

RESEARCH ARTICLE

# Evolution of intron splicing towards optimized gene expression is based on various *Cis*- and *Trans*-molecular mechanisms

Idan Frumkin<sup>1\*</sup>, Ido Yofe<sup>1</sup>, Raz Bar-Ziv<sup>1</sup>, Yonat Gurvich<sup>1</sup>, Yen-Yun Lu<sup>2</sup>, Yoav Voichek<sup>1</sup>, Ruth Towers<sup>1</sup>, Dvir Schirman<sup>1</sup>, Heike Krebber<sup>2</sup>, Yitzhak Pilpel<sup>1\*</sup>

**1** Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel, **2** Abteilung für Molekulare Genetik, Institut für Mikrobiologie und Genetik, Göttinger Zentrum für Molekulare Biowissenschaften (GZMB), Georg-August Universität Göttingen, Göttingen, Germany

These authors contributed equally to this work.

\* [frumkin.idan@gmail.com](mailto:frumkin.idan@gmail.com) (IF); [pilpel@weizmann.ac.il](mailto:pilpel@weizmann.ac.il) (YP)



**OPEN ACCESS**

**Citation:** Frumkin I, Yofe I, Bar-Ziv R, Gurvich Y, Lu Y-Y, Voichek Y, et al. (2019) Evolution of intron splicing towards optimized gene expression is based on various *Cis*- and *Trans*-molecular mechanisms. *PLoS Biol* 17(8): e3000423. <https://doi.org/10.1371/journal.pbio.3000423>

**Academic Editor:** Laurence D. Hurst, University of Bath, UNITED KINGDOM

**Received:** October 9, 2018

**Accepted:** August 8, 2019

**Published:** August 23, 2019

**Copyright:** © 2019 Frumkin et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All raw files are available in the Sequence Read Archive (accession numbers SRP156917 and SRP156917).

**Funding:** This study was supported by the Minerva Foundation, which funded the Minerva Center for Live Emulation of Evolution in the Lab (grant number AZ 5746940763 to YP). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Abstract

Splicing expands, reshapes, and regulates the transcriptome of eukaryotic organisms. Despite its importance, key questions remain unanswered, including the following: Can splicing evolve when organisms adapt to new challenges? How does evolution optimize inefficiency of introns' splicing and of the splicing machinery? To explore these questions, we evolved yeast cells that were engineered to contain an inefficiently spliced intron inside a gene whose protein product was under selection for an increased expression level. We identified a combination of mutations in *Cis* (within the gene of interest) and in *Trans* (in mRNA-maturation machinery). Surprisingly, the mutations in *Cis* resided outside of known intronic functional sites and improved the intron's splicing efficiency potentially by easing tight mRNA structures. One of these mutations hampered a protein's domain that was not under selection, demonstrating the evolutionary flexibility of multi-domain proteins as one domain functionality was improved at the expense of the other domain. The *Trans* adaptations resided in two proteins, Npl3 and Gbp2, that bind pre-mRNAs and are central to their maturation. Interestingly, these mutations either increased or decreased the affinity of these proteins to mRNA, presumably allowing faster spliceosome recruitment or increased time before degradation of the pre-mRNAs, respectively. Altogether, our work reveals various mechanistic pathways toward optimizations of intron splicing to ultimately adapt gene expression patterns to novel demands.

## Introduction

Throughout evolution, cells acquired regulatory mechanisms to tune gene expression, which have been the subject of intensive investigations—focusing mainly on transcription and translation. Among other known mechanisms, when cells are challenged to increase protein expression levels, the DNA sequence of genes can change so as to increase transcription [1,2], support more efficient mRNA translation [3,4], or result in greater mRNA transcript stability

**Competing interests:** The authors have declared that no competing interests exist.

**Abbreviations:** kan, kanamycin resistance gene; UMI, unique molecular identifier; WT, wild-type.

[5,6]. Additionally, the transcription and translation machineries themselves have been shown to adapt to environmental challenges by altering the cellular pools of transcription factors [7] or tRNAs [8,9].

In evolving expression programs, adaptation often occurs either directly on the genes under pressure (“evolution in *Cis*”) [10] or indirectly, e.g., on the expression machineries, typically transcription and translation (“evolution in *Trans*”)[11,12]. These two routes of evolution are profoundly different [13], as the first (*Cis*) provides a localized solution that in principle can affect only a certain gene, while the later (*Trans*) could be the method of choice if a coordinated change in many genes is needed.

Surprisingly, although the process of splicing is central to the maturation and regulation of mRNAs in eukaryotes [14–18], its role in adapting to novel demands on gene expression has not been thoroughly investigated. During mRNA splicing, precursor mRNAs are processed to remove introns while fusing exons together to create the mature transcript. This process can provide evolutionary means to diversify the proteome towards phenotypic novelty, as the choice of intron to be excluded, as well as the exons which are found in the mature transcript, can both be regulated based on the cell’s needs [16,19,20]. An aspect of splicing evolution that has been extensively studied is gain and loss of introns, for which several molecular models have been proposed, mainly reverse transcription and recombination-mediated intron loss, intron transposition, and also exonization and intronization via mutations [21–25].

While intron loss and gain have been demonstrated experimentally [26,27], other forms of evolution through changes in splicing, such as alterations in splicing efficiency under changing conditions, have not. Adaptation of splicing efficiency is presumably essential to cellular evolution given a recent finding that splicing efficiency increases with transcription rate [28], therefore making it likely that splicing efficiency of introns is under constant selection during evolution. Yet, the mechanisms that allow this adaptation are unknown.

Here, we set to reveal whether introns or the splicing apparatus can evolve so as to alter the expression levels of genes in an adaptive manner. To this end, we engineered yeast cells to express a reporter gene, to which we inserted an inefficiently spliced intron that was fused to an antibiotic resistance gene. We then carried out a lab-evolution experiment in which cells were exposed to the drug and followed their adaptation.

Our results demonstrate that adaptations were related to splicing, and they appear to have not addressed directly the transcription or translation of the gene under selection. Two alternative adaptive routes for evolution of splicing were observed. First, we found *Cis*-acting solutions in the form of adaptive mutations that occurred in the intron itself but also, surprisingly, in an upstream exon. These mutations resulted in increased splicing efficiency and higher expression levels of the antibiotic resistance gene. We then show how one such *Cis* mutation alters the predicted RNA structure of the intron to better support splicing.

Yet, in some other evolved cells there were no mutations in *Cis*, i.e., in the gene or in its surrounding regions, but rather *Trans*-acting adaptations that have increased cellular availability of the splicing machinery. Sequencing the genomes of *Trans*-evolved colonies revealed nonsynonymous mutations in the RNA recognition motifs of two SR-like proteins that are known to have a diverse set of cellular functions related to RNA splicing and maturation. In particular, SR-like proteins were shown to support splicing by co-transcriptional recruitment of splicing factors [29–32] and were also shown to be involved in quality control of nascent mRNAs by selectively exporting from the nucleus spliced mRNAs upon completion of splicing [33,34]. Here, we show that adaptations in *Trans* that occurred through this experiment have modified the affinity of these proteins to the transcript under selection in a way that could allow its more efficient splicing.

## Results

### Low splicing efficiency of a drug resistance gene leads to stressed cells in presence of antibiotics

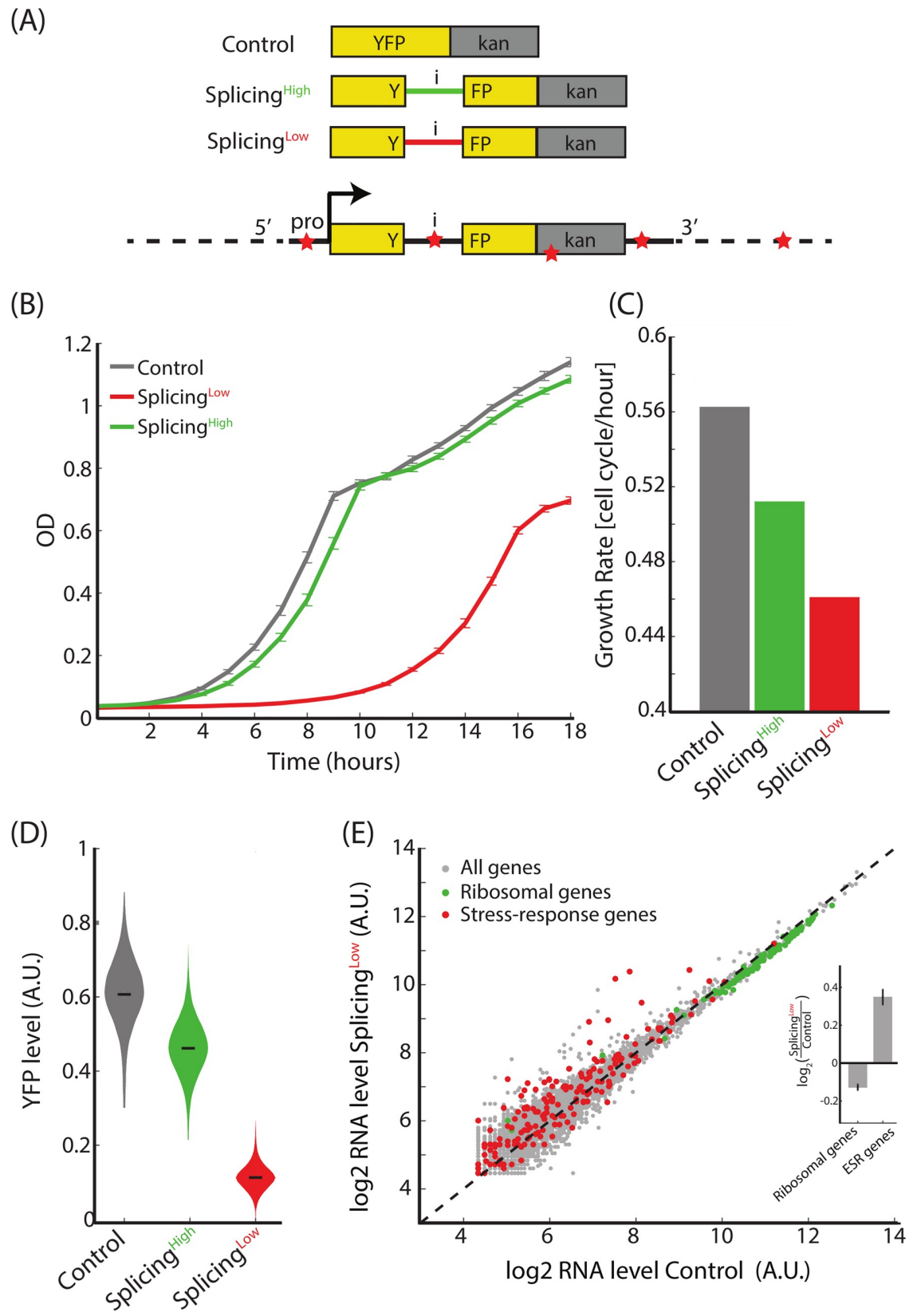
We hypothesized that splicing efficiency of genes could serve as a means to optimize their expression levels. To test this hypothesis, we used the yeast *Saccharomyces cerevisiae*, in which approximately 30% of the transcriptome is spliced to form mature mRNAs [35] at a range of splicing efficiencies [18,36]. We built a construct that consists of two fused domains. The first is a fluorescent reporter (YFP) that was engineered to include one of two alternative natural introns—with either high or low splicing efficiency. The intron was localized near the YFP's fluorescence site [36]. Downstream to this protein was fused an antibiotics resistance gene (kanamycin resistance gene [kan]). This general design consisted of three alternative strains: (i) "Control" with a YFP-Kan construct without an intron; (ii) "Splicing<sup>High</sup>" with a YFP-Kan gene that harbors the natural intron of *OSH7* that was previously reported to have high splicing efficiency within this YFP context [36]; and (iii) "Splicing<sup>Low</sup>" with a YFP-Kan gene that harbors the natural intron of *RPS26B*, with a low splicing efficiency [36] (see Fig 1A and S1 Table for a list of strains used in this study).

We first hypothesized that cellular growth of each strain in the presence of the antibiotic G418 will depend on YFP-Kan expression levels. We followed the growth of the three strains in the presence of the antibiotics and found that Control cells had the highest fitness, Splicing<sup>High</sup> grew slower, and Splicing<sup>Low</sup> demonstrated a severe growth defect compared with the two other strains (Fig 1B and 1C). We also measured fluorescence intensity of the YFP-Kan reporter in the presence of the drug and observed that Control cells demonstrated the highest fluorescence levels, followed by Splicing<sup>High</sup>, and with Splicing<sup>Low</sup> cells showing the lowest YFP-Kan levels (Fig 1D). These results demonstrate that the inefficiently spliced intron in Splicing<sup>Low</sup> reduces cellular levels of YFP-Kan and hence, presumably, leads to a reduced fitness.

Because YFP-Kan expression levels in Splicing<sup>Low</sup> were significantly lower compared with the other strains, we hypothesized that Splicing<sup>Low</sup> cells did not reach the needed concentration of the resistance protein to sufficiently neutralize the antibiotics, and hence resulted in stressed cells. To test this hypothesis, we performed mRNA sequencing of exponentially growing Control and Splicing<sup>Low</sup> cells in an antibiotics-containing medium and analyzed the transcriptome profiles of these cells. Indeed, we observed that ribosomal genes were down-regulated in Splicing<sup>Low</sup> compared with Control cells—a clear signature of stressed cells [37] (Fig 1E). Notably, we observed an averaged 8% reduction for mRNA levels of ribosomal proteins between Control and Splicing<sup>Low</sup> cells and, correspondingly, an 18% reduction in growth rate. Interestingly, this is consistent with the correlation observed in a recent study between growth rate and ribosomal expression levels in yeast cells [38]. In parallel, stress-related genes [39] were up-regulated in Splicing<sup>Low</sup> cells compared with Control cells (Fig 1E). We thus concluded that the general stress response was activated in Splicing<sup>Low</sup> cells.

### Rapid evolutionary adaptation increases expression level of the resistance gene

Our experimental system mimics an evolutionary scenario in which there is an immediate and continuous selection pressure to up-regulate the expression level of a specific gene in a particular environment. How would the system evolve to better resist the antibiotics? Possible means to adapt include mutations in the gene's promoter to increase transcription, mutations that increase translation initiation or efficiency, or mutations inside the gene itself that could increase the specific activity of the protein (Fig 1A). Additionally, the splicing machinery may



**Fig 1. Inefficient intron splicing leads to lower gene expression levels and compromised antibiotics resistance.** (A) We introduced two alternative introns into a YFP domain that was fused to a kanamycin resistance domain, to generate three strains: (i) Control without an intron; (ii) Splicing<sup>High</sup> with an efficiently spliced intron; and (iii) Splicing<sup>Low</sup> with an inefficiently spliced intron. Evolving cells at the presence of the antibiotics could adapt by mutating different parts of the YFP-Kan construct (evolution in *Cis*) or other loci, evolution in *Trans* (red stars represent potential locations of such putative mutation sites). (B,C) Splicing<sup>Low</sup> suffers from a severe growth defect compared with Control or Splicing<sup>High</sup> cells when the antibiotic is supplemented to the medium. The growth defect is manifested as both a longer lag phase and a lower maximal growth rate. (D) Florescence intensity of the YFP-Kan reporter for all three strains shows that Splicing<sup>Low</sup> cells have lower expression levels of YFP-Kan. This observation links between YFP-Kan expression levels and cellular fitness. (E) Transcriptome profiling shows that ribosomal genes were down-regulated (green dots,  $p = 4.62 \times 10^{-26}$ , paired *t* test) and stress-response genes were up-regulated (red dots,  $p = 3.40 \times 10^{-5}$ , paired *t* test) in Splicing<sup>Low</sup> compared with Control cells. This observation suggests that Splicing<sup>Low</sup> cells experience stress because of compromised resistance to the antibiotics and that the general stress response was activated in them. (Inset) Mean log<sub>2</sub> ratio of ribosomal and ESR gene groups. See numerical data for this figure in [S1 Data](#).

<https://doi.org/10.1371/journal.pbio.3000423.g001>

also take part in adaptation of gene expression levels. To find which evolutionary tracks are used by cells as they adapt, we evolved the three strains by daily serial dilution on a medium supplemented with G418 for approximately 560 generations, in four independent cultures for each strain. Interestingly, only the cultures of Splicing<sup>Low</sup> cells demonstrated a significant improvement in fitness when grown under the drug at the end of the experiment ([Fig 2A and 2B](#)). This observation suggests that only Splicing<sup>Low</sup> experienced a sufficiently strong selective pressure to adapt to the presence of the antibiotics in the medium, in contrast to the Control and Splicing<sup>High</sup> strains, which originally had much higher levels of the resistance proteins.

Consistent with the fitness measurements, YFP measurements of the evolved cultures showed that expression levels of the YFP-Kan fusion gene increased in all four evolved cultures of Splicing<sup>Low</sup> compared with the ancestral strain ([Fig 2C](#)). Conversely, the increase in YFP-Kan expression levels in the evolved Control and Splicing<sup>High</sup> populations was significantly smaller ([Fig 2C](#)). These results further indicate that Splicing<sup>Low</sup> cells experienced the strongest selective pressure to adapt rapidly to the presence of the antibiotics, and that they achieved this goal by increasing the levels of the YFP-Kan reporter. We next moved to reveal the molecular mechanisms underlying this evolutionary adaptation.

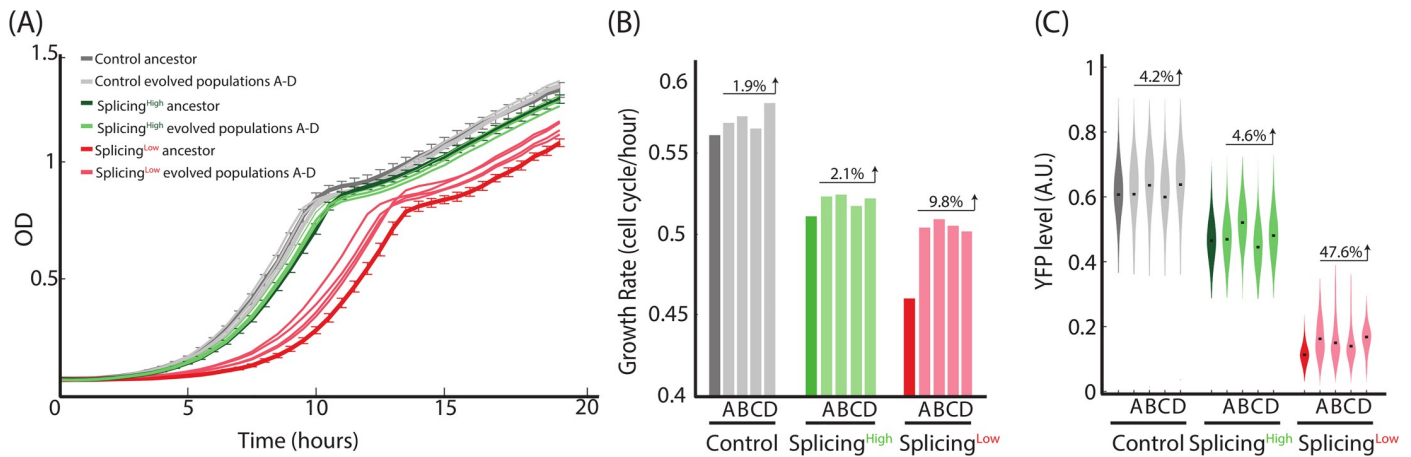
### Adaptations in both *Cis* and *Trans* lead to increased splicing efficiency

We hypothesized that improving the low splicing efficiency of the intron in Splicing<sup>Low</sup> could be natural selection's means to adapt towards increasing the resistance gene expression levels. We therefore sequenced the YFP-Kan locus, covering the entire gene from promoter to terminator, in 16 colonies from two evolved populations (termed here population A and population B) of Splicing<sup>Low</sup>. Interestingly, we found that the colonies were split into two types—either with or without a mutation in the YFP-Kan locus ([Fig 3A](#)). In population A, we found that the same mutation occurred in four out of eight colonies, changing adenine to cytosine inside the intron, 97 nucleotides upstream to its 3' end ([Fig 3B](#)). In population B, we identified an exonic nonsynonymous mutation that changed a thymine to cytosine 14 nucleotides upstream of the intron (V61A change in the YFP protein) in three out of eight colonies. In five other colonies from this population there were no mutations in the YFP-Kan locus.

Notably, none of the colonies demonstrated a mutation in the construct's promoter, terminator, or in the sequence of the Kan resistance gene itself. These results propose that different mutations in the intron, or its vicinity, were adaptive and might affect splicing efficiency of the intron. Surprisingly, the observed mutations did not occur in the 5' donor, 3' acceptor, nor in the intron branch point—suggesting that other positions of the intron can also be functional by affecting splicing efficiency.

While the intron- and exon-mutated colonies represent an evolutionary adaptation in *Cis*, the colonies that showed no mutation in the entire gene construct, that coexist with the *Cis*-





**Fig 2. Rapid adaptation to the presence of the antibiotics is observed only for Splicing<sup>Low</sup> cells.** (A,B) We evolved Control, Splicing<sup>High</sup>, and Splicing<sup>Low</sup> cells for approximately 560 generations with the presence of the antibiotics in four independent cultures for each strain. Growth measurements of evolved populations compared with the three ancestors shows that only evolved Splicing<sup>Low</sup> cells demonstrate significant improvement in growth for all four independent evolution lines. The number above each group of evolved populations represents the average improvement in growth rate compared with these populations' ancestor strains. These observations suggest that the inefficiently spliced intron led to a rapid adaptation of Splicing<sup>Low</sup> cells. (C) Fluorescence intensities of the YFP-Kan reporter for all evolved cultures show that expression levels were much increased in all four evolved cultures of Splicing<sup>Low</sup> compared with the ancestral strain (effect sizes = 78.67, 79.54, 75.17, 83.19). Conversely, the increase in expression levels in the evolved Control and Splicing<sup>High</sup> populations were smaller (Control effect sizes = 64.66, 68.44, 63.51, 67.74; Splicing<sup>High</sup> effect sizes = 54.33, 70.66, 52.43, and 58.27). The number above each group of evolved populations represents the average increase in YFP-Kan levels compared with these populations' ancestor strains. These observations suggest that adaptation of Splicing<sup>Low</sup> cells was based on their ability to increase expression levels of the resistance proteins. See numerical data for this figure in [S1 Data](#).

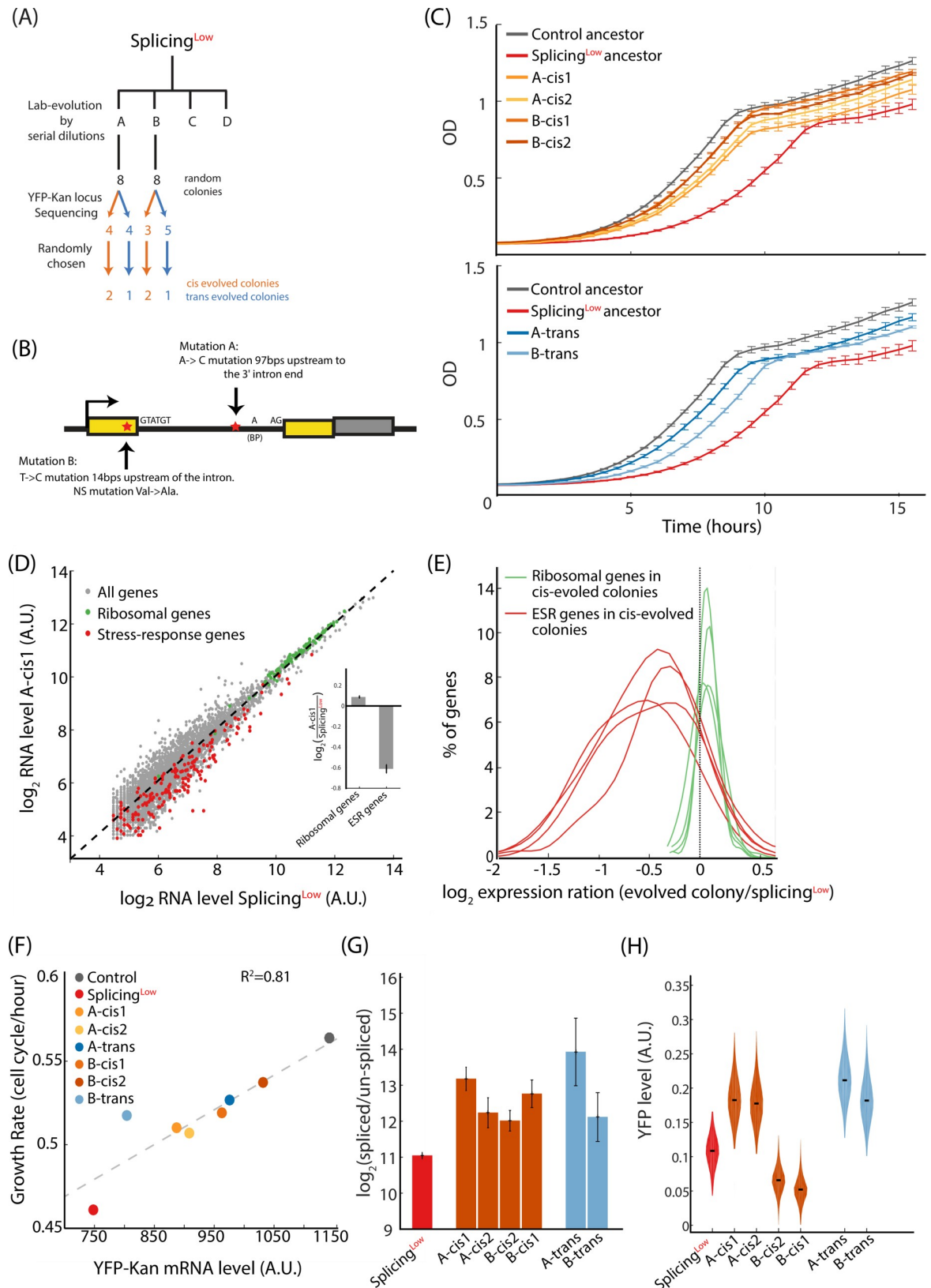
<https://doi.org/10.1371/journal.pbio.3000423.g002>

evolved colonies in the same populations, potentially found adaptive solutions in *Trans* that may have occurred elsewhere in the genome. We thus randomly chose six colonies: four colonies with a *Cis* mutation and two colonies that showed no mutations in *Cis*, as we reasoned that such colonies may have adapted in *Trans*. We termed these colonies according to the evolution lines from which they were derived: A-cis1, A-cis2, B-cis1, B-cis2, A-trans, and B-trans. We followed the growth of these evolved colonies in the presence of G418 and found, as expected, that all grew faster than the Splicing<sup>Low</sup> ancestor (Fig 3C).

We then performed mRNA sequencing and transcriptome analyses of all colonies and found that they indeed demonstrate relaxation of the stress state that was featured in the ancestor. In particular, the general stress response genes were reduced relative to their high levels in the Splicing<sup>Low</sup> ancestor, and ribosomal proteins were up-regulated relative to their low levels in this ancestral strain (Fig 3D for colony A-cis1). These dynamics were shared by all the *Cis*-evolved colonies (Fig 3E), demonstrating the robustness of this observation. Specifically, in all four *Cis*-evolved colonies, the expression levels of most ribosomal genes were increased compared with the Splicing<sup>Low</sup> ancestor (percentage of up-regulated ribosomal genes in the four colonies, 86%, 63%, 83%, and 84%) and levels of the ESR genes were mostly reduced compared with the ancestor (percentage of down-regulated ESR genes in four colonies, 90%, 76%, 87%, and 84%). These findings suggest that indeed these colonies adapted to the presence of the antibiotics in the environment and that the stress experienced by them was partially alleviated.

We next hypothesized that cellular fitness might correlate with mRNA levels of the YFP-Kan construct because increased transcript levels should result in higher concentrations of the YFP-Kan protein. Indeed, maximal growth rates of the Control and Splicing<sup>Low</sup> ancestors and of the six evolved colonies correlate with mRNA levels of their YFP-Kan construct, as deduced from the RNA-seq (Fig 3F).

Because the observed *Cis* mutations occurred at the vicinity of the intron, we hypothesized that they increased splicing efficiency of the YFP-Kan transcript. To test this possibility, we



**Fig 3. Evolved colonies demonstrate increased splicing efficiency that results in higher transcript levels and relieved stress.** (A) We randomly chose 16 colonies in total from two evolved lines of Splicing<sup>Low</sup> and sequenced the YFP-Kan locus of those colonies. We found

that approximately half had mutations in the YFP-Kan construct (indication of evolution in *Cis*) and the other half did not (indication of evolution in *Trans*). Of those colonies, we randomly chose two *Cis*-evolved and one *Trans*-evolved colonies from each evolved population for further examination. (B) Sequencing of the YFP-Kan construct in the evolved colonies revealed two mutation types: (i) in the intron itself and (ii) in the upstream exon. These mutations did not occur in the intron 5' donor, 3' acceptor, or the branching point—suggesting that other positions of the intron and its vicinity are functional and may affect splicing. (C) All *Cis*-evolved colonies (upper graph) and *Trans*-evolved colonies (lower graph) show increased fitness compared with the Splicing<sup>Low</sup> ancestor, yet still lower than the Control ancestor. (D) Transcriptome profiling reveals that ribosomal genes were up-regulated (green dots,  $p = 4.94 \times 10^{-18}$ , paired *t* test) and stress-related genes were down-regulated (red dots,  $p = 3.64 \times 10^{-15}$ , paired *t* test) in the evolved colony A-cis1 compared with the Splicing<sup>Low</sup> ancestor. (Inset) Mean log<sub>2</sub> ratio of ribosomal and ESR gene groups. (E) The four *Cis*-evolved colonies show similar trends, i.e., increased expression levels of ribosomal genes and decreased expression levels of stress-response genes ( $p$ -values for all cases  $< 0.005$ , paired *t* test). These observations suggest that the stress experienced by the evolved colonies was alleviated during their adaptation to the antibiotics in the medium. (F) mRNA levels of YFP-Kan transcripts correlate with growth rates ( $R^2 = 0.82$ ,  $p = 0.0023$ ), suggesting that cellular fitness in our setup is indeed determined by the availability of kanamycin resistance proteins to overcome the antibiotics. (G) All *Cis*- and *Trans*-evolved colonies demonstrate increased splicing efficiency of the YFP-Kan mRNA compared with the Splicing<sup>Low</sup> ancestor ( $p < 0.05$  for all colonies compared with Splicing<sup>Low</sup> ancestor). This result suggests that all adaptation trajectories led to the adaptation of the splicing process to better mature the un-spliced YFP-Kan transcript. (H) Fluorescence intensity of the YFP-Kan reporter shows increased levels for the two *Cis*-evolved colonies with the mutation in the intron and for the two *Trans*-evolved colonies. In contrast, the two *Cis*-evolved colonies with the nonsynonymous mutation in the exon demonstrate decreased YFP-Kan levels. This observation suggests that the nonsynonymous mutation hampered the ability of the YFP domain to fluorescent and reduced the fluorescence intensity per protein molecule (see text for full explanation). See numerical data for this figure in [S1 Data](#).

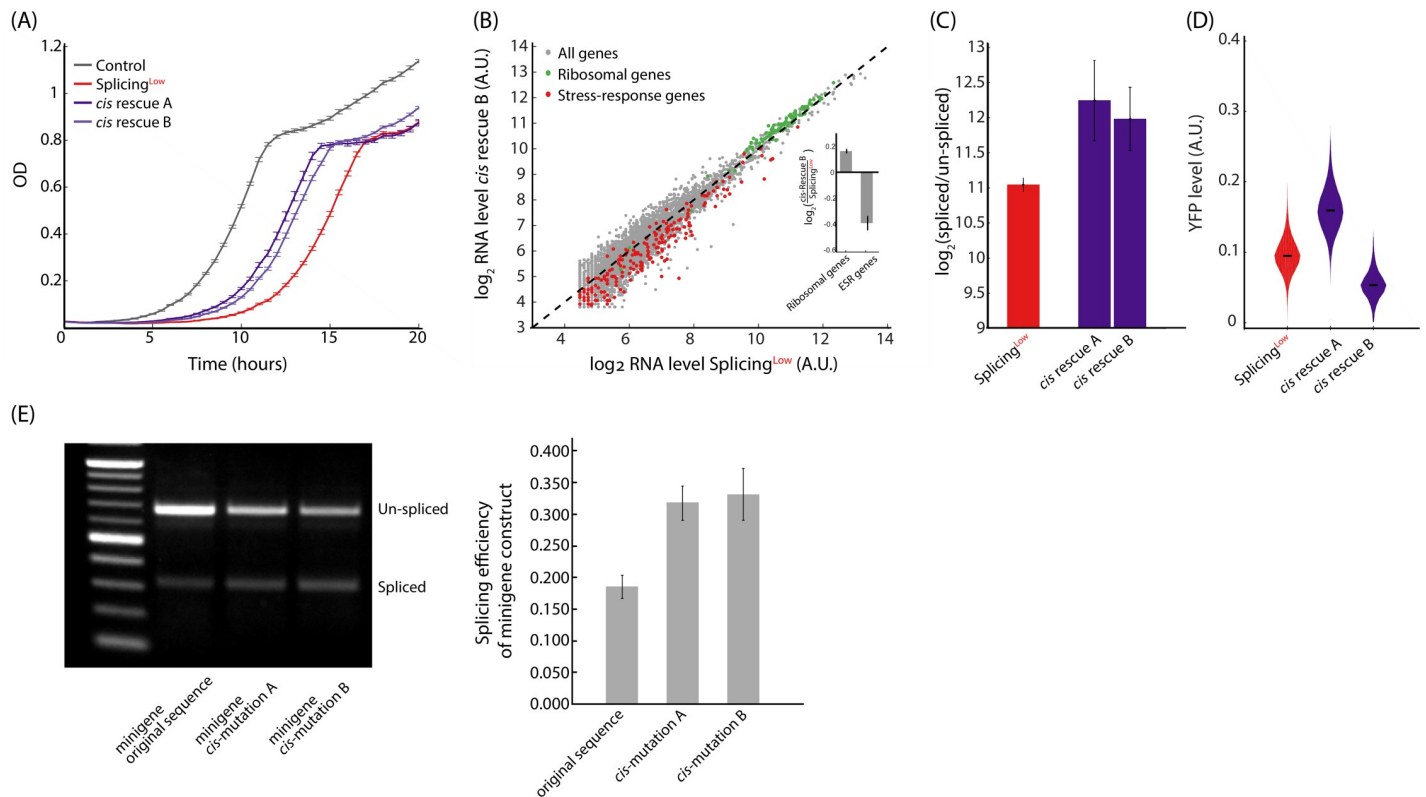
<https://doi.org/10.1371/journal.pbio.3000423.g003>

assayed splicing efficiency for both *Cis*- and *Trans*-evolved colonies with qPCR, targeting the un-spliced or spliced transcript versions. Interestingly, the ratio of spliced to un-spliced transcripts, a measure of splicing efficiency, was higher in all evolved colonies compared with the Splicing<sup>Low</sup> ancestor, suggesting that at least some of the increase in mRNA level we observed in the evolved colonies results from higher splicing efficiency (Fig 3G).

To prove that adaptation of the colonies actually led to higher protein levels of the fluorescence-resistance fused protein, we measured fluorescence intensity using flow cytometry. We found that the two *Cis*-evolved colonies from population A (A-cis1 and A-cis2) and the two *Trans*-evolved colonies (A-trans and B-trans) showed higher YFP-Kan levels compared with the ancestor. In contrast, the two *Cis*-evolved colonies from population B (B-cis1 and B-cis2) demonstrated decreased fluorescence intensity values relative to the ancestor (Fig 3H). These observations indicate that the nonsynonymous, exon mutation reduced the per-protein fluorescence value of the YFP component of the YFP-Kan construct in these colonies. Indeed, this exonic mutation occurred in a position that was recently reported to reduce fluorescence when mutated in the highly similar GFP [40]. Because YFP's functionality, i.e., fluorescence, was not selected for or against in our setup, it appears to have been free to obtain mutations that help achieve a higher expression level of the entire fusion construct. It thus seems that a modular domain architecture of a protein may increase its evolvability under relevant conditions, as it allows the trade-off and optimization of one domain at the expense of another.

We next wanted to confirm that the fitness gained in the *Cis*-evolved colonies was indeed due to these mutations. It is still possible that additional beneficial mutations exist in the genome of the *Cis*-evolved colonies, which account for at least part of the improved fitness we observed in these colonies. We thus generated two rescue strains, termed *cis*-Rescue-A and *cis*-Rescue-B, in which these *Cis*-acting mutations were introduced individually to the ancestral Splicing<sup>Low</sup> background. Notably, the two rescue strains grew better than Splicing<sup>Low</sup> cells in the presence of the antibiotics, although not as well as Control cells (Fig 4A). Additionally, the stress experienced by Splicing<sup>Low</sup> cells, as observed by changes in expression levels of ribosomal and stress genes, was alleviated upon insertion of each individual *Cis* mutation (Fig 4B). Then, we measured splicing efficiencies and fluorescence intensity levels for both rescue strains and found that they resembled the levels shown in the evolved single colonies (Fig 4C and 4D, in comparison with Fig 3G and 3H).





**Fig 4. Cis-acting mutations are sufficient to increase fitness by elevating splicing efficiency.** (A) We created two *Cis*-rescue strains, each harboring one of the mutations that appeared spontaneously in the evolved populations. Growth of the two *Cis*-rescue strains show that a single mutation in the YFP-Kan construct is sufficient to increase fitness compared with *Splicing*<sup>Low</sup>. (B) The exonic mutation is also sufficient to alleviate stress, as ribosomal genes were up-regulated (green dots,  $p = 1.02 \times 10^{-18}$ , paired *t* test) and stress-related genes were down-regulated (red dots,  $p = 9.02 \times 10^{-12}$ , paired *t* test) in *cis*-Rescue-B compared with *Splicing*<sup>Low</sup>. The same trend was also observed for the intronic mutation for *cis*-Rescue-A cells. (Inset) Mean  $\log_2$  ratio of ribosomal and ESR gene groups. (C) The two *Cis*-rescue strains demonstrate higher splicing efficiency of the YFP-Kan mRNA compared with the *Splicing*<sup>Low</sup> ancestor ( $p < 0.05$ ). This result suggests that a single mutation is sufficient to improve splicing efficiency. (D) Fluorescence intensity of the YFP-Kan reporter for the *cis*-Rescue-A and *cis*-Rescue-B strains show similar trends as the colonies in Fig 3D—supporting earlier conclusions. (E) The effects of *Cis* mutations on splicing tested with a mini-gene approach. We cloned the intron's original sequence and its two mutated versions together with 200 bps surrounding it into a high-copy number plasmid. RT-PCR assays for WT (BY4741) cells transformed with these plasmids show that, even in this context, the *Cis* mutations are sufficient to increase splicing efficiency of the intron for both mutated intron versions compared with the original sequence ( $p < 0.05$ , *t* test). See numerical data for this figure in S1 Data. WT, wild-type.

<https://doi.org/10.1371/journal.pbio.3000423.g004>

Our data so far demonstrate that the *Cis* mutations we identified in our lab-evolution experiment are beneficial because they improve splicing efficiency in the context of the entire YFP-Kan transcript. To further demonstrate that indeed these mutations increase splicing efficiency and to better exclude other mechanisms that might affect the YFP-Kan transcript, we used a mini-gene approach, a widely used methodology in splicing studies [41]. In this approach, the intron together with its adjacent exons are cloned to a new context, and splicing efficiency is measured. To this end, we cloned the original sequence of the intron plus 200 bases up- and downstream of it into a 2 $\mu$  plasmid to a locus driven by a strong promoter. We additionally created two mutated versions of this construct, each with one of the two *Cis* mutations we discovered. Because there are only two short exons surrounding the intron in this system, it is ideal to explore the effects of the mutations on splicing. We thus harvested mRNA from exponentially growing cells and measured the splicing efficiency of each variant using PCR, with primers flanking the intron that simultaneously amplify both spliced and un-spliced versions of the YFP-Kan transcripts. In agreement with the splicing assay done for the full gene (Fig 4C), in the mini-gene assay too, both variants with either mutation *cis* A or B showed increased splicing efficiency compared with the ancestral versions (Fig 4E).

Our results thus far provide direct evidence that intron splicing takes part in the adaptation and optimization of gene expression patterns to environmental needs. Although intron sequences are much less conserved, compared with exons, and are believed to be less functional, we demonstrate that their sequence can be used by natural selection as a molecular mechanism to regulate splicing efficiency and adjust gene expression patterns.

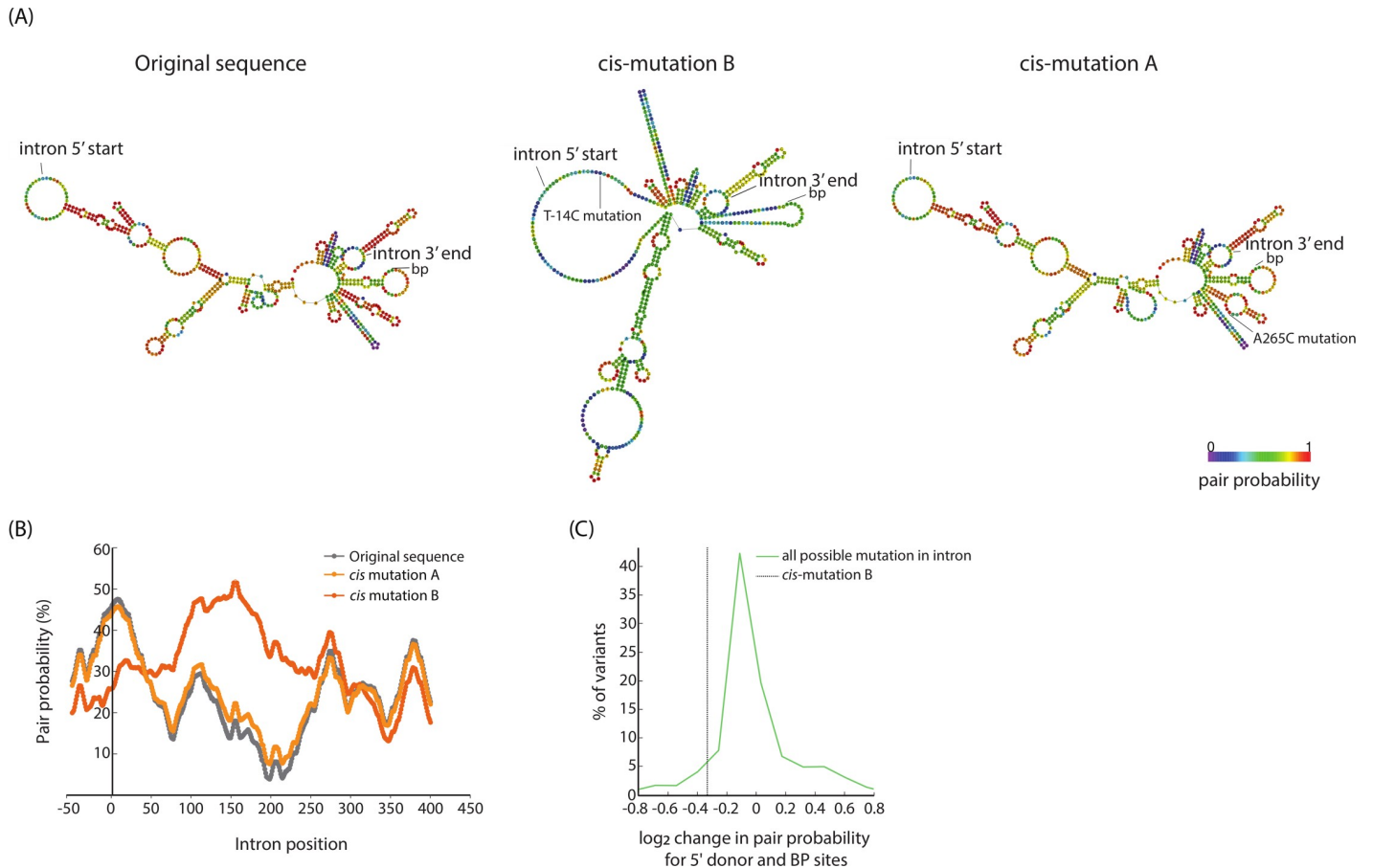
### ***Cis* mutation is adaptive through effects on mRNA structure that make 5' donor and branching point sites more accessible to splicing**

How can the mutations we identified in the intron facilitate splicing? One possibility is that these changes favorably alter the RNA structure of the un-spliced YFP-Kan transcript by making the splicing sites more accessible to the spliceosome. Indeed, RNA structures can inhibit or facilitate binding of spliceosome components to the pre-mRNA and affect splicing efficiency [42,43]. We thus computationally modeled the RNA structure of the intron and 50 bases on both its sides using the ViennaRNA algorithm [44]. We performed this analysis for the original and the two mutated sequences of the intron. Interestingly, mutation *cis* B, located near the 5' donor site, leads to massive changes in the predicted RNA structure, notably causing the structure near the 5' donor and the branching point sites to loosen (Fig 5A). Specifically, the pairing probability, a prediction for how likely it is for a position along the RNA molecule to associate with other positions, is decreased at the 5' donor and branching point positions between the original and the mutated sequence (Fig 5B). How likely is a single point mutation to change so drastically the predicted structure of an RNA? To ask that, we constructed a simple null model in which we calculated the predicted pair-probability difference between the original sequence and each of the other possible single nucleotide mutations in the intron. Notably, mutation *cis* B falls among the 5.5% of all mutations with the highest predicted potential to affect the secondary structure; i.e., it is among the 5.5% of point mutations that loosen the RNA secondary structure near the 5' donor and branching point the most (Fig 5C). These observations suggest a model in which mutation *cis* B may facilitate splicing due to increased accessibility of the splicing machinery to the functional splicing sites. In contrast, mutation *cis* A did not show similar patterns to mutation *cis* B (Fig 5A and 5B), raising the possibility that a different, still obscure, mechanism is causing the beneficial effects of this mutation.

### **Evolution in *Trans*: Increasing cellular availability of the splicing machinery can be adaptive**

We next aimed to decipher the mechanism behind the increased YFP-Kan levels in the *Trans*-evolved colonies that showed no mutations in *Cis*, i.e., within the reporter gene or in its vicinity. We reasoned that elevating availability of the splicing machinery could be a means to increase splicing efficiency of the YFP-Kan transcript and thus could be used as an adaptive mechanism to the antibiotics challenge. As with other cellular machineries whose functioning depends on supply-to-demand economy [4,8,45–47], increased splicing availability could be achieved by either increasing the expression of the splicing machinery genes or decreasing expression levels of other intron-containing genes, which collectively constitute the “demand” for the splicing machinery.

To test if any of these evolutionary routes were indeed taken by the evolved cells, we calculated the expression level ratio of genes between the evolved colonies and their Splicing<sup>Low</sup> ancestor. In colony *A-trans*, we observed increased expression ratio of splicing machinery genes (the “supply”) and decreased expression ratio of non-ribosomal intron-containing genes (the “demand,” Fig 6B, and S1 Fig for other colonies). In contrast, both supply and demand genes show similar expression levels between Splicing<sup>Low</sup> and Control ancestors, suggesting

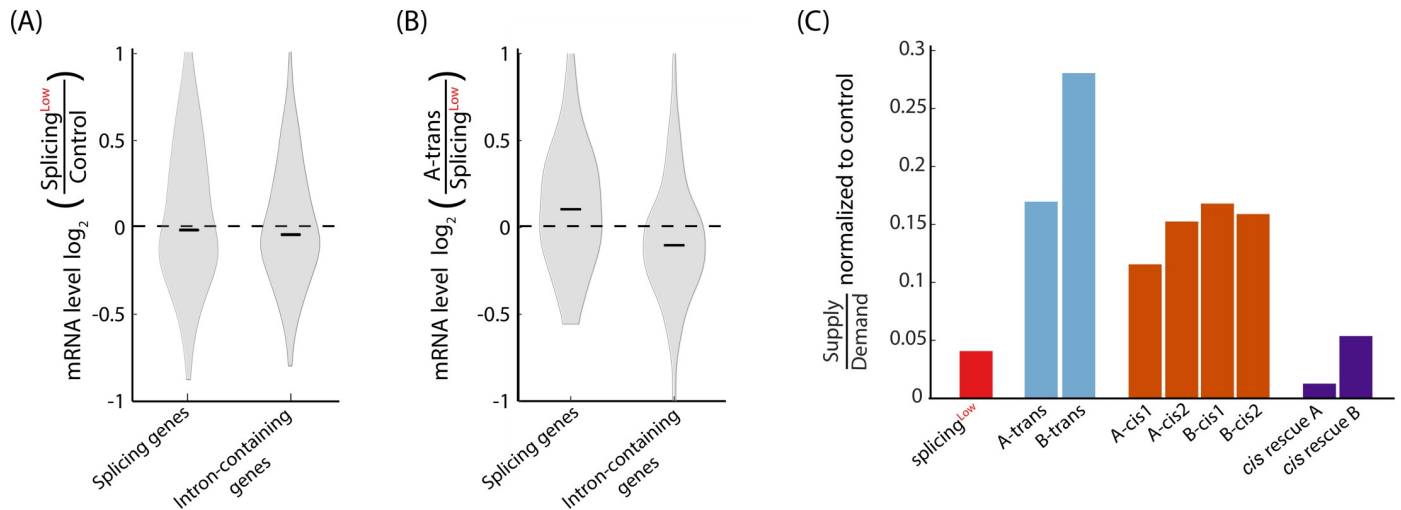


**Fig 5. A secondary structure model of the effects of *Cis* mutation on splicing efficiency.** (A) RNA secondary structure predictions using Vienna algorithm of the intron plus 50 bps of its surrounding exons for the original sequence and its two mutated versions. *Cis*-mutation B reshapes the predicted secondary structure and lowers the base pairing probability around the 5' donor and branching point sites. (B) Base pairing probability at all positions as calculated by the Vienna algorithm for the original intron sequence and its two mutated versions. Position of 5' donor site is 0, position of branching point site is 328, position of 3' acceptor site is 361, position of mutation A is 265, position of mutation B is -14. (C) Probability distribution of the log<sub>2</sub> of change in base pairing probability along a 20-base window at the 5' donor and branching point sites for all possible single nucleotide variations on the intron's sequence compared with the original sequence. Vertical line denotes the value of the observed *cis*-mutation B, which is at the bottom 5.5 percentage of the histogram; i.e., it is among mutations that mostly reduce the base pair probability at these sites.

<https://doi.org/10.1371/journal.pbio.3000423.g005>

that a mere physiological adaptation did not confer upon the Splicing<sup>Low</sup> ancestor the ability to adjust supply and demand to the intron burden (Fig 6A and see Discussion).

We then computed changes in “supply-to-demand splicing availability” by summing up for each strain the expression levels of the splicing machinery genes and intron-containing genes, separately, then dividing these values and normalizing the ratio with that of the control strain. Interestingly, the supply-to-demand difference has increased appreciably in the *Trans*-evolved colonies compared with the Splicing<sup>Low</sup> ancestor. The *Cis*-evolving colonies have also improved their supply-to-demand relative to the ancestor, indicating that they too may have evolved in *Trans* in addition to their *Cis* adaptations. Thus, we concluded that both *Cis* and *Trans* adaptation routes can co-occur in the same genome towards optimization of its gene expression patterns. Notably, the group of non-evolved strains (Splicing<sup>Low</sup> and the two *Cis*-rescue strains) showed lower supply-to-demand ratios compared with the group of evolved strains (both *Cis*- and *Trans*-evolved colonies) (Fig 6C).



**Fig 6. Increasing cellular availability of the splicing machinery is an adaptive mechanism of splicing.** (A) The groups of splicing genes (splicing supply) and intron-containing genes (splicing demand) show similar levels between Splicing<sup>Low</sup> and Control ancestors ( $p > 0.05$ , paired  $t$  test for both gene groups). (B) The groups of splicing genes and intron-containing genes were increased ( $p = 1.36 \times 10^{-3}$ , paired  $t$  test) and decreased ( $p = 1.67 \times 10^{-2}$ , paired  $t$  test), respectively, in colony A-trans compared with the Splicing<sup>Low</sup> ancestor. This observation suggests that the supply-to-demand ratio of the splicing machinery was increased in the A-trans colony, which allowed its increased splicing efficiency of the YFP-Kan transcript. (C) Supply-to-demand ratios for the splicing machinery were calculated to Control and Splicing<sup>Low</sup> ancestors, to all evolved colonies, and to the Cis-rescue strains as the sum expression level of all splicing-related genes over the sum expression level intron-containing genes. Values of all strains were then normalized to the value of the Control ancestor. Importantly, supply-to-demand ratios are similar for all strains that did not evolve (Control and Splicing<sup>Low</sup> ancestors and the two rescue strains) and were increased for all evolved colonies ( $p = 0.005$  for difference in supply-to-demand ratios between evolved and non-evolved strains,  $t$  test). These results suggest that indeed the cellular availability of the splicing machinery was elevated due to an evolutionary adaptation process and not because of other physiological mechanisms (see Discussion). See numerical data for this figure in S1 Data.

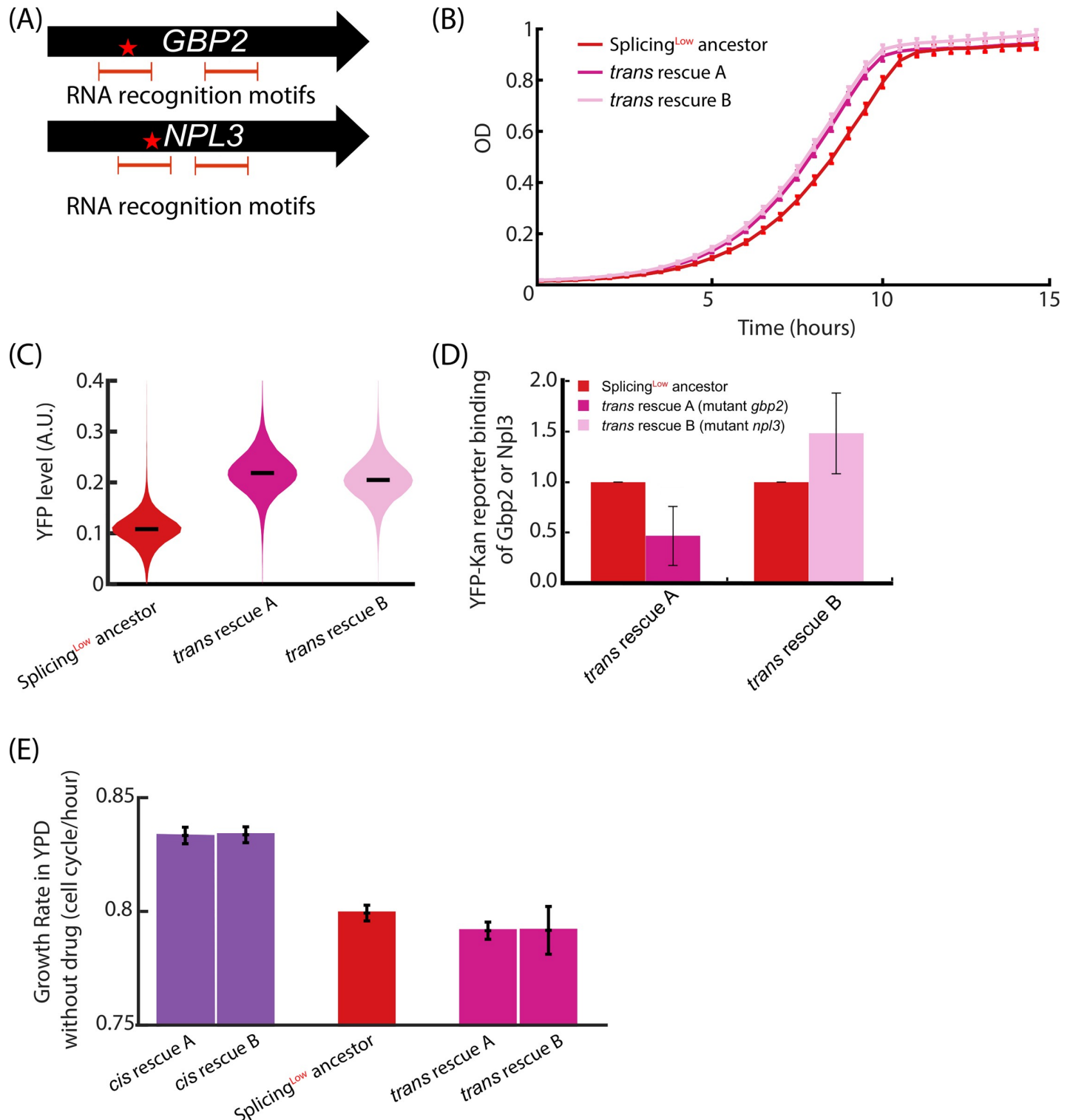
<https://doi.org/10.1371/journal.pbio.3000423.g006>

Taken together, we argue that increased cellular availability of the splicing machinery is the result of an evolutionarily adaptive process, which might have allowed for the improved splicing efficiency of the YFP-Kan gene in the evolved colonies. While genomic sequencing revealed interesting mutations in *Trans* (see below), we did not find mutations that can explain these supply-to-demand changes in a straightforward way.

### Adaptation in *Trans* is also based on mutations in SR-like proteins

We next aimed to reveal which genetic changes may have occurred in *Trans* to the fusion gene. To that end, we fully sequenced the genomes of the two *Trans*-evolved colonies (A-trans and B-trans) and identified, respectively, three and two mutations in each of their genomes (see S2 Table for list of mutations). Interestingly, in both colonies we found a mutation in a member of the SR-protein group in *S. cerevisiae*, which are splicing-related genes. Specifically, in colony A-trans we found a nonsynonymous mutation in *GBP2*, changing histidine into tyrosine, H160Y, and in colony B-trans we found a nonsynonymous mutation in *NPL3*, changing phenylalanine into valine, F160V (Fig 7A).

Npl3 was previously shown to be loaded onto pre-mRNAs during early stages of transcription [48,49] and to support the recruitment of the spliceosome [29] after it has completed its quality control function to monitor correct 5' capping (Schneider and Krebber, in preparation). Unlike for Npl3, deletion of Gbp2 does not result in accumulation of pre-mRNAs [29]. In contrast, Gbp2 was shown to associate with the late-stage spliceosome and appears to function as a quality control factor that oversees successful splicing of pre-mRNAs, which facilitates the recruitment of the nuclear exosome to incorrectly spliced pre-mRNAs in order to eliminate these faulty transcripts [33].



**Fig 7. *Trans* adaptation is based on mutations in SR-like proteins that change their affinity to the YFP-Kan transcript.** (A) Genome sequencing of the *Trans*-evolved colonies revealed two mutations in the SR-like proteins *GBP2* and *NPL3*. Both nonsynonymous mutations, H160Y in *GBP2* and F160V in *NPL3*, occurred in an RNA recognition motif. (B) We created two *Trans*-rescue strains, termed *trans*-Rescue-A and *trans*-Rescue-B, each harboring one of the two mutations as described in A. Growth of these two *Trans*-rescue strains in the presence of the drug show that the *Trans* mutations are sufficient to increase fitness compared with Splicing<sup>Low</sup>. (C) Fluorescence intensities of the YFP-Kan reporter for *trans*-Rescue-A and *trans*-Rescue-B strains are higher compared with Splicing<sup>Low</sup> cells. (D) Binding of mutant *gbp2* and *npl3* to the YFP-Kan mRNA was examined and compared with the respective WT proteins. RT-PCR results showed that mutant *gbp2* has a reduced binding affinity



to the reporter ( $p = 2.97 \times 10^{-3}$ ), whereas mutant *npl3* binds with a slightly higher affinity ( $p = 9.54 \times 10^{-2}$ ). These opposite effects on binding can be rationalized by the different functions of *GBP2* and *NPL3*. See text for full details. (E) Growth rates of all four rescue strains and the Splicing<sup>Low</sup> ancestor under permissive conditions, without the drug. Notably, *Cis*-rescue strains demonstrate higher growth rates compared with Splicing<sup>Low</sup>—indicating that they alleviate the burden of the inefficiently spliced intron that is independent of the drug. In contrast, both *Trans*-rescue strains show similar growth rates compared with Splicing<sup>Low</sup>—indicating that mutations in SR-like proteins are associated with additional costs to cells that cancel out the decreased burden of the inefficiently spliced intron. See numerical data for this figure in S1 Data. WT, wild-type.

<https://doi.org/10.1371/journal.pbio.3000423.g007>

To examine whether these mutations in the two SR-like proteins are sufficient to increase fitness and YFP fluorescence of cells that harbor the low-efficiency spliced intron as part of the YFP-Kan construct, we used CRISPR technology (see [Methods](#)) to introduce these point mutations, each individually, to the background of the Splicing<sup>Low</sup> ancestor. We termed these engineered strains *trans*-Rescue-A and *trans*-Rescue-B. Indeed, these two new rescue strains grew faster on the G418 antibiotics compared with ancestor cells ([Fig 7B](#)), and they both showed higher fluorescence intensity levels of YFP ([Fig 7C](#)). Hence, it appears that these *Trans* mutations result in increased levels of the YFP-Kan proteins.

Notably, both mutations we identified in *GBP2* and *NPL3* occurred in one of the two RNA recognition motifs of their encoded proteins ([Fig 7A](#)), suggesting that their affinity to the YFP-Kan transcript may have changed. We therefore tested if the mutant proteins would show modified YFP-Kan RNA-binding affinities compared with their wild counterparts. Mutant or wild-type (WT) Gbp2 and Npl3 were each immunoprecipitated with specific antibodies and the associated RNAs were purified (see [Methods](#)). Following reverse transcription, the levels of YFP-Kan mRNA bound to the different proteins were analyzed with RT-PCR. Interestingly, we obtained distinct results for Gbp2 and Npl3. While mutant *gbp2* showed significantly decreased binding to YFP-Kan compared with WT, mutant *npl3* seemed to have a slightly higher affinity to the reporter RNA ([Fig 7D](#)). These results are in agreement with our current understanding of the functions of the two proteins. The increase in YFP-Kan binding of the mutant *npl3* might result in better assembly of the spliceosome and consequently increased splicing efficiency. The decreased binding of the mutant *gbp2* to the YFP-Kan transcript might delay the assembly of the nuclear RNA degradation machinery and hence provide more time and higher chances to the YFP-Kan transcript to be successfully spliced and exported to the cytoplasm.

### ***Cis* and *Trans* mutations show distinct phenotypes under permissive conditions**

If the *Trans* mutations in the SR-like proteins are beneficial, why are they not the WT sequence? We speculated that these mutations are beneficial only under the drug pressure, and that they otherwise come with a cost when the antibiotic is absent. We thus measured the effects of these mutations on fitness in a medium lacking the antibiotics (rich YPD medium). We found that cells with either of the *Cis* mutations grew modestly faster than the Splicing<sup>Low</sup> ancestor strain ([Fig 7E](#)), indicating that they alleviate some drug-independent burden of the inefficiently spliced intron in Splicing<sup>Low</sup> cells. In contrast, cells with either of the *Trans* mutations grew similarly to Splicing<sup>Low</sup> cells when the drug was not supplemented to the medium ([Fig 7E](#)). It seems then that the *Trans* mutations do not alleviate cellular fitness in the absence of the drug, i.e., they do not reduce the drug-independent burden of the inefficiently spliced intron.

### **Discussion**

In this work, we study the role of the splicing machinery in optimization of gene expression programs by placing selective pressure on cells to improve the splicing efficiency of a specific

gene. Our results provide molecular evidence for the relevance of splicing as another instrument in the cellular toolbox towards adjusting its gene expression patterns. To the best of our knowledge, we demonstrate the first experimental evidence of splicing efficiency adaptation, confirming that this adaptation can occur in *Cis* and *Trans* similarly to adaptations of other stages of the gene expression regulation process. Adaptation of splicing efficiency might be very common among species evolution, given the observed correlation between splicing efficiency and transcription level of genes [28]. Hence, it is of interest that we elucidate more molecular mechanism that allow optimization of intron splicing.

Two potential solutions to the burden that we imposed on our Splicing<sup>Low</sup> ancestor lines were surprisingly not realized during our lab-evolution. First, considering previous studies of splicing evolution, one could have expected the intron to be lost by a genomic deletion or through reverse transcription [50,51]. Such a solution could have been an ideal evolutionary adaptation to alleviate the burden, as we show that the intron-less strain has the highest fitness. The fact that we did not observe an intron-loss event suggests that this is a less accessible solution in this case, in agreement with previous evidence in yeast that nucleotide mutations are 33 times more frequent than deletion events [52]. It further seems that our evolving strains did not evolve a transcription- or translation-based solution that in principle could have elevated the expression of the YFP-Kan fusion gene. Instead, adaptation appears to have affected the splicing of the gene. Given that some of the adaptations were found within the intron, another surprise was that none of the mutations occurred within any of the intron's three functional sites: the 5' donor, 3' acceptor, or the branch point of splicing. Indeed, one mutation that was verified here to affect splicing resides in a region of the intron not known to exert a major effect on splicing, and another splicing-improving mutation happened in the upstream exon. These observations indicate that various positions in the intron and its proximity may facilitate the splicing rate and take part in the evolution of introns, a phenomenon that was previously discussed only in regards to alternative splicing [53,54].

Interestingly, the *Cis* mutation that resides inside the intron results in considerable changes in the predicted RNA secondary structure of the intron, presumably lowering the association probability of the 5' donor and branching point sites. This finding goes in line with previous observations that show how RNA structures can inhibit or facilitate binding of spliceosome components to the pre-mRNA and affect splicing efficiency [42,43], specifically at the edges of introns [36,42,43,55,56]. Thus, it is tempting to speculate that many more positions than previously recognized of introns' sequences, rather than only its edges and the branching point, are under selection and are biologically relevant to the intron function.

Notably, the fluorescence intensity per protein molecule of the YFP domain was decreased due to the nonsynonymous mutation in the YFP first exon, suggesting that under certain evolutionary constraints, selection may hamper superfluous functions of certain protein domains so as to increase availability of the entire protein.

Adaptive changes also occurred in *Trans* to the YFP-Kan locus and increased availability of the splicing machinery to this gene. An alternative explanation for evolutionarily adaptive changes in splicing availability could be based on cellular physiological response. In such a model, the introduction of an inefficiently spliced intron into a gene that is on high demand may occupy a larger portion of splicing machinery, which in turn prevents splicing of other intron-containing genes, which ultimately leads to degradation at the pre-mRNA level, e.g., by the mRNA quality control machinery [33]. Although valid, this physiological adaptation appears less likely to explain our results. First, the physiological model predicts that the expression level changes that we see following the evolutionary adaptations will also be seen when comparing the expression of the splicing genes and intron-containing genes between Splicing<sup>Low</sup> and the Control strain (that harbors no intron in its YFP-Kan construct). However, we

did not observe such a difference, suggesting that a physiological regulatory model is less likely (Fig 6A). Second, the two *Cis*-rescue strains, which did not evolve and only harbor our artificially introduced *Cis*-acting mutation, did not demonstrate changes in splicing availability. Generally, evolved strains showed higher supply-to-demand ratios for splicing compared with non-evolved strains in our study (Fig 6C). These observations more forcefully support an evolutionarily adaptive process that relies on additional genetic or epigenetic changes that are accumulated during the continuous growth of the evolved populations.

Recently, the competition of pre-mRNAs for the splicing machinery was shown to affect cellular function, as splicing efficiency of multiple introns was influenced by changes in the composition of the transcript pool [57]. While this mechanism was elegantly shown to take part in physiological adaptation by maintaining the separation between meiotic and vegetative gene-expression states, it is also possible that it can be used as an adaptive mechanism in evolution to optimize gene expression levels of cells.

Our findings demonstrate how availability of the splicing apparatus may have been adaptively increased both by elevating the expression level of the machinery's genes and/or by reducing expression of other intron-containing genes that probably compete with the antibiotic resistance un-spliced RNA for the spliceosome. Thus, increase in supply-to-demand ratio, analogous to the case in translation systems [8,58], appears to have evolved in this case.

Additionally, we revealed a role for SR-like proteins, *GBP2* and *NPL3*, when the splicing machinery adapts to a new need of optimizing splicing of a specific intron. Notably, these proteins have various functions, including early recruitment of the spliceosome, quality surveillance of nascent mRNA quality, association with the nuclear RNA degradation machinery of faulty transcripts, and finally, assistance with nuclear export for mature mRNAs.

Interestingly, the nonsynonymous mutations we found in these proteins occurred in one of their RNA recognition motifs. Using an RNA immunoprecipitation assay, we showed that the mutations in *Gbp2* and *Npl3* decrease and increase, respectively, the binding capacities of the YFP-Kan construct. Yet, these surprising opposite effects can be explained by the different roles of *Gbp2* and *Npl3*—the former is a quality control factor of splicing that elicits degradation of un- or mis-spliced transcripts, and the latter is a spliceosome-recruitment agent. On the one hand, the lower binding capacity of the mutated *Gbp2* probably provides more time for the spliceosome to complete the maturation of the YFP-Kan pre-mRNA before it is degraded. The improved binding capacity of the mutated *Npl3*, on the other hand, facilitates the recruitment of the spliceosome and hence might improve splicing efficiency of the transcript. The *Trans*-evolution mechanism we reveal here is intriguing because it shows that under acute selection, a cellular machinery can evolve for the need of one gene only.

Ultimately, we revealed a fundamental difference for *Cis* and *Trans* evolution of the splicing machinery when this cellular process faces a need to adapt. While *Cis*-based adaptations are “local” and lowered the burdens of splicing for the intron under selection, *Trans*-based adaptations showed wider cellular effects that may be costly to cells when the original evolutionary challenge is lifted. Further investigations will reveal which of these solutions, *Cis* or *Trans*, proves to be more evolutionarily stable—to fully reveal the dynamics of splicing adaptation when cells optimize their gene expression.

## Methods

### Yeast strains and plasmids

All *S. cerevisiae* strains in this study have the following genetic background: *his3Δ1::TEF2-mCherry::URA3::RPS28Ap-YIFP-KAN::NAT*; *canΔ1::STE2pr-Sp\_his5*; *lypΔ1::STE3pr-LEU2*; *leu2Δ0*; *ura3Δ0*.

Strains of Y-intron-FP were taken from Yofe and colleagues [36] and were introduced with a Kan resistance gene fused 3' terminally to the YFP. To reconstitute the mutations discovered after lab evolution (rescue strains), we amplified cassettes of Y-<sub>i<sub>mut</sub></sub>-FP-KAN and transformed these into the ancestor Control strain, selecting with KAN. Notably, all strains also carry an mCherry-fluorescent protein driven by an independent *TEF2* promoter that was used to normalize cell-to-cell variability for the YFP-Kan expression levels.

## Media

Cultures were grown at 30°C in rich medium (1% bacto-yeast extract, 2% bacto-peptone, and 2% dextrose [YPD]). Throughout all experiments, G418 was supplemented to the medium at a concentration of 3 mg/mL, which is 10-fold higher than the standard.

## Evolution experiments

Lab-evolution experiments were carried out by daily serial dilution for 80 days. Cells were grown on 1.2 mL of YPD+G418 at 30°C until reaching stationary phase and then diluted by a factor of 1:120 into fresh media (approximately 7 generations per dilution, a total of about 560 generations).

## Liquid growth measurements

Cells were grown in YPD+G418 at 30°C overnight. The following day, they were diluted to an OD = 0.05 in YPD+G418, and optical density (600 nm) measurements were taken at 30-minute intervals. Growth comparisons were performed using 96-well plates, and the growth curve for each strain was obtained by averaging at least 15 wells.

## Flow cytometry measurements of YFP-Kan levels

Cells were grown in YPD+G418 at 30°C overnight. The following day, they were diluted to an OD = 0.05 in YPD+G418, placed at 30°C, and followed until they reached logarithmic growth phase at an optical density of between 0.4 and 0.5. Then, YFP and mCherry levels were measured for between 20,000 and 50,000 cells for each culture with flow cytometry. Gating was performed according to side and forward scatters, and YFP levels were normalized with the mCherry signal for each cell individually.

## Quantitative PCR measurements of splicing efficiency

Cultures were grown in YPD+G418 at 30°C until cells reached the logarithmic growth phase at an optical density of approximately 0.4. Then, RNA was extracted using MasterPure kit (Epicentre) and were reverse transcribed to cDNA using random primers. A total of 2 µL of cDNA was added to each reaction as template for qPCR using light cycler 480 SYBR I master kit and the LightCycler 480 system (Roche Applied Science), according to the manufacturer's instructions. For each strain, qPCRs were performed with two to three biological repetitions and three technical repeats. A first qPCR was performed targeting the transcript-spliced version, with a forward primer complementing the exon-exon junction and a downstream reverse primer. A second PCR targeted the un-spliced version of the transcript, with a forward primer complementing the intron and the same reverse primer of the first reaction:  $F_{\text{exon-exon}} = 5' \text{-CACTACTTTAGGTTATGGTTT-3'}$ ;  $F_{\text{intron}} = 5' \text{-CTTCAATTTACTGAATTTGTATG-3'}$ ;  $R_{\text{both}} = 5' \text{-GTCTTGTTAGTTACCGTCA-3'}$ .

Splicing efficiency is reported as the average  $C_p$  of the spliced transcript minus the average  $C_p$  of the un-spliced version.

## mRNA deep sequencing

Cultures were grown in YPD+G418 at 30°C until cells reached the logarithmic growth phase at an optical density of approximately 0.4. Cells were then harvested by centrifugation and flash-frozen in liquid nitrogen. RNA was extracted using a modified protocol of nucleospin 96 RNA kit (Machery-Nagel). Specifically, cell lysis was done in a 96-deep-well plate by adding to each well 450  $\mu$ L of lysis buffer containing 1 M sorbitol, 100 mM EDTA, and 0.45  $\mu$ L lyticase (10 IU/ $\mu$ L). The plate was incubated at 30°C for 30 minutes to break cell walls and centrifuged for 10 minutes at 3,000 rpm, followed by the removal of the supernatant. Then, extraction continued as in the protocol of nucleospin 96 RNA kit, only using  $\beta$ -mercaptoethanol instead of DTT. Poly(A)-selected RNA extracts of approximate size of 200 bps were reverse transcribed to cDNA using poly(T) primers that were bar coded with a unique molecular identifier (UMI). cDNA was then amplified and sequenced with an Illumina HiSeq 2500.

## Analysis of mRNA deep sequencing

Processing of RNA-seq data was performed as described in Voichek and colleagues [59]. Shortly, reads were aligned using Bowtie [60] (parameters:—best—a—m 2 —strata -5 10) to the genome of *S. cerevisiae* (R64 from SGD) with an additional chromosome containing the sequence of the YFP-Kan construct. For each sequence, we normalized for PCR bias using UMIs, as described in Kivioja and colleagues [61]. Next, reads for each gene end (400 bp upstream to 200 bp downstream of the ORF's 3' end) were summed up to estimate the gene's expression level. Genes with coverage lower than 10 reads were excluded. To normalize for differences in coverage among samples, we divided each gene expression by the total read count of each sample and then multiplied by  $10^6$ . Then, the expression ratio was calculated between an evolved/rescue colony to the ancestor, and a  $\log_2$  operation was performed on that ratio. These values were used to compare expression levels of gene groups (ribosomal genes, general stress response genes, splicing machinery genes, intron-containing genes) and of the YFP-Kan mRNA levels as described in the manuscript. When calculating the expression levels of splicing machinery and intron-containing gene groups, the ribosomal and general stress response genes were excluded from the analysis in order to avoid bias from cellular regulation due to changes in physiology and growth rate of the cells.

## CRISPR genome engineering

CRISPR protocol for *S. cerevisiae* was performed as described previously [62]. Shortly, gRNA sequences targeting a locus near (<50 bps) the position of the desired mutations were cloned into the vector bRA89 that allows co-expression of Cas9 and gRNA. CRISPR plasmid (1  $\mu$ g) was co-transformed to yeast cells (20  $\mu$ L) along with a repair dsDNA cassette with the desired mutations (500 ng) using the standard yeast transformation protocol. Cells were plated on selective plates (YPD+hygromycin, 50 mg/mL), colonies were screened with PCR and Sanger sequencing for the desired mutations, positive colonies were grown on permissive media for 24 hours to ascertain loss of the CRISPR plasmid, and copies were stored at -80°C.

The repair cassette holds 45-bp homology on each end, flanking the desired mutation. Because the mutation did not modify the PAM sequence, additional synonymous mutations were also introduced alongside the desired mutation to stop the cutting of the locus by the Cas9+gRNA complex. These synonymous mutations were also introduced to the relevant genetic background without the desired mutations and were confirmed to have no effect on cellular growth or YFP expression levels.



## RNA co-immunoprecipitation

Yeast cells were grown in YPD at 25°C until logarithmic phase and collected via centrifugation. Cells were resuspended in 1× pellet volume RNA-IP buffer (25 mM Tris-HCl [pH 7.5], 150 mM NaCl, 2 mM MgCl<sub>2</sub>, 0.5% [v/v] Triton X-100, 0.2 mM PMSF in isopropanol, 0.5 mM DTT, 40 units of RiboLock RNase Inhibitor [Thermo Fisher Scientific], cOmplete EDTA-free Protease Inhibitor Cocktail [Roche Diagnostics and Sigma-Aldrich]) and lysed with vigorous vortexing three times, 20 s, 6 m/s with the FastPrep-24 Instrument (MP Biomedicals) in the presence of 1× pellet volume of glass beads. For immunoprecipitation, cleared lysates were incubated with or without Gbp2 or Npl3 antibody (rabbit polyclonal, self-made) at 4°C for 1 hour followed by the addition of 10 μL Protein G sepharose beads (Amersham Biosciences) and further incubation at 4°C for 2 hours. RNase-Free DNase I (QIAGEN) was also added to digest DNA during incubation. A total of 100 μL of the lysate was treated with DNase I and incubated at 4°C in parallel. The beads were washed five times with RNA-IP buffer. RNA was isolated from both lysate and eluate samples using the TRIzol Reagent (Thermo Fisher Scientific) and eluted in 20 μL DEPC-treated ddH<sub>2</sub>O. Prior to reverse transcription, the RNA samples were treated with TURBO DNA-free Kit (Thermo Fisher Scientific) to eliminate residual DNA.

## Reverse transcription and qRT-PCR of precipitated RNA

The same concentration of RNA from all samples was reverse transcribed using the FastGene Scriptase II cDNA Kit (NIPPON Genetics EUROPE) with Random Hexamer Primer (Thermo Fisher Scientific). PCR samples were prepared with qPCRBIO Sygreen Mix Lo-ROX (PCR Biosystems), and qRT-PCR was performed using the CFX Connect Real-Time PCR Detection System (Bio-Rad Laboratories) for 45 cycles with an annealing temperature of 60°C. Forward primer (5'-TTATTCACCTGGTGTGTC-3') and reverse primer (5'-CATGGAACTGGCAATTACC-3') that bind specifically to *YFP-Kan* were used. Each PCR reaction was performed in triplicates, and the average Cq value was used in further analysis. From each eluate Cq value, the corresponding lysate Cq value was subtracted ( $\Delta Cq$ ). The  $\Delta Cq$  of the no-antibody control sample was subtracted from the  $\Delta Cq$  of the pull-down sample ( $\Delta\Delta Cq$ ). Binding of the RNA was calculated by  $2^{-\Delta\Delta Cq}$ .

## Supporting information

**S1 Fig. Supply-to-demand ratio in each of the evolved colonies.** Folding change ratios in log<sub>2</sub> are shown for splicing genes (left) and intron-containing genes (right). Black line represents the median of the distribution. See numerical data for this figure in [S1 Data](#).  
(PDF)

**S1 Data. Supporting data for all main figures.**  
(XLSX)

**S1 Table. List of strains used in this work.**  
(XLSX)

**S2 Table. List of genomic mutations in *Trans*-evolved colonies.**  
(XLSX)

## Acknowledgments

We thank Martin Mikl for help with the qPCR experiments and Emmanuel Levy and Meta Heidenreich for help with CRISPR protocol in yeast. We thank Xiaoxue Snow Zhou, Daniel

Richard Corbi, and Angelika Amon for providing help and essential materials. We also thank Shlomit Gilad and Sima Benjamin from the Nancy & Stephen Grand Israel National Center for Personalized Medicine (G-INCPM) for assistance with genome sequencing. Our gratitude goes to Dana Bar-Zvi for help with the analysis of genome sequencing data. We also greatly appreciate comments about the text from Tslil Ast, Orna Dahan, and Avihu Yona and the entire Pilpel lab for discussions of the project. IF thanks the Azrieli Foundation for an Azrieli PhD Fellowship. YP is the incumbent of the Ben May Professorial Chair.

## Author Contributions

**Conceptualization:** Idan Frumkin, Ido Yofe, Yitzhak Pilpel.

**Data curation:** Idan Frumkin, Ido Yofe, Raz Bar-Ziv, Yonat Gurvich, Yen-Yun Lu, Ruth Towers, Dvir Schirman, Heike Krebber.

**Formal analysis:** Idan Frumkin, Yen-Yun Lu, Yoav Voichkek, Dvir Schirman, Heike Krebber.

**Funding acquisition:** Yitzhak Pilpel.

**Investigation:** Idan Frumkin, Ido Yofe, Yitzhak Pilpel.

**Methodology:** Idan Frumkin, Ido Yofe, Raz Bar-Ziv, Yonat Gurvich, Yen-Yun Lu, Yoav Voichkek, Heike Krebber, Yitzhak Pilpel.

**Project administration:** Idan Frumkin.

**Resources:** Yen-Yun Lu.

**Supervision:** Yitzhak Pilpel.

**Validation:** Idan Frumkin, Heike Krebber, Yitzhak Pilpel.

**Visualization:** Idan Frumkin.

**Writing – original draft:** Idan Frumkin, Yitzhak Pilpel.

**Writing – review & editing:** Idan Frumkin, Yitzhak Pilpel.

## References

1. Weingarten-Gabbay S, Segal E. The grammar of transcriptional regulation. *Hum Genet.* 2014; <https://doi.org/10.1007/s00439-013-1413-1> PMID: 24390306
2. Amorós-Moya D, Bedhomme S, Hermann M, Bravo IG. Evolution in regulatory regions rapidly compensates the cost of nonoptimal codon usage. *Mol Biol Evol.* 2010; 27: 2141–51. <https://doi.org/10.1093/molbev/msq103> PMID: 20403964
3. Quax TEF, Claassens NJ, Söll D, van der Oost J. Codon Bias as a Means to Fine-Tune Gene Expression. *Mol Cell.* 2015; 59: 149–161. <https://doi.org/10.1016/j.molcel.2015.05.035> PMID: 26186290
4. Gingold H, Pilpel Y. Determinants of translation efficiency and accuracy. *Mol Syst Biol.* 2011; 7: 481. <https://doi.org/10.1038/msb.2011.14> PMID: 21487400
5. Presnyak V, Alhusaini N, Chen Y-H, Martin S, Morris N, Kline N, et al. Codon optimality is a major determinant of mRNA stability. *Cell.* 2015; 160: 1111–24. <https://doi.org/10.1016/j.cell.2015.02.029> PMID: 25768907
6. Bazzini AA, Del Viso F, Moreno-Mateos MA, Johnstone TG, Vejnar CE, Qin Y, et al. Codon identity regulates mRNA stability and translation efficiency during the maternal-to-zygotic transition. *EMBO J.* 2016; 35: 2087–2103. <https://doi.org/10.15252/embj.201694699> PMID: 27436874
7. Babu M, Teichmann SA. Evolution of transcription factors and the gene regulatory network in *Escherichia coli*. *Nucleic Acids Res.* 2003; 31: 1234–44. <https://doi.org/10.1093/nar/gkg210> PMID: 12582243
8. Yona AH, Bloom-Ackermann Z, Frumkin I, Hanson-Smith V, Charpak-Amikam Y, Feng Q, et al. tRNA genes rapidly change in evolution to meet novel translational demands. *Elife.* 2013; 2: e01339. <https://doi.org/10.7554/eLife.01339> PMID: 24363105

9. Kirchner S, Ignatova Z. Emerging roles of tRNA in adaptive translation, signalling dynamics and disease. *Nat Rev Genet.* 2014; <https://doi.org/10.1038/nrg3861> PMID: 25534324
10. Wray GA. The evolutionary significance of cis-regulatory mutations. *Nat Rev Genet.* 2007; 8: 206–16. <https://doi.org/10.1038/nrg2063> PMID: 17304246
11. Lemos B, Araripe LO, Fontanillas P, Hartl DL. Dominance and the evolutionary accumulation of cis- and trans-effects on gene expression. *Proc Natl Acad Sci U S A.* 2008; 105: 14471–6. <https://doi.org/10.1073/pnas.0805160105> PMID: 18791071
12. Wittkopp PJ, Haerum BK, Clark AG. Evolutionary changes in cis and trans gene regulation. *Nature.* 2004; 430: 85–8. <https://doi.org/10.1038/nature02698> PMID: 15229602
13. Wong ES, Schmitt BM, Kazachenka A, Thybert D, Redmond A, Connor F, et al. Interplay of cis and trans mechanisms driving transcription factor binding and gene expression evolution. *Nat Commun.* 2017; 8: 1092. <https://doi.org/10.1038/s41467-017-01037-x> PMID: 29061983
14. Matera a. G, Wang Z. A day in the life of the spliceosome. *Nat Rev Mol Cell Biol.* 2014; 15: 108–121. <https://doi.org/10.1038/nrm3742> PMID: 24452469
15. Merkin J, Russell C, Chen P, Burge CB. Evolutionary dynamics of gene and isoform regulation in Mammalian tissues. *Science.* 2012; 338: 1593–9. <https://doi.org/10.1126/science.1228186> PMID: 23258891
16. Barbosa-Morais NL, Irimia M, Pan Q, Xiong HY, Gueroussov S, Lee LJ, et al. The evolutionary landscape of alternative splicing in vertebrate species. *Science.* 2012; 338: 1587–93. <https://doi.org/10.1126/science.1230612> PMID: 23258890
17. Petibon C, Parenteau J, Catala M, Elela SA. Introns regulate the production of ribosomal proteins by modulating splicing of duplicated ribosomal protein genes. *Nucleic Acids Res.* 2016; 44: gkw140. <https://doi.org/10.1093/nar/gkw140> PMID: 26945043
18. Parenteau J, Durand M, Morin G, Gagnon J, Lucier J-F, Wellinger RJ, et al. Introns within ribosomal protein genes regulate the production and function of yeast ribosomes. *Cell.* 2011; 147: 320–31. <https://doi.org/10.1016/j.cell.2011.08.044> PMID: 22000012
19. Reyes a., Anders S, Weatheritt RJ, Gibson TJ, Steinmetz LM, Huber W. Drift and conservation of differential exon usage across tissues in primate species. *Proc Natl Acad Sci.* 2013; 2–7. <https://doi.org/10.1073/pnas.1307202110> PMID: 24003148
20. Bush SJ, Chen L, Tovar-Corona JM, Urrutia AO. Alternative splicing and the evolution of phenotypic novelty. *Philos Trans R Soc Lond B Biol Sci.* 2017; 372. <https://doi.org/10.1098/rstb.2015.0474> PMID: 27994117
21. Irimia M, Roy SW. Origin of spliceosomal introns and alternative splicing. *Cold Spring Harb Perspect Biol.* 2014; 6. <https://doi.org/10.1101/cshperspect.a016071> PMID: 24890509
22. Roy SW, Irimia M. Mystery of intron gain: new data and new models. *Trends Genet.* 2009; 25: 67–73. <https://doi.org/10.1016/j.tig.2008.11.004> PMID: 19070397
23. Hooks KB, Delneri D, Griffiths-Jones S. Intron evolution in Saccharomycetaceae. *Genome Biol Evol.* 2014; <https://doi.org/10.1093/gbe/evu196> PMID: 25364803
24. Shabalina S a Ogurtsov AY, Spiridonov AN Novichkov PS, Spiridonov N a, Koonin E V. Distinct patterns of expression and evolution of intronless and intron-containing mammalian genes. *Mol Biol Evol.* 2010; 27: 1745–9. <https://doi.org/10.1093/molbev/msq086> PMID: 20360214
25. Mitrovich QM, Tuch BB, De La Vega FM, Guthrie C, Johnson AD. Evolution of yeast noncoding RNAs reveals an alternative mechanism for widespread intron loss. *Science.* 2010; 330: 838–41. <https://doi.org/10.1126/science.1194554> PMID: 21051641
26. Lee S, Stevens SW. Spliceosomal intronogenesis. *Proc Natl Acad Sci U S A.* 2016; 201605113. <https://doi.org/10.1073/pnas.1605113113> PMID: 27217561
27. Derr LK, Strathern JN, Garfinkel DJ. RNA-mediated recombination in *S. cerevisiae*. *Cell.* 1991; 67: 355–64. [https://doi.org/10.1016/0092-8674\(91\)90187-4](https://doi.org/10.1016/0092-8674(91)90187-4) PMID: 1655280
28. Ding F, Elowitz MB. Constitutive splicing and economies of scale in gene expression. *Nat Struct Mol Biol.* 2019; 26. <https://doi.org/10.1038/s41594-019-0226-x> PMID: 31133700
29. Kress TL, Krogan NJ, Guthrie C. A Single SR-like Protein, Npl3, Promotes Pre-mRNA Splicing in Budding Yeast. *Mol Cell.* 2008; 32: 727–734. <https://doi.org/10.1016/j.molcel.2008.11.013> PMID: 19061647
30. Chen Y-C, Milliman EJ, Goulet I, Côté J, Jackson CA, Vollbracht JA, et al. Protein arginine methylation facilitates cotranscriptional recruitment of pre-mRNA splicing factors. *Mol Cell Biol.* 2010; 30: 5245–56. <https://doi.org/10.1128/MCB.00359-10> PMID: 20823272
31. Moehle EA, Ryan CJ, Krogan NJ, Kress TL, Guthrie C. The yeast SR-like protein Npl3 links chromatin modification to mRNA processing. *PLoS Genet.* 2012; 8: e1003101. <https://doi.org/10.1371/journal.pgen.1003101> PMID: 23209445

32. Muddukrishna B, Jackson CA, Yu MC. Protein arginine methylation of Npl3 promotes splicing of the SUS1 intron harboring non-consensus 5' splice site and branch site. *Biochim Biophys Acta*. 2017; 1860: 730–739. <https://doi.org/10.1016/j.bbtagrm.2017.04.001> PMID: 28392442
33. Hackmann A, Wu H, Schneider U-M, Meyer K, Jung K, Krebber H. Quality control of spliced mRNAs requires the shuttling SR proteins Gbp2 and Hrb1. *Nat Commun*. 2014; 5: 3123. <https://doi.org/10.1038/ncomms4123> PMID: 24452287
34. Zander G, Hackmann A, Bender L, Becker D, Lingner T, Salinas G, et al. mRNA quality control is bypassed for immediate export of stress-responsive transcripts. *Nature*. 2016; 540: 593–596. <https://doi.org/10.1038/nature20572> PMID: 27951587
35. Ares M, Grate L, Pauling MH. A handful of intron-containing genes produces the lion's share of yeast mRNA. *RNA*. 1999; 5: 1138–9. <https://doi.org/10.1017/s135583829991379> PMID: 10496214
36. Yofe I, Zafrir Z, Blau R, Schuldiner M, Tuller T, Shapiro E, et al. Accurate, Model-Based Tuning of Synthetic Gene Expression Using Introns in *S. cerevisiae*. *PLoS Genet*. 2014; 10: e1004407. <https://doi.org/10.1371/journal.pgen.1004407> PMID: 24968317
37. de Nadal E, Ammerer G, Posas F. Controlling gene expression in response to stress. *Nat Rev Genet*. 2011; 12: 833–45. <https://doi.org/10.1038/nrg3055> PMID: 22048664
38. Metzl-Raz E, Kafri M, Yaakov G, Soifer I, Gurvich Y, Barkai N. Principles of cellular resource allocation revealed by condition-dependent proteome profiling. *Elife*. 2017; 6. <https://doi.org/10.7554/eLife.28034> PMID: 28857745
39. Gasch AP, Spellman PT, Kao CM, Carmel-Harel O, Eisen MB, Storz G, et al. Genomic expression programs in the response of yeast cells to environmental changes. *Mol Biol Cell*. 2000; 11: 4241–57. <https://doi.org/10.1091/mbc.11.12.4241> PMID: 11102521
40. Sarkisyan KS, Bolotin DA, Meer M V., Usmanova DR, Mishin AS, Sharonov G V., et al. Local fitness landscape of the green fluorescent protein. *Nature*. 2016; 533: 397–401. <https://doi.org/10.1038/nature17995> PMID: 27193686
41. Webb M, editor. *Cancer Susceptibility* [Internet]. Totowa, NJ: Humana Press; 2010. <https://doi.org/10.1007/978-1-60761-759-4>
42. Warf MB, Berglund JA. Role of RNA structure in regulating pre-mRNA splicing. *Trends Biochem Sci*. 2010; 35: 169–78. <https://doi.org/10.1016/j.tibs.2009.10.004> PMID: 19959365
43. Gahura O, Hammann C, Valentová A, Púta F, Folk P. Secondary structure is required for 3' splice site recognition in yeast. *Nucleic Acids Res*. 2011; 39: 9759–67. <https://doi.org/10.1093/nar/gkr662> PMID: 21893588
44. Lorenz R, Bernhart SH, Höner Zu Siederdisen C, Tafer H, Flamm C, Stadler PF, et al. ViennaRNA Package 2.0. *Algorithms Mol Biol*. 2011; 6: 26. <https://doi.org/10.1186/1748-7188-6-26> PMID: 22115189
45. Qian W, Yang JR, Pearson NM, Maclean C, Zhang J. Balanced codon usage optimizes eukaryotic translational efficiency. *PLoS Genet*. 2012; 8. <https://doi.org/10.1371/journal.pgen.1002603> PMID: 22479199
46. Brackley CA, Romano MC, Thiel M. The dynamics of supply and demand in mRNA translation. *PLoS Comput Biol*. 2011; 7: e1002203. <https://doi.org/10.1371/journal.pcbi.1002203> PMID: 22022250
47. Frumkin I, Lajoie MJ, Gregg CJ, Hornung G, Church GM, Pilpel Y. Codon usage of highly expressed genes affects proteome-wide translation efficiency. *Proc Natl Acad Sci U S A*. 2018; 115: E4940–E4949. <https://doi.org/10.1073/pnas.1719375115> PMID: 29735666
48. Lei EP, Krebber H, Silver PA. Messenger RNAs are recruited for nuclear export during transcription. *Genes Dev*. 2001; 15: 1771–82. <https://doi.org/10.1101/gad.892401> PMID: 11459827
49. Lee MS, Henry M, Silver PA. A protein that shuttles between the nucleus and the cytoplasm is an important mediator of RNA export. *Genes Dev*. 1996; 10: 1233–46. <https://doi.org/10.1101/gad.10.10.1233> PMID: 8675010
50. Sverdlov A V, Babenko VN, Rogozin IB, Koonin E V. Preferential loss and gain of introns in 3' portions of genes suggests a reverse-transcription mechanism of intron insertion. *Gene*. 2004; 338: 85–91. <https://doi.org/10.1016/j.gene.2004.05.027> PMID: 15302409
51. Cohen NE, Shen R, Carmel L. The role of reverse transcriptase in intron gain and loss mechanisms. *Mol Biol Evol*. 2012; 29: 179–86. <https://doi.org/10.1093/molbev/msr192> PMID: 21804076
52. Zhu YO, Siegal ML, Hall DW, Petrov D a. Precise estimates of mutation rate and spectrum in yeast. *Proc Natl Acad Sci U S A*. 2014; 111: E2310–8. <https://doi.org/10.1073/pnas.1323011111> PMID: 24847077
53. Rosenberg AB, Patwardhan RP, Shendure J, Seelig G. Learning the sequence determinants of alternative splicing from millions of random sequences. *Cell*. 2015; 163: 698–711. <https://doi.org/10.1016/j.cell.2015.09.054> PMID: 26496609

54. Julien P, Miñana B, Baeza-Centurion P, Valcárcel J, Lehner B. The complete local genotype-phenotype landscape for the alternative splicing of a human exon. *Nat Commun.* 2016; 7: 11558. <https://doi.org/10.1038/ncomms11558> PMID: 27161764
55. Soemedi R, Cygan KJ, Rhine CL, Glidden DT, Taggart AJ, Lin C-L, et al. The effects of structure on pre-mRNA processing and stability. *Methods.* 2017; 125: 36–44. <https://doi.org/10.1016/j.ymeth.2017.06.001> PMID: 28595983
56. Hiller M, Zhang Z, Backofen R, Stamm S. Pre-mRNA secondary structures influence exon recognition. *PLoS Genet.* 2007; 3: e204. <https://doi.org/10.1371/journal.pgen.0030204> PMID: 18020710
57. Munding EM, Shiue L, Katzman S, Donohue JP, Ares M. Competition between Pre-mRNAs for the Splicing Machinery Drives Global Regulation of Splicing. *Mol Cell.* 2013; 51: 338–348. <https://doi.org/10.1016/j.molcel.2013.06.012> PMID: 23891561
58. Kudla G, Murray AW, Tollervey D, Plotkin JB. Coding-sequence determinants of gene expression in *Escherichia coli*. *Science.* 2009; 324: 255–8. <https://doi.org/10.1126/science.1170160> PMID: 19359587
59. Voichek Y, Bar-Ziv R, Barkai N. Expression homeostasis during DNA replication. *Science.* 2016; 351: 1087–90. <https://doi.org/10.1126/science.aad1162> PMID: 26941319
60. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 2009; 10: R25. <https://doi.org/10.1186/gb-2009-10-3-r25> PMID: 19261174
61. Kivioja T, Vähärautio A, Karlsson K, Bonke M, Enge M, Linnarsson S, et al. Counting absolute numbers of molecules using unique molecular identifiers. *Nat Methods.* 2011; 9: 72–4. <https://doi.org/10.1038/nmeth.1778> PMID: 22101854
62. Dekker B, Anand R, Memisoglu G, Haber J. Cas9-mediated gene editing in *Saccharomyces cerevisiae*. *Protoc Exch.* 2017; <https://doi.org/10.1038/protex.2017.021a>