctcgctgaaggcct
5	60	tcgggctggtctagggagaa	35	55	tcgcagcgcttgtcgtagat
6	55	gccaaagtcagagtctccgt	36	60	aagatgacgcccaggctcca
7	65	atgtcagggcgggctgagca	37	50	gtagccgctgagtaggatat
8	55	ttggcgtccgggatgtcaat	38	60	atgttctggcaggcagggca
9	50	cttcttcttcttgcccctct	39	50	ctcctggatgctctcaaaca
10	65	tgcccgagaagctgtcggtg	40	50	ttgtcggggaactcgtactt
11	45	agctggtagacgtcttcaaa	41	65	gcaggagatgtgggcccagt
12	60	ctcccccagcacatcttcct	42	50	tggagatgaggtctttggca
13	60	aggtctgcactcgggcatga	43	65	ttggcgtcacggaccagcag
14	50	tggctggtgatcaggttgat	44	65	ttgggcggcactcagcctct
15	45	aatgatcttgacggcgtact	45	65	tgaacccaggggtgctgcag
16	60	gaatgtggcctggctgcttc	46	65	aaggtgttctccggggcgca
17	55	acctccctgaaaaccctgct	47	60	tctgcaggaccatgggagtg
18	55	ctggcactggtacagcatct	48	55	aggaggaagtgactgtccca
19	55	gctctaggacgttcctgtgt	49	65	accagtcctccaggtcgcac
20	45	tcctcctcgaagaactcaat	50	45	ctcattcacagtaacggttc
21	50	aaacaccaggtagaagcggt	51	60	ccagggtcctccaggatctt
22	60	atggagcctccccgcatctt	52	45	attaacacatggccgttcac
23	60	gcgcttgtggatgtggctca	53	25	tgtcttttaaaaacatcgta
24	60	cctccagctcgttgaagtgc	54	35	cccccactttaaaaaaactt
25	65	acgtcctgcaccaccacgct	55	35	gaattaaagtgcttggatgt
26	60	cagaaagtccaaggcgctgg	56	50	cgtcagttcacctggtacat
27	50	tgtgggcgatgcctttgtta	57	40	gtaaaggaaaacttctgagc
28	45	atgttttccggctttaggtc	58	45	aattccggcattgacagttg
29	60	ctggttggggtgctcacaga	59	25	atctttacaaaacagaatac
30	45	tcgaagtcacagatcttcac	60	25	gcaagttttttgacttttat

Table S4. List of Alexa555-coupled probes sequences for MKNK2. Probes are ordered from 5' to 3' end on the mRNA.

Probe #	GC%	Sequence	Probe #	GC%	Sequence
1	50	tctttggcacagctgttcct	37	50	tggggtcagctgcagtttaa
2	65	ctcagccgcgaaggacgtga	38	60	tctcaggccttggtgtggag
3	55	agctgccggttcatggcaat	39	60	tgctgcggtggaaacggctt
4	65	agccaggtcctcgtcgtgct	40	55	attggctgacagtgcctgtc
5	60	tgaggtagctcggaccagga	41	60	aaggggtgtcttcaggacc
6	65	gtggagacagctgcaggcag	42	65	tgagaaggggcactcgggtg
7	65	tgcgccagcttggactggga	43	60	tggctacggtcagactgagc
8	60	aggacagactggccctttgc	44	60	tgacggagcttccaggtcag
9	65	tctcccaccaggaccactgg	45	50	tcagctctaaggacaaggac
10	65	gagatgggagggtcaggcgt	46	65	tgaggattcggccagacccc
11	55	cgggtgacctatgtacagag	47	50	agaccaaaaacactgccctg
12	40	acctttagatttgattgggg	48	55	cagccgttttcctgaaggtg
13	55	ccggctggcgatagcttaaa	49	60	actgagcagacccccgcatt
14	55	ttagtgcctgggaatccagg	50	65	ttggagcatggagggtgggg
15	55	aaaacacccccctcagctga	51	50	attgcatagaacgtccccac
16	40	cccaaaagcaaaaaccttct	52	45	cacttctaaagtggaaccct
17	45	ggggtggaaacaggaaaaaa	53	50	atttagaagccagaggggct
18	40	ccaaaggaaaaaataacggg	54	50	aacaggaggaagaatcccga
19	45	tgcctgcaatgcttttaacc	55	20	ttaaggaaaaactaaaaaac
20	55	acctaccctctgctcacctt	56	40	gtccatcgatgtgttgtttt
21	45	ttcatagtagaggtgagcag	57	45	tcttcaagaacagggaagca
22	65	atggcgggcagcacaggtga	58	60	aggccccccgacattcaagt
23	55	tgagtcactgcaagccacgt	59	60	agattcaggaggggcagcag
24	60	aaacgctgtgggcctgctct	60	50	accggactgtgaccatcata
25	60	ctccccagctctgcaagatg	61	65	accagggaaggcagagcacc
26	65	acacgagggcagggtcctgt	62	50	agtccacagatatgggcagt
27	55	tgaaatactgcgggctggga	63	65	cccaggggtctttggaaggg
28	55	atagaggggaagagcctgtc	64	65	aaaggggtgggcggaatgcc
29	65	aggggtgctctcagggtgag	65	55	tgacaggcgagaagtgatgg
30	60	tcccttccccaaccaagcca	66	45	ccgttacgaaacatggaatc
31	60	acacctccagagacaggcag	67	35	atacaaaaaggcagagaatc
32	65	aagggctgggggatcccatg	68	45	gggtccatttgcttgttctt
33	50	actatcatggggacaaggca	69	65	atggctatgggcgcctgcag
34	55	ttgcaggaagccccgattgt	70	55	gccactcagctttagagacc
35	55	aggagtcctggacagacaga	71	40	aacaaacaaacaaacgctgc
36	60	aaggtcggaggccgaatctg			

Table S5. List of Cy5-coupled probes sequences for CAPRIN1. Probes are ordered from 5' to 3' end on the mRNA.

Probe #	GC%	Sequence	Probe #	GC%	Sequence
1	40	caccaccacaaaatattacc	26	35	gaaatgctagctctaattgt
2	45	aacagaagctccacttagga	27	30	catttagaatttcagactac
3	40	tattaacagctcttccaagg	28	45	aacctccctcaaatcagaga
4	35	aatgtattcctaacatgcag	29	40	agaggaaacctgtttgatgt
5	45	agcaacaagtctggaaagga	30	30	ttttatacatgttatggcca
6	35	cagagcatttcatttaatcc	31	40	cgttgtaattcctccttaac
7	35	ggtccaagattaagttttag	32	40	gccagtttctactagtattc
8	30	cagggaaaattaaatcattc	33	30	agaaaaatgtacctttttcc
9	35	gttctagttttcaagacagt	34	35	cccaagtgatccattaattt
10	35	cacttttctgagaagagaa	35	35	acttctagtcaagtcagtaa
11	30	ggacctgaaaaataattttc	36	40	tcacatgcaaacatcctttg
12	35	cctaaaaacatttagcaggt	37	45	tgggaaagaacataacccac
13	30	gtcttacaaaaatgtttcag	38	35	ctttcatcgaatatgctaca
14	30	tgttcaatctcattccaaaa	39	30	ttttagctatcagttaactc
15	30	ggcatgtttcaaaaaaatga	40	30	tttacatgctgttaaaacag
16	40	gtgtctggccctaaaatata	41	30	ctacataacattcaaaactg
17	40	tgttaactatggcttatccg	42	30	ccttattacctaaactgtta
18	35	cacttctaaatggttctcta	43	35	aaagcaccagaatgaaaaca
19	40	ggcattgcaaattccattag	44	30	gtaagtaacatcatactatc
20	35	tatcactaatagaggtccaa	45	40	cactctagcttacatttcag
21	35	gggagttttgtttaaaagtc	46	40	atggagcttacattctagtg
22	35	gtatagctcaataagttagg	47	45	agaagacagacaaggtacct
23	35	caaaaaactctctaaacctg	48	45	agcgttgggaatagatacag
24	40	ggaagtagttttccaacagt	49	50	tatgtgccaggcaccatcat
25	40	aggttaacttcctgccaaaa	50	30	caaatatttattgagtgcct

Table S6. List of Alexa555-coupled probes sequences for CAPRIN1. Probes are ordered from 5' to 3' end on the mRNA.

Probe #	GC%	Sequence	Probe #	GC%	Sequence
1	40	acttgctgagtgttcatttg	49	40	ctgggctaaatgctgaaaat
2	40	ggcgattaaacataatcctg	50	40	taaatagctgagcacctact
3	40	atattatggtacactggcca	51	40	ctccagttgtctaatgatga
4	40	ctgaaagggagaacaaatag	52	40	gtaggttacaaaaccagcaa
5	40	gatgaaaacagtcccttac	53	40	gttcaacagcctatccattt
6	40	gacaattgtagtcctgtctt	54	40	ccaccctacaaaacttttga
7	40	cttccatatccaggtaatag	55	40	acattgaagatcccccattt
8	40	gacagaacatgcagagtaaa	56	40	cacacatgagaaaagtcaag
9	40	atctgagtatcatgtgcaag	57	40	cttccaatatgatgtaccac
10	40	ttttactcctaagcaagggt	58	40	catttgccaaagtaacaggt
11	40	ggagattattatcaccctgt	59	40	ggaggtgctagcaaaaaaat
12	40	ccttttcaagccacttcaa	60	40	gatgtcatttaaagcacgca
13	40	gagggctgatttgtagattt	61	40	acatatggttgtggtacatc
14	40	catcttcctttctagggaaa	62	40	atccattcccctaagataca
15	40	aatgcagttgttctgccaaa	63	40	gattacccagtaaaccacaa
16	40	gtgctccaatttaactggaa	64	40	ctgcaagcatacatcatcta
17	40	ttggtatgcatctgaacgtt	65	40	cagtcctgattgtaaatgtc
18	40	tatgtgtgattaggacccat	66	40	tcactgtgtgaagaacttac
19	40	atgaacagtgctgtgtcaaa	67	40	cttctctgaaactgattcag
20	40	tttaaccacagttgtttggc	68	40	aacctagaaggcattttctc
21	40	gtcttatacagtatcagctg	69	40	gcactaatgcatagaagttc
22	40	cctaattttgcatcttggct	70	40	caccttacacattcacaac
23	40	cggattacatcccatatttg	71	40	ggtctgagtagagacataaa
24	40	tcacattaagtacagaagcg	72	40	tgaggcacccctttaaaaat
25	40	gacagacaatgtacatgtca	73	40	tattgaaaagttctctgccc
26	40	caatacagcaaaatccaagg	74	40	gcttagaaccaagattcaga
27	40	gtacctgtaccaatttaggt	75	40	ctaggacaagcaatcacata
28	40	cgagagtgttatgattatcc	76	40	ttgaccaaacaaaaggcctt
29	40	agctgaaggaaaaacatgtg	77	40	catttccttgaaactctgct
30	40	ggcatttggtatcttttcct	78	45	ctacagtgacatgtgtgaca
31	40	aattgaaaagggtggtagca	79	40	ttgtatgtgagcttgctaca
32	40	ctggtgcctttcaaaagata	80	40	gccctaatcaaggtaatcaa
33	40	ctgaaacagggaaatcaatc	81	40	gaactgtagaagctaactga
34	40	gaaactactgtccgtgattt	82	40	gctgcatataaaaggagaca
35	40	ccagtgtttcaagatacaac	83	40	ttaagtggaaaagtctccca
36	40	ataagctgctagctgttgaa	84	40	aagcacacattctatgtctc
37	40	aatttgtgacactaacgtgg	85	40	cagttaatgaaccttctgag
38	40	tgttgctggagataaaccat	86	40	gctcaactagttgttaactc
39	40	cggagaattcagttaaggtt	87	40	ctaaaacacttaggaagctg
40	40	atataaatgcctccaggaga	88	40	cctctagttatatgacggaa
41	40	ctaagggaaggaattatcac	89	40	attcggcacttatctgcatt
42	40	atttagagactctccctatg	90	40	cagattcaggagcttaaaca
43	40	actcaagtgtccatttccat	91	40	cttaaaaacaacccaagtgg
44	40	gagtacataaggctaagtca	92	40	ggcaatactgattctgtgaa
45	40	actgctagcacaaattccaa	93	40	ccgttaactccataaaaagg
46	40	attaacttcctaggcacaca	94	40	ttactttcgttttcctgagg
47	40	gtcagaatgagacaataagc	95	40	gtatgaagatgagccttaac
48	40	aggtgggattaattctccat	96	40	cattcactagtcttttgcag

Supplementary Theory

S-I Stochasticity in the alternative splicing decision alone does not increase mRNA abundance variability

To familiarize ourselves with the effect of alternative splicing on heterogeneity in mRNA levels, we begin by addressing its relationship to Poisson (or 'shot') noise in mRNA synthesis. For a large class of systems, including the one we shall consider here, the random and independent timing of consecutive transcription initiation events lead to fluctuations in mRNA abundance, n, that result in a variance of the form[1],

$$\mathrm{Var}[n] = \underbrace{\langle n \rangle}_{\text{Poisson noise}} + \quad \text{additional sources of variability.} \tag{1}$$

In the absence of splicing factor fluctuations, alternative splicing events do not affect this Poisson variability. This means, for example, that a gene that is constitutively transcribed at some average rate, say 1 transcript/minute, will have the same cell-to-cell variability in transcript number as the isoform of a gene that is transcribed at double the rate (2 transcripts/minute), and has the same mRNA lifetime — but where only half of the transcripts are randomly spliced into that isoform.

This is the case because the defining feature of the Poisson process remains unchanged: the probability of a transcription initiation event occurring in a small period of time is constant. By introducing an alternative splicing step, only a fraction of pre-mRNA transcripts become a given isoform, however the probability of transcription initiation for any particular isoform remains time-invariant. Therefore, in the absence of splicing factor fluctuations, mRNA synthesis and splicing preserve Poisson statistics.

S-II Defining a model of stochastic mRNA transcription and splicing

Recent years have seen considerable progress in the development of theoretical approaches to elucidate the origins of variability in mRNA levels [supp. ref. *2–10*]. Here we adapt these approaches to the case of alternative splicing. We make few model-specific assumptions

[1] This well-known form of the variance holds for many, but not all, stochastic systems. For example it fails for molecules that stimulate their own degradation [supp. ref. *1*].

about the nature of fluctuations in mRNA synthesis rates and alternative splicing outcomes.

We focus on a homogeneous cell population that is in steady-state, characterized by a mean pre-mRNA synthesis rate, $\bar{\nu}$. Each pre-mRNA molecule is rapidly spliced into one of two mature isoforms (hereafter isoforms 1 and 2), which are subsequently degraded after a mean mRNA lifetime τ_1 or τ_2, respectively. Given that each splicing event occurs independently, the outcome of alternative splicing is a random event (a Bernoulli trial) with a mean probability, $\bar{p}$, of generating isoform 1. Taken together, these events can be represented by the following processes,

$$\begin{aligned} n_1 &\xrightarrow{p(t)\nu(t)} n_1+1 \\ n_2 &\xrightarrow{[1-p(t)]\nu(t)} n_2+1 \\ n_1 &\xrightarrow{n_1/\tau_1} n_1-1 \\ n_2 &\xrightarrow{n_2/\tau_2} n_2-1 \end{aligned} \tag{2}$$

where n_1, n_2 denote the number of mRNA molecules of each of the two isoforms, and $\nu(t),\ p(t)$ are respectively the instantaneous mRNA synthesis rate and the instantaneous alternative splicing probability at a time t during the experiment, with expectation values $\langle \nu(t) \rangle = \bar{\nu}$ and $\langle p(t) \rangle = \bar{p}$. In many studies the degradation processes shown in Eq. (2) are typically taken to reflect an exponential decay of mRNA molecules. However we will also consider other distributions of the mRNA lifetime, for example assuming that mRNA lifetimes are near-uniform.

To complete the model, we must characterize the fluctuations in mRNA synthesis rates and splicing factors. We assume a general form for the auto-covariance function of the synthesis rate that is suitable for describing a wide range of sources of synthesis rate fluctuations characterized by just a single relevant timescale, viz.

$$G_\nu(\Delta t) = \langle \nu(t)\nu(t+\Delta t) \rangle - \bar{\nu}^2 = \sigma_\nu^2 e^{-|\Delta t|/\tau_{\rm syn}}. \tag{3}$$

Here the magnitude of the fluctuations is $\sigma_\nu^2 = \mathrm{Var}[\nu(t)]$, and $\tau_{\rm syn}$ is the correlation time. This exponential form of synthesis rate fluctuations describes stochastic gene activation-inactivation, [supp. ref. *5*, *7*], in addition to more subtle origins of cell-to-cell variability in mRNA synthesis.

Similarly, fluctuations in splicing factors may affect the splicing probability $p(t)$, and are

characterized by the auto-covariance function,

$$\begin{aligned} G_p(\Delta t) &= \langle p(t)p(t+\Delta t)\rangle - \bar{p}^2 \\ &= \sigma_{\rm sp}^2 e^{-|\Delta t|/\tau_{\rm sp}}\,, \end{aligned} \tag{4}$$

where $\sigma_{\rm sp}^2 = {\rm Var}[p(t)]$ gives the magnitude of fluctuations in the mean splicing probability, and $\tau_{\rm sp}$ is the correlation time of these fluctuations. Thus, outcome of alternative splicing events, $P = 1$ for isoform 1 or $P = 0$ for isoform 2, has an auto-correlation function

$$G_P(\Delta t) = \begin{cases} \bar{p}(1-\bar{p}) & \text{for } |\Delta t| = 0 \\ \sigma_{\rm sp}^2 e^{-|\Delta t|/\tau_{\rm sp}} & \text{for } |\Delta t| > 0 \end{cases}\,, \tag{5}$$

The term at $|\Delta t| = 0$ reflects the fact that splicing is a binomial process, which requires that ${\rm Var}[P] = G_P(0) = \bar{p}(1-\bar{p})$. This completes the definition of the model.

S-III Measurements of mean splicing probabilities and mRNA synthesis rates

From process (2) one may show that the average number of mRNA molecules per cell is given by,

$$\langle n_1\rangle = \bar{p}\bar{\nu}\tau_1\,,\ \ \langle n_2\rangle = (1-\bar{p})\bar{\nu}\tau_2\,. \tag{6}$$

When mRNA lifetimes are comparable to or longer than the cell cycle length, the mean lifetimes τ_1, τ_2 in Eq. (6) represent the effective mRNA lifetimes within a cell, which takes into account the halving of mRNA abundance at each cell division [supp. ref. *11*, *12*].

This study has also measured the average fractional isoform 1 abundance,

$$\langle {\rm FIA}_1\rangle = \left\langle \frac{n_1}{n_1+n_2} \right\rangle .$$

For the observed isoform abundances we find, $\langle n_1/(n_1+n_2)\rangle \approx \langle n_1\rangle/\langle n_1+n_2\rangle$, which together with Eq. (6) gives,

$$\langle {\rm FIA}_1\rangle \approx \frac{\bar{p}}{\bar{p}+(1-\bar{p})(\tau_2/\tau_1)}\,, \tag{7}$$

and

$$\bar{p} = \langle n_1\rangle\tau_2/[\langle n_1\rangle\tau_2 + \langle n_2\rangle\tau_1].$$

One may also see that the mean mRNA synthesis rate can be measured as, $\bar{\nu} = \langle n_1 \rangle / \tau_1 + \langle n_2 \rangle / \tau_2$.

We will treat the empirical measurements of isoform lifetimes of *MKNK2* and *CAPRIN1* (Fig. S3) as the effective lifetimes, including any cell division effects[2]. This approach is approximate since the isoform lifetimes have a length similar to the cell cycle duration. However, any error introduced by this estimate will have only a minor effect on measurements of the splicing probability $\bar{p}$, but may result in larger corrections to the synthesis rate, $\bar{\nu}$, (of order 50%). With this limitation, we have calculated $\bar{p}$ and $\bar{\nu}$ for *CAPRIN1* and *MKNK2* in the table below (errors in $\bar{p}$ show standard error of the mean from uncertainties in empirical measurements; errors in $\bar{\nu}$ show SD between experiments; these reflect the most significant sources of error for each quantity):

Gene	Cell line	$\bar{p}$	$\bar{\nu}$ (1/hour)
CAPRIN1	Rpe1	0.94 ± 0.02	8.15 ± 0.75
	HeLa	0.96 ± 0.06	14.66 ± 0.58
MKNK2	Rpe1	0.21 ± 0.05	1.36 ± 0.19
	HeLa	0.34 ± 0.04	2.99 ± 0.42

Interestingly, one may see that the mean synthesis rate of both genes is approximately two-fold larger in HeLa cells compared to Rpe1 cells, perhaps reflecting increased chromosome abundance in the former cell line.

S-IV Variability and covariance of mRNA isoform abundances

Analysis of process (2) using the auto-correlation functions (3) and (5) allows one to derive expressions for the CVs and the covariance of the mRNA isoforms, and reveals the relationship between fluctuations in splicing factor activity and gene expression fluctuations. Technical details of the calculation are provided in section S-VII; for simplicity, we omit corrections that arise from mRNA partitioning upon cell division.

[2]The effect of cell division on the effective mRNA lifetimes is highly sensitive to fluctuations in the cell cycle length and to intrinsic noise in mRNA degradation rates. For example, if both the cell cycle length T_c and mRNA lifetimes τ are highly uniform (i.e. have no fluctuations), then the effective lifetimes acquire the expressions, $\tau^{(\mathrm{eff})} = T_c(1 - 2^{-\tau/T_c})/\ln(2)$, where T_c is the cell cycle time. By contrast, if mRNA molecules should undergo exponential decay, then $\tau^{(\mathrm{eff})} = \tau/[1 + (\tau/T_c)\ln(2)]$ [supp. ref. *11*].

We find that the CVs of the mRNA isoforms acquire the form,

$$\begin{aligned} CV_k^2 &= \frac{\mathrm{Var}[n_k]}{\langle n_k\rangle^2} \\ &= \underbrace{\frac{1}{\langle n_k\rangle}}_{\text{Poisson}} + \underbrace{CV_{\text{syn}}^2 \underbrace{\mathcal{A}\left[\frac{\tau_k}{\tau_{\text{syn}}}\right]}_{\text{time-avg.}}}_{\text{synthesis fluctuations}} + \underbrace{(CV_{\text{sp}}^{(k)})^2 \underbrace{\mathcal{A}\left[\frac{\tau_k}{\tau_{\text{sp}}}\right]}_{\text{time-avg.}}}_{\text{splicing fluctuations}} + \underbrace{CV_{\text{syn}}^2 (CV_{\text{sp}}^{(k)})^2 \mathcal{A}\left[\frac{\tau_k}{\tau_{\text{syn}}} + \frac{\tau_k}{\tau_{\text{sp}}}\right]}_{\text{synthesis and splicing}} \end{aligned} \tag{8}$$

for $k = 1, 2$. Here, the Poisson term corresponds to the intrinsic noise that persists even in the absence of all sources of mRNA synthesis and splicing variability. The second term arises from fluctuations in the mRNA synthesis rate (see also [supp. ref. *5*] and references therein). The term $\mathcal{A}[\tau_k/\tau_{\text{syn}}]$ arises from time-averaging of fluctuations over the mRNA lifetime. The precise form of $\mathcal{A}[x]$ depends on the detailed mRNA lifetime distribution: for exponentially-distributed mRNA lifetimes one finds, $\mathcal{A}[x] = 1/(1+x)$, which is typically considered in the literature [supp. ref. *4*, *5*, *7*]. However for near-uniform mRNA lifetimes, which might better approximate the degradation mechanism of eukaryotic mRNA (discussed later in section S-V) [supp. ref. *10*], gives $\mathcal{A}[x] = 2(e^{-x} - 1 + x)/x^2$.

Turning to the splicing term in Eq. (8), one may see that this term closely mimics the structure of the synthesis term, effectively acting as a second, independent, source of synthesis rate fluctuations for each mRNA isoform. The quantity $CV_{\text{sp}}^{(k)} = \{\sigma_{\text{sp}}/\bar{p}$ for $k = 1$; $\sigma_{\text{sp}}/(1-\bar{p})$ for $k = 2\}$ reveals the magnitude of splicing factor fluctuations from the perspective of each isoform: for example, if $\bar{p}$ is small then these fluctuations will appear considerably larger for isoform 1 compared to isoform 2.

The final term in Eq. (8) arises from the multiplicative effect of synthesis and splicing fluctuations on mRNA abundance. For example, a short burst in mRNA synthesis at some time in the experiment will lend more weight to the splicing probability $p(t)$ at that time, whereas a pause in mRNA synthesis will mask the effects of $p(t)$ fluctuations, as no mRNA can be spliced until synthesis resumes. This is a second-order effect that undergoes strong time-averaging, and can therefore be neglected provided that fluctuations remain small.

Next, the normalized covariance between the number of mRNA molecules of the two

isoforms is found to be (see section S-VII),

$$\frac{\mathrm{Cov}[n_1,n_2]}{\langle n_1\rangle\langle n_2\rangle} = \overbrace{CV^2_{\mathrm{syn}}\left(\frac{\tau_1}{\tau_1+\tau_2}\underbrace{\mathcal{A}\left[\frac{\tau_1}{\tau_{\mathrm{syn}}}\right]}_{\mathrm{time-avg.}}+\frac{\tau_2}{\tau_1+\tau_2}\underbrace{\mathcal{A}\left[\frac{\tau_2}{\tau_{\mathrm{syn}}}\right]}_{\mathrm{time-avg.}}\right)}^{\text{synthesis rate fluctuations}} \underbrace{-CV^{(1)}_{\mathrm{sp}}CV^{(2)}_{\mathrm{sp}}\left(\frac{\tau_1}{\tau_1+\tau_2}\mathcal{A}\left[\frac{\tau_1}{\tau_{\mathrm{sp}}}\right]+\frac{\tau_2}{\tau_1+\tau_2}\mathcal{A}\left[\frac{\tau_2}{\tau_{\mathrm{sp}}}\right]\right)}_{\text{splicing factor fluctuations}} + (\text{synergistic synth. \& splicing terms})\,. \tag{9}$$

These terms have the same interpretation as in Eq. (8). We see here that fluctuations in splicing factors have an opposite effect on covariance compared to synthesis rate fluctuations: while the latter acts similarly on the two isoforms, the former does not.

S-V Intrinsic noise in mRNA degradation leads to time-averaging of synthesis and splicing fluctuations, but does not contribute independently to isoform variability

mRNA degradation in lower organisms is typically modelled as a single stochastic event characterized by an exponential distribution in mRNA lifetimes, akin to a process of radioactive decay. Such mRNA exponential decay gives rise to a wide variance in lifetimes ($CV = 100\%$). However, mammalian mRNA degradation typically requires sequential steps of de-adenylaton at the 3' end prior to rapid decay by nucleases [supp. ref. *13*], a mechanism which may lead to a tighter distribution of mRNA lifetimes [supp. ref. *14*]. Therefore, stochastic (or 'exponential') degradation is unlikely. An important question is therefore the degree to which stochastic noise in mRNA lifetimes may affect isoform abundance distributions.

Referring to Eqs. (8) and (9), we see that mRNA degradation noise does not contribute independent terms to the isoform variability or covariance. Rather, such noise appears only indirectly in the synthesis and splicing terms through the time-averaging function $\mathcal{A}[x]$. Thus, if synthesis and splicing factor fluctuations were both absent, intrinsic noise in mRNA degradation would have no effect on isoform variability.

When synthesis or splicing factor fluctuations are present, the intrinsic noise in mRNA degradation determines their time-averaging profiles, $\mathcal{A}[\tau_k/\tau_{\mathrm{syn}}]$ and $\mathcal{A}[\tau_k/\tau_{\mathrm{sp}}]$, which serve

to dampen the effects of these upstream fluctuations. Paradoxically, intrinsic noise in mRNA lifetimes *increases* time-averaging and thus reduces mRNA variability. To see this, let us compare the case of exponential mRNA degradation (i.e. the case of high intrinsic noise) with a case in which all mRNA molecules have a fixed lifetime with no variability. For the former case we find $\mathcal{A}_{\rm exp}[x] = 1/(1+x)$ [supp. ref. *5*], whereas in the latter $\mathcal{A}_{\rm fix}[x] = 2(e^{-x} - 1 + x)/x^2$ (see Section S-VII). These time-averaging functions both approach unity when the mRNA lifetime is short compared to the timescale of synthesis or splicing fluctuations. However, when the mRNA lifetime is long, we find that the time-averaging is twice as strong when mRNA lifetimes are exponentially distributed, viz. $\lim_{x\gg 1} \mathcal{A}_{\rm fix}[x] = 2\mathcal{A}_{\rm exp}[x]$. In neither case, however, do intrinsic fluctuations in mRNA lifetime contribute independently to cell-to-cell variability in isoform abundance [supp. ref. *5*].

S-VI Variability and covariance of mRNA isoform levels suggest the presence of splicing factor uctuations in HeLa cells

From Eq. (8), it is evident that the CVs of two mRNA isoforms of a single gene differ in their Poisson terms, lifetime-averaging, and the effect of fluctuations in splicing factor activity. If splicing factor fluctuations are negligible (i.e. $CV_{\rm sp}^{(1,2)} \ll CV_{\rm syn}$) one may see that the isoform statistics satisfy the empirical relationship provided by Eq. (1) of the main text,

$$\boxed{\frac{\mathrm{Cov}[n_1, n_2]}{\langle n_1\rangle\langle n_2\rangle} = \left(CV_1^2 - \frac{1}{\langle n_1\rangle}\right)\frac{\tau_1}{\tau_1+\tau_2} + \left(CV_2^2 - \frac{1}{\langle n_2\rangle}\right)\frac{\tau_2}{\tau_1+\tau_2}\,.} \tag{10}$$

When splicing factor fluctuations are significant the equality breaks down, giving instead,

$$\begin{aligned}\frac{\mathrm{Cov}[n_1, n_2]}{\langle n_1\rangle\langle n_2\rangle} \;\simeq\; & \left(CV_1^2 - \frac{1}{\langle n_1\rangle}\right)\frac{\tau_1}{\tau_1+\tau_2} + \left(CV_2^2 - \frac{1}{\langle n_2\rangle}\right)\frac{\tau_2}{\tau_1+\tau_2} \\ & - \left[w_1\left(CV_{\rm sp}^{(1)}\right)^2 + w_2\left(CV_{\rm sp}^{(2)}\right)^2 + (w_1+w_2)CV_{\rm sp}^{(1)}CV_{\rm sp}^{(2)}\right]\,,\end{aligned} \tag{11}$$

where, $w_k = \mathcal{A}\left[\tau_k/\tau_{\rm sp}\right] \times \tau_k/(\tau_1+\tau_2)$. Thus, in the presence of splicing factor fluctuations the covariance is predicted to drop below the value predicted by the isoform CVs, which results in the inequality in Eq. (1) of the main text[3].

[3]If present, it would also be possible for fluctuations in the mean isoform lifetimes to diminish covariance while increasing isoform CVs. This would lead to a similar inequality in Eq. (1) of the main text. However, for this to occur, fluctuations in the mean isoform lifetimes would have to be independent, or anti-correlated, for the two isoforms. As extrinsic lifetime fluctuations should affect both mRNA species in the same manner, it is

S-VII Analysis of the stochastic model

This section provides an outline of the technical considerations used to obtain Eqs. (8) and (9) above.

S-VII.1 Approximate solution using the fluctuation-dissipation theorem (FDT)

The FDT is a powerful method to derive the approximate form of Eqs. (8) and (9) [supp. ref. *5*]. Briefly, we define the normalized steady-state covariance matrix, η, as a 4×4 symmetric matrix of the covariances of $\nu(t)$, $p(t)$, n_1 and n_2. The matrix acquires the form,

$$\eta = \begin{pmatrix} CV_{\rm syn}^2 & 0 & \cdot & \cdot \\ 0 & (CV_{\rm sp}^{(1)})^2 & \cdot & \cdot \\ \cdot & \cdot & CV_1^2 & C_{12} \\ \cdot & \cdot & C_{12} & CV_2^2 \end{pmatrix} \begin{matrix} \leftarrow \nu(t) \\ \leftarrow p(t) \\ \leftarrow n_1 \\ \leftarrow n_2 \end{matrix} ,$$

where $C_{12} = \mathrm{Cov}[n_1, n_2]/\langle n_1\rangle\langle n_2\rangle$, and the arrows on the right show the row/column assignments of the matrix. For clarity, we have only shown the matrix elements that are relevant to our analysis. With this covariance matrix, the FDT can be expressed as,

$$M\eta + (M\eta)^T = D\,. \tag{12}$$

Here M is the normalized Jacobian matrix and D is the diffusion matrix of process (2), which have the particular forms (see [supp. ref. *5*] for details on their construction),

$$M = \begin{pmatrix} \frac{1}{\tau_{\rm syn}} & 0 & 0 & 0 \\ 0 & \frac{1}{\tau_{\rm sp}} & 0 & 0 \\ -\frac{\bar{p}^2}{\tau_1} & -\frac{\bar{\nu}^2}{\tau_1} & -\frac{1}{\tau_1} & 0 \\ -\frac{(1-\bar{p})^2}{\tau_2} & -\frac{\bar{\nu}^2}{\tau_2} & 0 & -\frac{1}{\tau_2} \end{pmatrix}, \; D = 2\begin{pmatrix} \frac{CV_{\rm syn}^2}{\tau_{\rm syn}} & 0 & 0 & 0 \\ 0 & \frac{CV_{\rm sp}^2}{\tau_{\rm sp}} & 0 & 0 \\ 0 & 0 & \frac{1}{\bar{p}\bar{\nu}\tau_1^2} & 0 \\ 0 & 0 & 0 & \frac{1}{(1-\bar{p})\bar{\nu}\tau_2^2} \end{pmatrix}.$$

Substituting these matrices into Eq. (12) and solving for η, one directly obtains the leading terms of Eqs. (8) and (9).

Given that the FDT is exact only for linear noise processes, this approach does not generate the synergistic (second-order) terms in Eqs. (8) and (9), which may be relevant if fluctu-

unlikely that these fluctuations could explain the observed failure of Eq. (10).

ations in splicing probabilities and synthesis rates are both large. To obtain the exact results including all terms, in the following section we outline a more detailed analysis using linear response theory.

S-VII.2 Exact solution for fixed isoform lifetime

Using linear response theory [supp. ref. *4*], the number of mRNA molecules of isoform $k = 1, 2$, in a single cell at a time t of the experiment is given by the convolution,

$$n_k(t) = \int_{-\infty}^{t} H_k(t - t')S_k(t')dt' \equiv [H_k * S_k](t)\,, \tag{13}$$

where $S_k(t)$ is the instantaneous synthesis rate, and $H_k(t) = \{1 \text{ for } 0 < t < \tau_k;\ 0 \,\text{otherwise}\}$ is the linear response function describing a fixed mRNA lifetime. The fixed-lifetime approximation will provide a reasonable estimate of mRNA abundance statistics provided that the intrinsic variability between mRNA lifetimes is small (say CV$\lesssim$50%). As specified in process (2), the synthesis rate $S_k(t)$ is a Poisson process with fluctuations in its average rate $\nu(t)$, multiplied by the fluctuating splicing outcome $P(t)$. The steady-state average values are thus $\langle S_1 \rangle = \bar{\nu}\bar{p}$, and $\langle S_2 \rangle = \bar{\nu}(1 - \bar{p})$.

Eq. (13) only provides an exact description of mRNA degradation when the lifetime is taken to be deterministically fixed[4]. To derive Eqs. (8) and (9), we make use of the relations,

$$\mathrm{Var}[n_k] \;=\; \frac{1}{2\pi}\int_{-\infty}^{\infty} |\tilde{H}_k(\omega)|^2 \tilde{G}_{S_k,S_k}(\omega)d\omega\,, \tag{14}$$

and

$$\mathrm{Cov}[n_1, n_2] = \frac{1}{2\pi}\int_{-\infty}^{\infty} \tilde{H}_1^*(\omega)\tilde{H}_2(\omega)\tilde{G}_{S1,S2}(\omega)d\omega\,, \tag{15}$$

where $\tilde{H}_k(\omega) = (1 - e^{-i\omega\tau_k})/i\omega$ is the Fourier transform of the response function, and the terms $\tilde{G}_{S_i,S_j}(\omega)$ are the Fourier transforms of the auto-correlation functions of the isoform synthesis rates (for $i = j$) and of their cross-correlation function (for $i \neq j$).

[4]Previous studies have applied the same formalism assuming an exponential decaying response function, $H_k(t) = e^{-t/\tau_k}$, which makes the approximation that mRNA molecules are capable of continuous (i.e. fractional) decay [supp. ref. *4*, *7*]. This correctly captures most aspects of the statistics, but it introduces errors into the Poisson noise term.

To obtain expressions for the synthesis rate correlation functions, the synthesis rates can be written explicitly as $S_1(t) = s(t)P(t)$ and $S_2(t) = s(t)[1-P(t)]$, where $s(t)$ is a Poisson process with an auto-covariance $G_s(\Delta t) = \bar{\nu}\delta(\Delta t) + G_\nu(\Delta t)$, and $G_\nu(\Delta t)$, $G_P(\Delta t)$ are defined in Eqs. (3) and (5). Consequently, multiplication of the synthesis and splicing noise terms gives,

$$G_{S_1,S_1}(t) = \underbrace{\bar{p}\bar{\nu}\delta(t)}_{\text{Poisson noise}} + \underbrace{\bar{p}^2 G_\nu}_{\text{Synth.}} + \underbrace{\bar{\nu}^2 G_P}_{\text{Spl.}} + \underbrace{G_\nu G_P}_{\text{Synth.}\times\text{Spl.}}, \tag{16}$$

$$G_{S_1,S_2}(t) = \underbrace{\bar{p}[1-\bar{p}]G_\nu}_{\text{Synth.}} - \underbrace{\bar{\nu}^2 G_P}_{\text{Spl.}} - \underbrace{G_\nu G_P}_{\text{Synth.}\times\text{Spl.}}. \tag{17}$$

The expression for $G_{S_2,S_2}(t)$ has the same form as $G_{S_1,S_1}(t)$, with the sole change $\bar{p} \mapsto 1-\bar{p}$. Finally, taking the Fourier Transform of these expressions using Eqs. (3) and (5), and integrating Eqs. (14) and (15), we obtain the full solutions for the isoform variances and covariance in Eqs. (8) and (9).

Supplementary theory references

1. J. PAULSSON and A. HILFINGER, Private correspondence.

2. Z. WAKS and P. A. SILVER, Nuclear origins of cell-to-cell variability, *Cold Spring Harb Symp Quant Biol.* **75** (2010).

3. J. PECCOUD and B. YCART, Markovian modeling of gene-product synthesis, *Theoretical Population Biology* **48**, 222 (1995).

4. E. M. OZBUDAK, M. THATTAI, I. KURTSER, A. D. GROSSMAN, and A. VAN OUDENAARDEN, Regulation of noise in the expression of a single gene, *Nat Genet* **31**, 69 (2002).

5. J. PAULSSON, Summing up the noise in gene networks, *Nature* **427**, 415 (2004).

6. D. ZENKLUSEN, D. R. LARSON, and R. H. SINGER, Single-RNA counting reveals alternative modes of gene expression in yeast, *Nat Struct Mol Biol* **15**, 1263 (2008).

7. A. RAJ, C. S. PESKIN, D. TRANCHINA, D. Y. VARGAS, and S. TYAGI, Stochastic mRNA synthesis in mammalian cells, *Plos Biol* **4**, e309 (2006).

8. S. IYER-BISWAS, F. HAYOT, and C. JAYAPRAKASH, Stochasticity of gene products from transcriptional pulsing, *Phys. Rev. E* **79**, 31911 (2009).

9. J. RAUSENBERGER, C. FLECK, J. TIMMER, and M. KOLLMANN, Signatures of gene expression noise in cellular systems, *Prog Biophys Mol Biol* **100**, 57 (2009).

10. J. M. PEDRAZA and J. PAULSSON, Effects of molecular memory and bursting on fluctuations in gene expression, *Science* **319**, 339 (2008).

11. D. HUH and J. PAULSSON, Non-genetic heterogeneity from random partitioning at cell division, *Nature Genetics, in press* (2010).

12. D. HUH and J. PAULSSON, Random partitioning of molecules at cell division, *pre-print* (2010).

13. J. G. BELASCO, All things must pass: contrasts and commonalities in eukaryotic and bacterial mRNA decay, *Nat Rev Mol Cell Biol* **11**, 467 (2010).

14. C. J. DECKER and R. PARKER, A turnover pathway for both stable and unstable mRNAs in yeast: evidence for a requirement for deadenylation, *Genes & Development* **7**, 1632 (1993).