

ADVANCED REVIEW

Environmental influences on RNA processing: Biochemical, molecular and genetic regulators of cellular response

Athma A. Pai¹  | Francesca Luca² ¹RNA Therapeutics Institute, University of Massachusetts Medical School, Worcester, Massachusetts²Center for Molecular Medicine and Genetics, and Department of Obstetrics and Gynecology, Wayne State University, Detroit, Michigan**Correspondences**

Athma Pai, RNA Therapeutics Institute UMass Medical School Worcester, MA.

Email: athma.pai@umassmed.edu

Francesca Luca, Center for Molecular Medicine and Genetics, and Department of Obstetrics and Gynecology, Wayne State University, Detroit, MI. Email: fluca@wayne.edu

Funding information

University of Massachusetts Medical School; National Institute of General Medical Sciences, Grant/Award Number: NIH R01GM109215; American Heart Association, Grant/Award Number: (AHA 14SDG20450118)

RNA processing has emerged as a key mechanistic step in the regulation of the cellular response to environmental perturbation. Recent work has uncovered extensive remodeling of transcriptome composition upon environmental perturbation and linked the impacts of this molecular plasticity to health and disease outcomes. These isoform changes and their underlying mechanisms are varied—involving alternative sites of transcription initiation, alternative splicing, and alternative cleavage at the 3' end of the mRNA. The mechanisms and consequences of differential RNA processing have been characterized across a range of common environmental insults, including chemical stimuli, immune stimuli, heat stress, and cancer pathogenesis. In each case, there are perturbation-specific contributions of local (cis) regulatory elements or global (trans) factors and downstream consequences. Overall, it is clear that choices in isoform usage involve a balance between the usage of specific genetic elements (i.e., splice sites, polyadenylation sites) and the timing at which certain decisions are made (i.e., transcription elongation rate). Fine-tuned cellular responses to environmental perturbation are often dependent on the genetic makeup of the cell. Genetic analyses of interindividual variation in splicing have identified genetic effects on splicing that contribute to variation in complex traits. Finally, the increase in the number of tissue types and environmental conditions analyzed for RNA processing is paralleled by the need to develop appropriate analytical tools. The combination of large datasets, novel methods and conditions explored promises to provide a much greater understanding of the role of RNA processing response in human phenotypic variation.

This article is categorized under:

RNA Processing > RNA Editing and Modification

RNA Evolution and Genomics > Computational Analyses of RNA

RNA Processing > Splicing Mechanisms

RNA Processing > Splicing Regulation/Alternative Splicing

KEYWORDS

environmental perturbation, RNA processing response, splicing QTL

1 | INTRODUCTION

The ability to interact with and respond to a range of environmental stimuli is an intrinsic property of any organism. On a cellular level, this response must be rapid, tunable, and dynamic to allow a cell to adapt quickly to a new environmental condition. While there is already a solid body of literature focusing on transcriptional gene regulatory responses, it is becoming increasingly clear that co- or post-transcriptional processing of mRNA molecules may also play large roles in regulating the

composition and concentration of both the transcriptome and proteome upon a cellular response. Core RNA processing mechanisms and machinery are highly conserved (Gerstberger, Hafner, & Tuschl, 2014; Lee & Rio, 2015), with misregulation of constitutive patterns often having severe cellular consequences. For instance, mutations underlying mRNA splicing defects are implicated as causal effectors of many debilitating human diseases, such as muscular dystrophy, spinal muscular atrophy, cystic fibrosis, retinal degenerative disorders and many others (as reviewed in Scotti and Swanson (2016)). Analysis of differential exon usage across cellular contexts (175 individuals and 43 body sites) represented in the Genotype-Tissue Expression (GTEx) Consortium pilot dataset (Mele et al., 2015), identified gene expression as the main mechanism for cell specificity, with splicing variability likely playing a secondary or complementary role. Brain tissues are a notable exception, since they have very distinct splicing programs characterized by preferential exon inclusion. The high level of splicing variability observed in the brain (specifically in the cerebellum) is supported by mechanisms such as the brain-specific expression of particular RNA-binding proteins. RNA processing decisions can be defining features of particular tissues and cell types and have the potential to determine cell fate (Hu et al., 2015; Turner & Díaz-Muñoz, 2018; Xu, Wang, Luo, Zhao, & Zhou, 2017). However, large variability in RNA processing across closely related species despite conservation of gene expression levels (Barbosa-Morais et al., 2012; Merkin, Russell, Chen, & Burge, 2012) suggests that RNA processing mechanisms are good substrates for regulatory plasticity to create fine-tuned transcriptome patterns. By extension, these pathways are likely prime candidates for coordinating the cascade of molecular events that occur when cells are faced with a change in their environment.

Broadly, mechanisms of RNA processing influence the final composition or function of an RNA molecule. This encompasses mechanisms such as 5' capping, 3' cleavage, 3' polyadenylation, RNA modification, RNA secondary structure, and mRNA splicing, all of which are likely to play a role in how cells respond to changing environments. In this review, we focus only on mechanisms influencing mRNA sequence or isoform composition, particularly alternative transcription initiation, alternative splicing, and alternative 3' cleavage. These processes have been extensively studied, with well-characterized regulatory machinery, auxiliary factors, and sequence elements canonically associated with each mechanistic step (Figure 1). Alternative transcription initiation is governed by a decision to initiate transcription at an alternative start site. This process is likely determined by the local chromatin and transcription factor (TF) environment (Lai & Pugh, 2017), which is mediated by DNA sequence elements governing TF binding. RNA Polymerase II (PolII) is recruited to initiate transcription, undergoing multiple

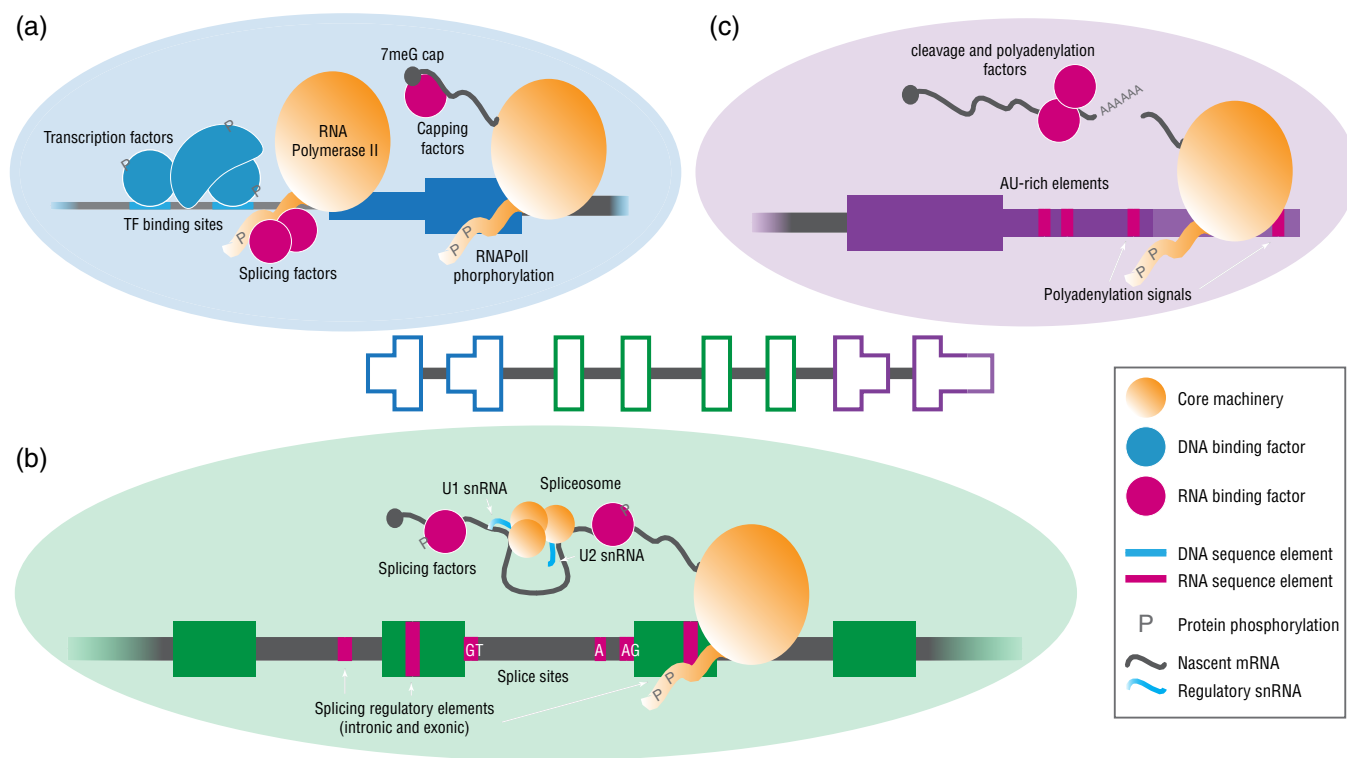


FIGURE 1 Mechanistic details of RNA processing. Isoform composition and complexity is regulated by core cellular machinery, DNA/RNA binding factors, and DNA/RNA regulatory sequences. A second layer of regulation is added by the dynamic mutational and modification events occurring at both RNA and protein levels. (a) Alternative transcription initiation is governed by transcriptional mechanisms regulating the usage of DNA promoters and transcription start sites. (b) Alternative splicing is governed by the spliceosome machinery and auxiliary splicing factors, which guide the recruitment of core snRNAs such as U1 and U2 to the nascent RNA molecule and direct 5' or 3' splice site selection. (c) Alternative cleavage and polyadenylation sites and factors direct decisions involved in 3' end processing of nascent molecules

phosphorylation checkpoints prior to entering productive transcript elongation (Harlen & Churchman, 2017). Processing at the 5' end is completed with the addition of a 7-methylguanosine cap on the RNA molecule shortly after exit from the PolII exit channel (Bentley, 2014). During the pause between initiation and elongation, splicing regulatory factors (SRFs) are loaded onto the C-terminal domain of PolII and their presence is often necessary for productive elongation (Hsin & Manley, 2012). It is unclear whether the SRFs loaded onto PolII directly regulate downstream splice site choice, though the same factors are associated with the regulation of alternative splicing (Hsin & Manley, 2012). RNA splicing occurs through the recognition of three sites on an mRNA molecule (Lee & Rio, 2015): a 5' splice site (nine nucleotide motif spanning the exon–intron boundary, represented by a mostly invariant GU defining the 5' end of the intron), a branchpoint (an A nucleotide within a degenerate motif located approximately 20–50 nucleotides upstream of the 3' end of an intron), and a 3' splice site (approximately 15–20 nucleotide motif spanning the intron–exon boundary, starting with a polypyrimidine track and containing a mostly invariant AG defining the 3' end of the intron). These sequences are initially bound by U1 and U2 small nuclear ribonucleoproteins (snRNPs; bound to the 5' splice site and branchpoint, respectively), which then recruit the larger spliceosome complex that catalyzes the two sequential steps of splicing and excises an intron (Lee & Rio, 2015). Splice site choice can be regulated by the binding of SRFs to splicing regulatory elements (SREs) located in the intron or flanking exons, resulting in either enhanced usage or repression of given splice sites (Fu & Ares, 2014; Wang & Burge, 2008). Finally, cleavage and polyadenylation specificity factor and cleavage stimulation factor bind to a polyadenylation signal (PAS), usually found in the 3' untranslated regions (UTRs) of mRNAs, to initiate cleavage of the mRNA approximately 30–40 nucleotides downstream of the PAS (Elkon, Ugalde, & Agami, 2013; Tian & Manley, 2013). An mRNA molecule is fully mature after the addition of the polyA tail to the 3' end of a molecule.

Each of these stages in RNA processing are comprised of complex mechanisms involving multiple cis-acting nucleotide sequence elements and trans-acting regulatory factors, any of which can be targets in the process of regulating a cellular phenotype. Furthermore, while the processes of transcription initiation, mRNA splicing and RNA cleavage are commonly studied independently, their temporal relationships (Bentley, 2014; Brugiolo, Herzel, & Neugebauer, 2013), spatial proximity (Bentley, 2014; Han, Xiong, Wang, & Fu, 2011) and patterns of coordinated decisions (Anvar et al., 2018) suggest that they are coregulated to influence final isoform levels. Here, we review the current literature on how RNA processing is affected upon changes in cellular environment (Figure 2), as well as the roles that various cis- or trans-acting regulators of RNA processing play in regulating the cellular response to environmental perturbations.

2 | SHIFTS IN RNA PROCESSING ARE TREATMENT-SPECIFIC

Environmental perturbations have large effects on both organismal and cellular traits, including gene expression and RNA processing. Analyses of RNA processing response have largely focused on genome-wide studies of response to individual chemical and biological agents or large-scale screens of drug compound libraries with reporter minigenes.

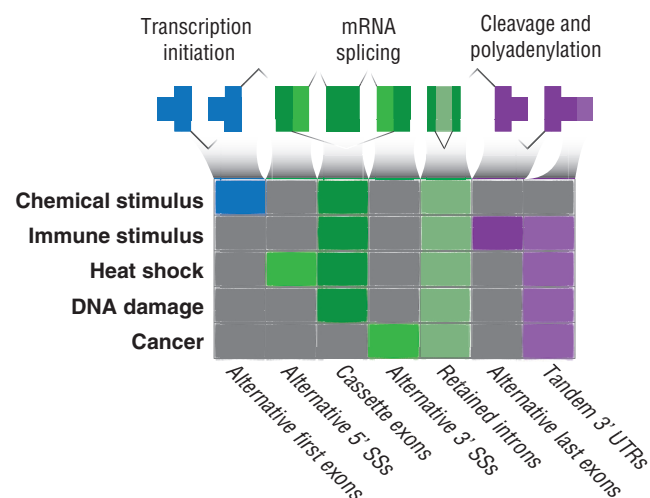


FIGURE 2 The diversity of RNA processing changes across environmental perturbations. While each environmental context that has been studied so far shows substantial evidence for complex transcriptome changes across all RNA processing categories (gray boxes), each context also has characteristic global changes (dark colored boxes). These involve particular event types that are frequently observed to change, change with a greater magnitude, or change in a directional pattern after the environmental perturbation. Notably, retained introns are consistently enriched across all environmental perturbations, indicating that changes in environmental context affect splicing efficiency rather than only isoform choice. Global retention of annotated retained introns may serve as a sign that both regulated and canonically constitutive introns are being retained due to overall perturbation of the splicing process

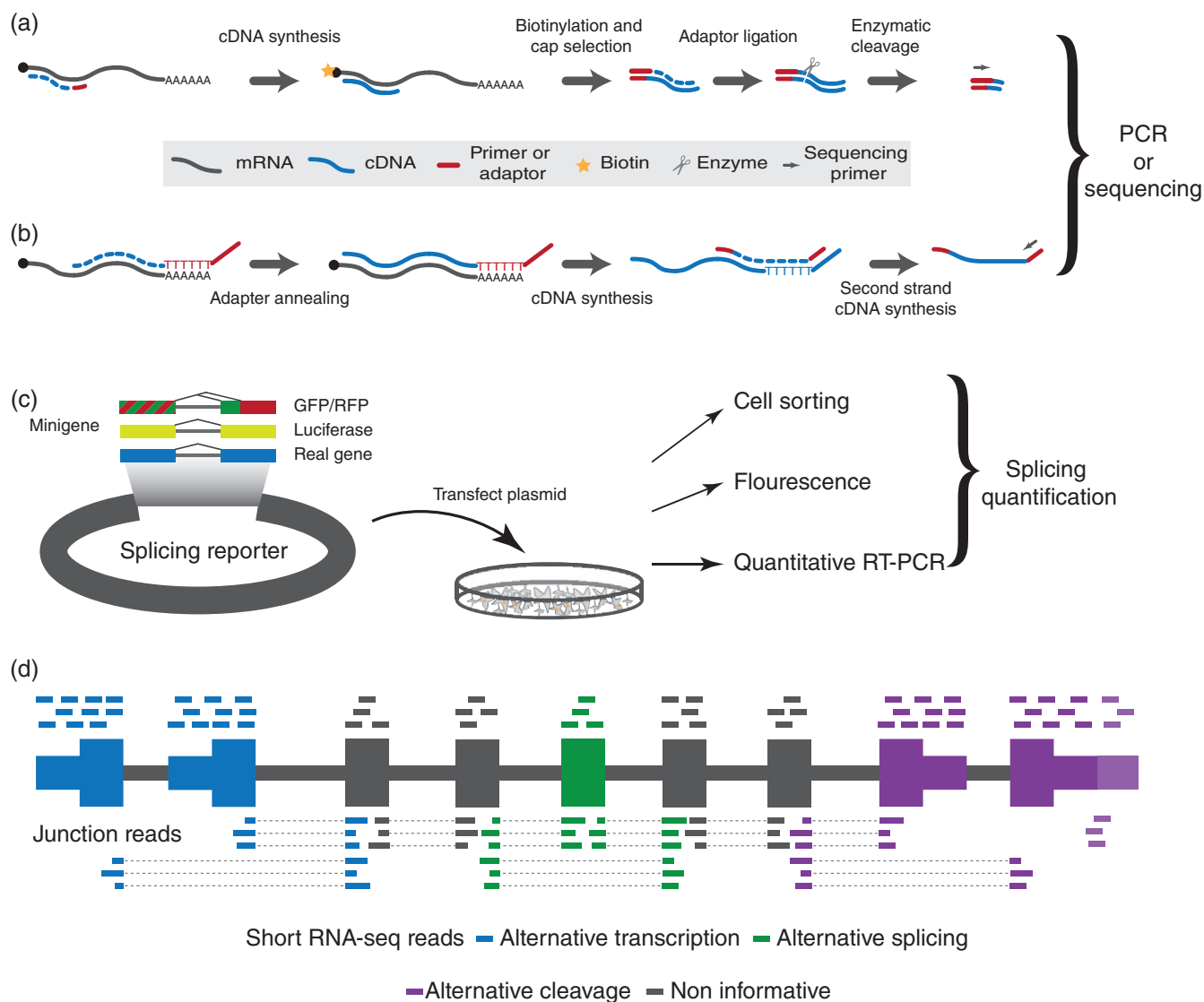


FIGURE 3 Methods to assay isoform composition. (a) Cap Analysis of Gene Expression protocols are commonly used to interrogate the location and usage of transcription initiation sites. The technique relies on capture of molecules with a 5' 7-methylguanosine cap, commonly through selective biotinylation of the cap. Following selection, a 5' adaptor is added, containing an enzyme recognition site that allows cleavage of the cDNA at a given distance away from the adaptor. The resulting product can be assayed through PCR and sequencing with either conventional or high-throughput methods. (b) Rapid Analysis of cDNA Ends protocols allow the interrogation of the location and usage of mRNA cleavage and polyadenylation sites. Synthesis of cDNA is initiated by annealing of an oligo-dT primer that contains a known 3' adaptor sequence. Second strand synthesis is initiated using either a gene-specific or random primer, and the resulting product can be assayed through PCR and sequencing with either conventional or high-throughput methods. (c) Splicing reporter assays are used to validate or interrogate sequences that affect splice site usage. A minigene construct containing either a real intron flanked by its corresponding exons or a synthetic construct encoding either luciferase, green fluorescent protein (GFP) or red fluorescent protein (RFP) is cloned into a reporter plasmid. The splicing of the introns in any of these cases can be assayed using qRT-PCR. For luciferase constructs, the increased splicing of intervening intron would result in increased luciferase production, which can be assayed using fluorescence-based assays. Finally, the GFP/RFP construct often interrogates usage of individual splice sites, such that usage of an upstream splice site would encode GFP, while usage of the downstream splice site would encode RFP. The relative proportion of splice site usage can then be assayed by sorting cells based on markers for GFP and RFP. (d) Short reads from RNA-sequencing techniques can be used to identify and quantify RNA processing patterns across multiple event types. Quantification of whole isoforms or exons often rely on the quantification reads within all regions present in an isoform, while methods specifically interrogating alternative patterns rely on reads that are informative for usage of particular exons. The most informative reads are junction reads that are split between two exons and span the resulting junction. We note that reads for terminal exons could indicate either differential TSS or polyA site usage (for AFEs and ALEs, respectively) or differential splicing of terminal exons to internal exons

Most of the individual agents investigated are relevant for cancer therapy due to their ability to induce DNA damage (UV radiation treatment has also been investigated in this context; Tresini, Martijn, & Vermeulen, 2016). A genome-wide study of the splicing changes in human colon and breast cancer cell lines (HCT116 and MCF7) in response to the plant alkaloid camptothecin (CPT) used an exon microarray to study splicing (Box 1) and identified alternative exon usage predominantly within splicing factors (Solier et al., 2010). The DNA topoisomerase Top1 is the primary target of CPT, and its clinical

BOX 1

THE CHALLENGES OF MEASURING ALTERNATIVE RNA PROCESSING: EXPERIMENTAL TECHNIQUES

There are many experimental approaches that have been developed to assess RNA processing patterns. The usage and location of transcription start sites (TSSs) is typically measured using a cap-pulldown method, such as the Cap Analysis of Gene Expression, either for a single gene or in a high-throughput fashion (Kodzius et al., 2006; Figure 3a), or newer and more precise methods such as Cap-seq that isolate RNA based on pulldown of cap-binding complexes (Xie et al., 2013). Similarly, the position of the 3' end of a transcript can be assayed using Rapid Analysis of cDNA Ends methods (Frohman, Dush, & Martin, 1988; Figure 3b), or newer methods like 3p-seq that provide more resolution for 3' end identification (Jan, Friedman, Ruby, & Bartel, 2011). The majority of technology development has focused on measuring alternative splice site and exon usage. Classically, these experiments were performed with splicing reporter assays involving plasmids containing minigenes designed to resemble introns and exons from a real gene with known alternative splicing (i.e., Somarelli et al., 2013; Stoilov, Lin, Damoiseaux, Nikolic, & Black, 2008), or measure splicing in a synthetic gene construct (i.e., inserting introns into Luciferase (Younis et al., 2010) or a dual-signal reporter such as that encoding both green fluorescent protein (GFP) and red fluorescent protein (RFP) (Nasim & Eperon, 2006; Orengo, Bundman, & Cooper, 2006; Figure 3c). Minigenes contain all the sequence elements necessary to create a mature transcript (and often peptide fragment) and are quite useful to: (a) perform mutagenesis experiments to dissect the cis-elements responsible for splicing of the given intron or (b) screen the splicing response of the specific intron to treatment with many stimuli or conditions. The read-out for all perturbations (either mutagenesis or environmental conditions) can be quickly measured in a high-throughput fashion using either qRT-PCR, fluorescence assays, or cell sorting. However, these techniques are limited to assaying the response of a small number of genes and exons in a synthetic construct. More recently, microarrays—especially exon arrays—allowed for the measurement of hundreds of exons at a time. However, the limited sensitivity of exon arrays hindered the measurement of small changes in splicing and these arrays were still reliant on knowing all isoform possibilities.

derivatives, topotecan and irinotecan are widely used as anticancer agents. Validation experiments of the genes with the largest splicing changes showed that the alternative exon usage induced by CPT is not observed with cisplatin or vinblastine, which are other two widely used chemotherapeutic agents. Genome-wide alternative splicing has also been observed in murine intestinal organoids in response to ER-stress (thapsigargin treatment) and nutrient starvation (Tsalikis et al., 2016). RNA-seq analysis showed that splicing changes shared between both treatments also occurred predominantly in splicing factors. A study of intron retention showed that certain introns that are stably detained in nuclear transcripts (without triggering nonsense-mediated decay even an hour after transcription) are more sensitive to drug-inhibition of Clk, a stress-responsive kinase (Boutz, Bhutkar, & Sharp, 2015). Similarly, these retained introns were often sensitive to changes after DNA damage, with consequences for effects on gene expression of the gene containing the retained intron. Overall, the results of these studies suggest an auto-regulatory feedback loop regulating the splicing response that is common to different types of stress.

A large body of literature reports single gene, microarray and RNA-seq analyses of the transcriptional response to nuclear receptor ligands, including steroid hormones. The transcriptional response to nuclear receptor ligands such as progesterone, estrogen and vitamin D is predominantly the result of the nuclear receptor acting as a TF to directly regulate gene expression changes. However, several studies have demonstrated that NR ligands can control both transcription and mRNA splicing by recruiting receptor coregulators (reviewed in Zhou et al. (2015)).

Minigene reporter assays have typically been used to screen large panels of treatments for their effects on mRNA splicing (Box 1). One study screened a library of 1,100 compounds (Prestwick, France) of Food and Drug Administration-approved and other drugs and a smaller 340-compound library (BioMol, Germany) consisting of enzyme inhibitors and ion-channel antagonists. This study successfully identified many compounds that cause significant differential exon usage (Stoilov et al., 2008) of the minigene construct in HEK293 cells. Several of these compounds are cardiotonic steroids, including digoxin, which is used to treat heart failure. A much larger library of compounds (>23,000) was screened with a rapid-response splicing reporter assay in HeLa cells (Younis et al., 2010), identifying compounds with splicing inhibitor function and validating these results with follow-up experiments.

Larger scale studies of RNA processing across different environmental contexts have been made possible by the widespread availability of high-throughput sequencing technology. *In vitro* studies of cellular response to environmental perturbations have the distinct advantage of allowing for tight control of the environmental conditions studied. Furthermore, the ability to perform different functional genomic assays in the same controlled context allows for a better understanding of the cellular mechanisms leading to variation in RNA processing. Recently, a high-throughput and cost-effective approach was developed to investigate transcriptional response to environmental perturbations in 250 cellular environments (Moyerbrailean et al.,

2015, 2016). The high coverage RNA-seq data collected in 89 environmental conditions (including 5 cell types and 35 treatments that were chosen as the cell types and treatments with the largest transcriptional changes) provided the largest dataset to date to analyze RNA processing response (Richards et al., 2017). Using the probabilistic framework implemented in the software mixture of isoforms (Katz, Wang, Airolidi, & Burge, 2010), Richards et al. identified 15,300 changes in RNA processing across conditions, representing a unique set of 8,489 events that significantly differ between at least one treatment and control conditions. This study cemented the idea that cellular response to environmental perturbations involves several different RNA processing mechanisms. Eight event types were considered: skipped exons (SEs), retained introns (RIs), alternative 3' or 5' splice sites (A3SSs, A5SSs), mutually exclusive exons, alternative first or last exons (AFEs, ALEs), and tandem UTRs (TandemUTRs). The largest number of changes across all perturbations considered (after correcting for the number of events tested) occurred in AFEs, ALEs, and RIs, while A3SSs had the smallest number of changes. Overall, AFE usage changed the most across environments, though there was substantial variation in the extent to which each event type changed across treatments. For example, vitamin E, tunicamycin, and cadmium were enriched for different event types: vitamin E is enriched for A5SS while cadmium is depleted for RI.

The detailed analysis of a large number of environmental contexts in this work facilitated the evaluation of shared patterns of shifts across treatments and investigation of common regulatory mechanisms. AFEs, RIs, and SEs consistently showed a directionality in their usage patterns, such that within an environmental context (cell type-treatment combination), there was often a preference for or against usage of the upstream TSS leading to different AFEs, retention of an intron, or inclusion of a cassette exon. Eighteen environments were enriched for increased inclusion of RIs, while only two contexts led to increased splicing of RIs. For AFEs, 34 environmental contexts led to increased usage of an upstream TSS, while only seven contexts showed shifts towards usage of a downstream TSS. Richards et al. also analyzed the cis sequence elements and trans regulatory elements that may regulate these concerted RNA processing shifts. Changes in RIs correlated with differential gene expression of *LARP7* and 13 other splicing factors in response to environmental perturbations, suggesting that increased expression of *LARP7* could lead to more intron retention. Additionally, the presence of putative splicing factor binding sites near the RNA processing events is a predictive feature of SE and RI usage and provides a plausible cis-regulatory mechanism for these RNA processing events. Analogously, this study found that differential expression and binding of TFs may regulate AFE usage in response to environmental perturbations. A total of 328 (out of 1,342) TF gene expression changes are correlated with shifts in AFE usage due to differential TSS preferences, and TF binding within 1 Kb of the TSS is significantly predictive of TSS usage. Functional support for the importance of TF binding in determining AFE usage was provided by chromatin accessibility data collected in lymphoblastoid cell lines (LCLs) treated with selenium. Data from an Assay for Transposase-Accessible Chromatin using sequencing (Buenrostro, Giresi, Zaba, Chang, & Greenleaf, 2013), can be used to identify TF binding sites by inferring the footprint that they leave on chromatin (Pique-Regi et al., 2011). Significant differences in chromatin accessibility after selenium treatment were identified for 64 TF motifs located near alternative TSSs. Of these 64 motifs, 26 correspond to ETS TF family members, including ELF2. These results provide support for a mechanism whereby environmental perturbations induce chromatin accessibility and differential TF binding at sites that regulate AFE usage.

3 | WIDESPREAD RNA PROCESSING CHANGES UPON IMMUNE STIMULATION MAY BE KINETICALLY REGULATED

The immune response is perhaps the best known cellular program that an organism uses to mount responses to changes in environmental conditions. Infectious pathogenic agents or immune stimuli initiate a cascade of cell surface receptor interactions that trigger signaling responses to regulate a multitude of gene expression programs (Medzhitov and Janeway 1998). Numerous studies have characterized the extensive transcriptional changes upon immune activation (Huang et al., 2001; Medzhitov & Horng, 2009; Smale, 2010), but less is known about the full extent to which changes in RNA processing impact cellular responses and how these transcriptome changes are regulated.

Initial studies focused on a few prominent immune genes and characterized how altered isoform usage after an immune response directly regulates cellular function. For instance, the *CD45* antigen is alternatively spliced upon T-cell activation to exclude three alternative exons (Jacobsen et al., 2000, 2002). Increased skipping of these exons has been directly linked to reduction of CD45 phosphatase activity and attenuation of antigen receptor signaling in T-cells (Lynch & Weiss, 2000), with a broader role in preventing the development of auto-immune disease mediated by T-cell hyperactivation (Hermiston, Xu, Majeti, & Weiss, 2002; Lynch & Weiss, 2000). Similarly, the MAPK kinase *MKK7* is also alternatively spliced upon T-cell activation to exclude its second exon. The resulting *MKK7* isoform enhances JNK pathway activity by increasing c-Jun phosphorylation and ultimately leads to upregulation of *TNF-alpha* (Martinez et al., 2015). Notably, this splicing event is part of a positive feedback loop promoting JNK-dependent signaling (Martinez et al., 2015), since the splicing event itself is dependent

on JNK-induced expression of the SRF *CELF2* (Mallory et al., 2015; Martinez et al., 2015). Approximately 25% of splicing changes after T-cell activation are likely regulated by this JNK-*CELF2* autoregulatory axis during crucial stages of the immune response (Mallory et al., 2015; Martinez et al., 2015). These studies highlight how the regulation of RNA processing can attenuate the function of several individual regulators of the immune response (Martinez & Lynch, 2013).

On a genome-wide level, many studies have used either exon arrays or high-throughput sequencing (see Boxes 1 and 2) to profile changes in RNA processing after various immune activations and stimulations. Upon T-cell activation, global profiling of alternative cassette exon usage identified 150–200 robust changes in exon usage—encompassing approximately 10% of known alternatively spliced exons expressed in T-cells (Ip et al., 2007; Martinez et al., 2012). These exons were enriched within genes related to both the immune response and cell-cycle regulation and were immediate responders upon activation (Ip et al., 2007). Strikingly, genes with activation-mediated alternative splicing were often distinct from genes with differential mRNA abundance despite sharing similar functional annotations, pointing to two independent regulatory programs affecting the transcriptome during immune activation (Ip et al., 2007). Similarly, a profile of human macrophages independently infected with either *Listeria monocytogenes* or *Salmonella typhimurium* also found that between 10 and 15% of alternatively processed exons expressed in macrophages change in usage upon infection, distributed across changes in all types of RNA processing events and enriched within genes with immune-related biological functions (Pai et al., 2016). Changes in 3' UTRs and RIs were the most prominent shifts, though the latter may be an indication of stress response similar to that observed during heat shock (described in detail below). Though genes with changes in isoform usage were also often differentially expressed, as many as 12–17% of genes expressed in macrophages change their isoform usage following infection without any detectable difference in gene expression levels (Pai et al., 2016).

Another RNA processing hallmark of both T-cell activation and bacterial infection is a shift towards usage of upstream PASs, leading to shorter 3' UTRs (Pai et al., 2016; Sandberg, Neilson, Sarma, Sharp, & Burge, 2008). Though this phenomenon is often associated with proliferative transitions (such as T-cell activation or cancer, Mayr & Bartel, 2009; Sandberg et al., 2008), prominent 3' UTR shortening after bacterial infection of macrophages—which are not proliferating—indicates that the immune response shares this regulatory strategy (Pai et al., 2016). In both T-cell activation and bacterial infection of macrophages, shorter 3' UTRs were regulated to exclude portions of the 3' UTR containing target sites for specific miRNAs (Pai et al., 2016; Sandberg et al., 2008). Thus, 3' UTR shortening in the immune response might allow key genes to evade repression by immune-activated miRNAs, which otherwise could downregulate or translationally repress these genes.

Despite the numerous RNA processing changes that occur following immune activation or stimulation, the mechanisms regulating the immune-mediated changes in exon, splice site, or polyadenylation site usage are yet unclear. Though it is

BOX 2

THE CHALLENGES OF MEASURING ALTERNATIVE RNA PROCESSING: HIGH-THROUGHPUT ADVANCES

The advent of high-throughput sequencing of cDNA (RNA-seq) led to an explosion in splicing-related research, especially for assessing the extent to which there are global RNA processing changes in varying cellular conditions. RNA-seq enabled the genome-wide assessment of changes in isoform usage, coupled with fine-scaled measurements of isoform levels and identification of novel isoform patterns. However, isoform analyses of RNA-seq data presented new computational and statistical challenges (Kakaradov, Xiong, Lee, Jojic, & Frey, 2012). This is epitomized by the multitude of algorithms that are still being designed to identify new isoforms, quantify their relative usage across multiple samples, and calculate statistical significance (Garber, Grabherr, Guttman, & Trapnell, 2011; Hooper, 2014; Liu, Lora, & Dickerson, 2014; Steijger et al., 2013). These algorithms can be broadly categorized into three classes: (a) quantifying expression levels and differential expression of entire isoforms (Bray, Pimentel, Melsted, & Pachter, 2016; Patro, Duggal, Love, Irizarry, & Kingsford, 2017; Trapnell et al., 2010), (b) quantifying exon usage and differential usage levels, based on locally informative reads (Katz et al., 2010; Vaquero-Garcia et al., 2016), and (c) quantifying sites of exon excision and differential splicing, based on split-junction reads spanning two noncontiguous transcript regions (Li et al., 2018; Vaquero-Garcia et al., 2016; Figure 3d). Despite the ongoing development of computational tools in this area and lack of consensus about the ideal statistical models, all of these approaches have yielded reproducible results and generated many testable hypotheses. Combining the advantages of classical single-gene assays and more recent high-throughput approaches, studies are now using high-throughput sequencing to screen causal variants for RNA processing changes in reporter assays (i.e., massively parallel reporter assays, Bogard, Linder, Rosenberg, & Seelig, 2018; Rosenberg, Patwardhan, Shendure, & Seelig, 2015) and assess the broader influences of *in vivo* mutagenesis of SREs or factors (using techniques such as CRISPR; Kapahnke, Banning, & Tikkanen, 2016; Mou et al., 2017; Yue & Ogawa, 2018).

possible that each specific stimulus might trigger a different regulatory pathway, such as the upregulation of *CELF2* mediated by JNK-pathway signaling following T-cell activation, it is more likely that there are some shared regulatory responses across stimuli that affect the transcriptome. For instance, most immune responses lead to widespread changes in transcriptional activation and, thus, the timing of transcriptional events, both of which could influence the regulation of downstream RNA processing. Studies investigating nascent RNA production following pro-inflammatory response in macrophages induced by Lipid A stimulation found that there was an increased rate of transcript production, coupled with a large increase in chromatin-associated nascent transcripts (Bhatt et al., 2012; Pandya-Jones et al., 2013). Intriguingly, these chromatin-associated transcripts were enriched for full-length transcripts that had not yet been spliced—especially at time points shortly after stimulation—suggesting that the limiting step for mature mRNA production following immune stimulation was the timing or rate of mRNA splicing (Pandya-Jones et al., 2013). Similarly, nascent RNA profiling across a time course of T-cell activation found that splicing was initially lagging or inhibited, with increased efficiency at later stages (Davari et al., 2017). Productive splicing was associated with kinase-dependent phosphorylation of PolII that likely promoted co-transcriptional splicing (Davari et al., 2017). Thus, the kinetics of RNA processing might play a large role in fine-tuning the waves of gene expression and transcriptional activity characteristic of the immune response.

4 | DIFFERENTIAL PROTEIN PHOSPHORYLATION AND AVAILABILITY REGULATES RNA PROCESSING UPON HEAT SHOCK

Delayed or inhibited mRNA splicing is also a key feature of RNA processing regulation during the cellular response to heat shock. The profound influence that heat shock has on splicing regulation was first discovered in flies, where unspliced pre-mRNAs of two factors—*Hsp83* (heat shock factor) and *Adh* (alcohol dehydrogenase enzyme)—accumulated following heat shock of *Drosophila melanogaster* cells (Yost & Lindquist, 1986). Broad inhibition of splicing after heat shock was later observed across many organisms and cell types, including yeast, trypanosomes, and HeLa cells (Bond, 1988; Yost & Lindquist, 1991; Yost, Petersen, & Lindquist, 1990). Severe heat shock in mammalian cells leads to widespread intron retention that causes nuclear retention of resulting transcripts (Shalgi, Hurt, Lindquist, & Burge, 2014). Introns in genes involved in protein folding or energy production are an exception and often escape this splicing inhibition (Shalgi et al., 2014), though heat shock also pauses global translational elongation (Shalgi et al., 2013). Introns that escape splicing inhibition tend to be co-transcriptionally spliced (Shalgi et al., 2014), suggesting heat-shock alters the timing or efficiency of splice site recognition and intron excision. The inhibition of splicing and widespread retention of transcripts may prevent production of aberrant proteins, reducing the burden on chaperone and proteasome machineries that are already highly burdened during heat shock. Instead, the stabilization of nuclear retained transcripts (as suggested by their detection at high levels) may serve as a reservoir of fully transcribed mRNA molecules that could later be rapidly spliced, exported and translated to recover protein levels upon relief from stress conditions.

Though the details of the mechanisms underlying splicing inhibition during heat shock are unknown, there is evidence that Serine-rich (SR) splicing factors may play a crucial role. The loading of SR proteins onto the CTD of PolII during transcription is often necessary for proper transcription elongation and downstream splicing. They are known to bind SREs in exons and introns, acting as either splicing activators or repressors depending on their binding context. One SR protein, SRSF10, is customarily bound by 14-3-3 proteins (Shi & Manley, 2007), whose dissociation upon heat shock allows de-phosphorylation of SRSF10 by the phosphatase PP1 (Shi & Manley, 2007), which itself is activated by the dissociation of phosphatase inhibitors (Shi & Manley, 2007). SRSF10 de-phosphorylation upon heat shock is broadly correlated with splicing inhibition, since dephosphorylated SRSF10 interacts with the U1 snRNP and interferes with U1 snRNP recognition of 5' splice sites (splice donors) (Shin, Feng, & Manley, 2004). Cells deficient for SRSF10 have increased temperature sensitivity and cannot recover after heat shock (Shin et al., 2004), which suggests a crucial role for SRSF10 in cell survival after heat-induced stress. As a group, SR proteins are broadly differentially phosphorylated after heat shock, due to heat-induced translocation of SR protein-specific kinases (i.e., SRPK1) from the cytoplasm to the nucleus (Zhong, Ding, Adams, Ghosh, & Fu, 2009). Thus, it is possible that much of the differential splicing and widespread splicing inhibition seen upon heat shock is regulated by differential protein phosphorylation triggered by stress signaling pathways. Furthermore, some SR proteins, hnRNPs, and other splicing factors have been shown to be sequestered into nuclear stress bodies during heat shock (Biamonti & Cáceres, 2009; Biamonti & Vourc'h, 2010), while others have remained soluble in the nucleus. Since the state of co- versus post-transcriptional splicing affects the fate of transcripts in heat shock, the ability of splicing factors to associate with chromatin under these conditions (and perhaps more broadly in cellular perturbation) may be critical. Thus, not only differential phosphorylation, but perhaps sequestration and/or availability of SRFs could regulate the specificity of heat-shock-induced splicing inhibition or escape from it.

Similarly, 3' cleavage and polyadenylation of mRNA molecules is also highly regulated in the context of stress responses. In heat shock, polyadenylation has been shown to be inhibited through the stress-induced posttranslational modification of Poly(A) polymerase (PAP) leading to abrogation of PAP activity (Di Giammartino, Shi, & Manley, 2013). More recently, transcriptional read-through has been shown to occur in a very widespread, albeit differentially regulated manner in response to heat shock and several other stress conditions (Vilborg et al., 2017; Vilborg, Passarelli, Yario, Tycowski, & Steitz, 2015). These patterns represent a failure of polymerases to terminate properly, most likely due to inhibition of the cleavage and polyadenylation machinery. Read-through transcripts may actually function to maintain open chromatin upon stress, suggesting a beneficial role for polyadenylation inhibition (Vilborg et al., 2017). This regulated defect in 3' end processing, which is strongly associated with stress response, is similar to observations of 3' UTR shortening seen in proliferating cells, immune response, and cancer cells (Mayr & Bartel, 2009; Pai et al., 2016; Sandberg et al., 2008).

Splicing factors may also be moonlighting as regulators of the DNA damage response induced by UV radiation. There are many transcriptome changes upon DNA damage, with a spate of apoptotic genes expressing alternative isoforms that have functions antagonistic to the pre-damage isoform (Muñoz et al., 2009). One famous example is the alternative splicing of the *BCL2L1* gene in humans upon DNA damage, where there is a splicing factor regulated shift away from the longer isoform that encodes an apoptotic inhibitor (BCL-xl) and increased expression of the shorter isoform that encodes an apoptotic activator (BCL-xs) (Shkreta, Toutant, Durand, Manley, & Chabot, 2016). In addition to their role in regulating RNA processing changes, splicing factors may also play a more direct, splicing-independent role in maintaining genome stability and repair, as reviewed in Naro, Bielli, Pagliarini, and Sette (2015). The dual role of splicing factors in regulating splicing and protection against DNA damage is likely facilitated by their colocalization and simultaneous interaction with both chromatin-wrapped DNA and nascent transcripts. This coordination may be enhanced by independent regulation of RNA PolIII activity upon DNA damage. One study investigating transcriptional dynamics after UV irradiation found that UV induces hyperphosphorylation of the CTD of RNA PolIII, which inhibits both transcription elongation and mRNA 3' end formation (Muñoz et al., 2009). Due to kinetic coupling between elongation rates and co-transcriptional splicing, UV-induced PolIII modifications are likely to directly regulate a portion of the alternative RNA processing seen during a DNA damage response.

5 | MUTATIONS IN THE SPLICING MACHINERY ARE A HALLMARK OF CANCER

Global transcriptional regulation has long been seen as a hallmark of cancer proliferation and tumor progression. Upregulation of the TF and proto-oncogene *MYC* during cancer progression results in an overall increase of the abundance of RNA molecules in a cell (Lin et al., 2012; Nie et al., 2012). The increase in pre-mRNA transcripts resulting from c-*MYC* upregulation may cause an undue burden on the splicing machinery (Munding, Shiue, Katzman, Donohue, & Ares, 2013), exemplified by the widespread intron retention commonly observed in tumors or cancerous cell lines. Abnormal splicing patterns in cancer include frequent retention of both alternative and constitutive introns (Dvinge & Bradley, 2015), with different sets of introns exhibiting alternative splicing in different cancer types. Notably, introns in genes encoding RNA splicing and export factors are frequently retained across many cancer conditions and often differentially expressed (Dvinge & Bradley, 2015), which could influence their trans-regulatory functions (i.e., Shapiro et al., 2011). Genes within these pathways have been previously shown to be more likely to be differentially expressed and spliced during heat shock (Shalgi et al., 2014). Similarly, cancer cells may also be hijacking the use of widespread intron retention (Shalgi et al., 2014), which is an intrinsic aspect of the heat shock response (the pathway that governs the response to many proteotoxic stress conditions). Consistent with these observations, the increase in *MYC* activity have also been linked to the specific upregulation of spliceosome components, core snRNP assembly genes, ribosomal proteins, and RNA binding protein genes (Koh et al., 2015), all of which are critical for the biogenesis of core snRNA and RNA-based regulatory machinery. In lymphomagenesis, *MYC* overexpression may increase the splicing fidelity for exons with weak 5' splice sites (splice donors), resulting in splicing defects underlying proliferative and anti-apoptotic phenotypes commonly seen in lymphoma progression (Koh et al., 2015). Finally, cancer cells exhibit a shift towards increased usage of upstream polyadenylation sites, leading to shorter 3' UTR regions (Mayr & Bartel, 2009). The shortening of 3' UTRs, a hallmark of proliferative cellular phenotypes and immune activation (Pai et al., 2016; Sandberg et al., 2008), may lead to oncogene activation with consequences for cancer progression (Mayr & Bartel, 2009). Thus, cancer progression pathways may be co-opting stress or immune response regulatory mechanisms as a way of adapting to constant growth and proliferation during stressful environmental perturbations.

In addition to the widespread dysregulation of transcriptional activity and resulting consequences on the transcriptome in cancer, there is growing consensus that cancer progression is aided by the accumulation of somatic mutations within cancerous cells (Dvinge, Kim, Abdel-Wahab, & Bradley, 2016; Martincorena & Campbell, 2015; Tomasetti, Li, & Vogelstein, 2017). One of the most recurrently mutated genes in cancer is *SF3B1* (Darman et al., 2015; Quesada et al., 2011), a key component of the spliceosome involved in branchpoint recognition (Gozani, Potashkin, & Reed, 1998). Mutated *SF3B1* often promotes usage of alternative 3' splice sites through the selection of aberrant branchpoints, resulting in tumor-specific RNA processing changes that often

target transcripts for downregulation by nonsense-mediated decay (Alsafadi et al., 2016; Darman et al., 2015). Mutations in another commonly mutated spliceosome component, *U2AF1*, similarly influence altered 3' splice site selection in cancerous cells and tumors, contributing to altered transcriptome patterns but not widespread splicing failure (Ilagan et al., 2015). Somatic *U2AF1* mutations have also been associated with tumor-specific alternative splicing of genes and regulatory RNAs involved in RNA processing (Shirai et al., 2015), including ribosomal RNAs, perhaps, indirectly influencing the transcriptome patterns of hundreds of downstream targets. Interestingly, the most prevalent cancer-associated mutations in *U2AF1* may also perturb the kinetic balance between the timing of splicing and 3' cleavage decisions, causing splicing to occur entirely posttranscriptionally after cleavage and release of the RNA from the chromatin (Coulon et al., 2014). More broadly, the preponderance of somatic mutations that accumulate in RNA processing genes in cancerous cells, including many within tissue-specific splicing factors (Dvinge et al., 2016; Kim et al., 2015), point to a coupling between proliferation-driven mutagenesis and misregulation of RNA processing in order to drive cancer pathogenesis. This relationship underscores the need to better characterize the genetic regulation of RNA processing and the impact of both cis- and trans-acting polymorphisms on RNA processing regulation.

6 | GENETIC REGULATION OF RNA PROCESSING

Genetic variation in noncoding DNA can influence several molecular processes and phenotypes. Quantitative trait locus (QTL) mapping of molecular phenotypes is a popular approach to identify genetic variants associated with interindividual variation in gene expression, chromatin accessibility and additional molecular traits—including RNA processing—collected through genomic and other high-throughput approaches. Although genetic studies of RNA processing are still limited, they are contributing important insights into two major areas: (a) understanding molecular mechanisms that are also used for RNA processing responses and (b) determining the genetic contribution to interindividual variation in RNA processing response, which has critical consequences in clinical phenotypes (e.g., response to drugs).

We can obtain an orthogonal assessment of the molecular mechanisms underlying RNA processing plasticity in different environmental contexts by considering interindividual variation of this molecular phenotype. Natural genetic variation can be considered nature's perturbations and often mimic mild to extreme environmental perturbations. This is the framework that was used to interrogate the hypothesis that differential TF binding can influence AFE usage (Richards et al., 2017) in response to treatment with selenium. QTL mapping on AFE usage (measured as percent-spliced-in values; AFE-QTL mapping) was performed across 200 individuals from the GEUVADIS study (Lappalainen et al., 2013) and demonstrated that AFE-QTLs in ELF2 motifs are enriched for single nucleotide polymorphisms (SNPs) computationally predicted to disrupt ELF2 binding.

Splicing QTLs (sQTLs) have been identified in humans and in model organisms (Gutierrez-Arcelus et al., 2015; Hasin-Brumshtein et al., 2016; Li et al., 2016; Ongen & Dermitzakis, 2015; Qu, Gurdziel, Pique-Regi, & Ruden, 2017; Saha et al., 2017). Analysis of alternative splicing in LCLs, fibroblasts and T-cells from 204 individuals showed that most splicing events are shared across cell types (up to 80%) and identified genetic variants correlated with interindividual variation in splicing (Gutierrez-Arcelus et al., 2015). Furthermore, a significant correlation was found between splicing and promoter DNA methylation signals across individuals. Regions with DNA methylation and/or sQTLs were also enriched for CTCF binding, providing support for a mechanism where methylation-sensitive CTCF binding affects alternative splicing (Shukla et al., 2011). These results were replicated in a subsequent study by Li et al. that performed sQTL mapping (in addition to QTL mapping for other molecular phenotypes) in a panel of 70 LCLs (Li et al., 2016). Using a computational method that analyzes split-reads indicative of splice junctions in short sequencing reads (see Box 2), this study identified almost 2,900 sQTLs, demonstrated that these QTLs are mostly independent from eQTLs (expression quantitative trait loci), and provided strong evidence that genetic variation can affect splicing by altering chromatin-level traits. Finally, this study showed that sQTLs detected in LCLs are enriched for genetic variants associated with auto-immune diseases—similar to enrichments previously observed among eQTLs—further supporting the potential importance of splicing misregulation in complex traits. The prevalence of natural variation that influences splicing decisions also indicates that cells are able to tolerate some variance in splicing outcomes, though the magnitude of differential exon usage between two genetic alleles is often lower than that exhibited after a strong cellular perturbation or misregulation (Figure 4). All of these studies specifically focused only on mRNA splicing (alternative exon or splice site usage) and did not consider other types of RNA processing events. Thus they were limited in their evaluation of the full potential for genetic variation to affect RNA processing and isoform composition.

Several studies have shown that genetic variants can alter molecular mechanisms that are important for response to environmental perturbations, thus contributing to interindividual variation in response phenotypes. Pharmacogenetic studies have highlighted the clinical importance of such genotype-by-environment interactions that can modify an individual's response to pharmacological treatment. The functional relevance of sQTLs has also been long recognized in the context of interindividual differences in response to pharmaceutical drugs (for a review of sQTLs and pharmacogenetics see Sadee et al., 2011). Despite the increasing interest in splicing as an important mechanism for human health, the number of studies analyzing genetic and

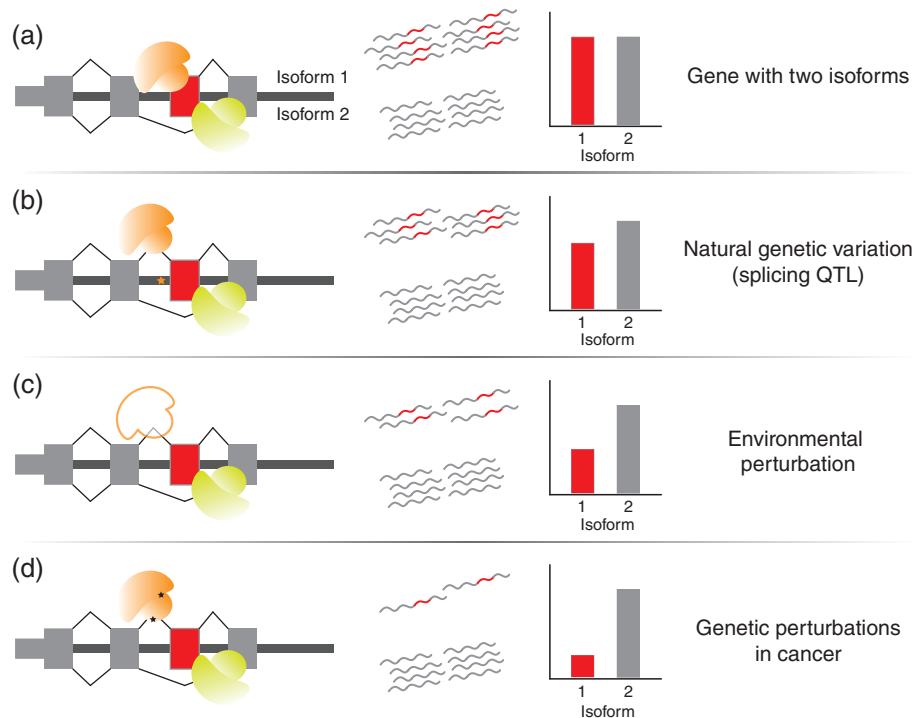


FIGURE 4 The magnitude of RNA processing changes across different contexts. (a) Any two isoforms for a given gene are often both expressed at high levels, such that an isoform excluding a cassette exon (gray) may be expressed at an equal proportion to an isoform including the exon (red). (b) Natural genetic variation impacting isoform composition often has a small but consistent effect upon relative isoform levels. Most QTLs detected with current methods are common variants acting in cis, which are presumed to be drifting under nearly neutral selection rather than strong selective pressures. (c) RNA processing responses to common environmental perturbations activating a cellular defense system (such as stimuli or stresses) often have moderate effect sizes. These changes are characterized by global shifts in the usage of given sites or exons, putatively regulated by differential expression or posttranslational modification of trans-regulatory factors. (d) RNA processing changes in proliferative cancer cells are often quite severe and widespread, coupled with and likely regulated by deleterious somatic mutations in core spliceosome machinery or splicing factors

environmental effects causing differential splicing is still limited, especially relative to eQTL mapping. This is even more evident when considering studies of response sQTLs. Response eQTL mapping studies have demonstrated that genetic variation can modulate gene regulatory responses to environmental perturbations, such that the regulatory potential of a variant is unveiled only within a specific environmental context. Response QTL mapping studies not only uncovered context-specific functional genetic variants, but also uncovered the importance of personalized genomic characterizations in the study of environmental risk for human health (e.g., Barreiro et al., 2012; Knowles et al., 2017, 2018; Mangravite et al., 2013; Maranville et al., 2011; Moyerbrailean et al., 2016; Nedelec et al., 2016; Taylor et al., 2018). Overall, our understanding of the role of genetic variation in modulating the RNA processing response to environmental perturbations is still quite limited. There are a limited number of studies that have analyzed sQTLs in contexts relevant for human health. In two recent studies (Nedelec et al., 2016; Ye et al., 2017) over 1,000 sQTLs were identified in dendritic cells and macrophages treated with agents that stimulate the immune system (Salmonella, Listeria, Influenza, IFN-beta). The functional relevance of these sQTLs is strengthened by an enrichment for signatures of positive selection among these sites and high overlap with SNPs associated with autoimmune and neurodegenerative diseases as assessed by genome-wide association studies. Another study identified over 100 cis-sQTLs in response to lead exposure in *D. melanogaster* heads (Qu et al., 2017). The study also identified a potential trans-sQTL hotspot that regulated a cluster of transcripts in response to lead, indicating that trans-acting polymorphisms may have large impacts on isoform composition in different environmental contexts.

7 | CONCLUSIONS

Eukaryotic transcriptomes are dynamic entities that react to the smallest changes in the cellular environment with rapid and consequential responses. It is clear that the interplay between transcriptional, RNA processing, and posttranslational regulatory mechanisms is at the heart of the cellular response. While recent studies have started to scratch the surface of characterizing environment-specific transcriptomes and how they regulate cellular function in various contexts, there are still many challenges remaining. One question that remains unaddressed is whether RNA processing response to a certain environmental

perturbation is cell type specific. The studies performed so far used different cell lines, therefore, confounding comparisons across RNA processing responses in the few cases where the same treatment was used for more than one cell type. With the advent of new high-throughput sequencing technology, it is easier to profile hundreds of changes in RNA processing events across many cell types and environments. However, we are often unable to predict which changes have a direct phenotypic impact, their consequences, and the mechanisms by which these effects occur.

Mapping natural genetic variants that modulate RNA processing responses (response QTLs) represents a promising avenue to characterize mechanisms underlying isoform variation and gain a deeper understanding of interindividual variation in the RNA processing response to environmental perturbations. Examples of such context-specific genetic effects on splicing have been identified for over 50 genes in dendritic cells treated with Influenza and IFN-beta (Ye et al., 2017). However, further studies with larger sample sizes and a wider range of environmental contexts are needed to paint a complete picture of these context-specific effects and their relevance for human health. As the cost of generating deep sequencing data continues to decrease, future efforts should focus on developing proper statistical and computational methods to robustly identify isoforms and quantify their abundance (Box 2). Current methods have many technical gaps—for instance, they often lack the ability to account for experimental replicates, have limited power for differentiation of full length-transcripts, and results are often difficult to reproduce or interpret, especially across platforms or statistical methods. While many of these caveats are starting to be solved for differential expression quantification and analysis, computational approaches to study RNA processing are lagging behind in this regard. The development and availability of such tools will enable the study of how environmental and genetic perturbations coordinately affect all aspects of RNA processing—beyond splicing and isoform expression—thus gaining a deeper understanding of cellular mechanisms relevant for health and disease in different environmental contexts.

Though it is clear that both *cis*- and *trans*-regulatory factors of RNA processing play a role in the coordinated response to environmental perturbations, we are still far from a complete characterization of the relative importance of such mechanisms. Results from sQTL mapping can help to identify novel regulatory sequences and therefore *cis*-regulatory mechanisms in the RNA processing response. Additionally, these studies can inform predictive computational models for regulatory sequences, similar to those already available to predict noncoding variants that regulate gene expression. Ultimately, such catalogs of regulatory variants influencing RNA processing response will contribute to the interpretation of association signals for complex traits in humans. These *cis*-regulatory sequences are the targets of *trans* factors that are often unknown. A full understanding of the mechanistic underpinnings of RNA processing response will require that in parallel to *cis*-regulatory sequences, *trans* factors are characterized by means of experimental assays. High throughput assays to identify DNA and RNA binding sites and factors are already available and have enabled studies of steady-state binding patterns in many cell lines and tissues (i.e., work of the ENCODE consortium and parallel efforts). Thus future studies should expand upon this work by focusing on performing these assays in a variety of cellular environments and contexts.

More broadly, it appears that variability in the gene regulatory response to environmental perturbations is reliant on a balance between *cis*- and *trans*-regulation. Differential usage of genetic loci influencing isoform composition (i.e., TSSs, splice sites, polyadenylation sites) is canonically thought of as being regulated locally, such that the decision made at a particular locus is independent of genome-wide decisions. However, it is increasingly clear that many changes are happening in a concerted fashion across many loci. For instance, context-specific shifts towards usage of upstream sites across genes or increased inclusion of exons indicate a global regulatory change. Previous studies have found that these changes may be due to: (a) changes in the stoichiometry of regulatory factors and their regulatory targets (Berg et al., 2012; Munding et al., 2013); (b) differential expression (or isoform usage) of RNA binding proteins that have downstream effects on RNA processing patterns (see above); and (c) signaling pathways that influence kinase activity to change protein phosphorylation of key RNA processing factors (see above). These observations underlie a need to better characterize how the transcriptional, RNA processing, and posttranscriptional or translational regulatory responses are all coordinated upon cellular perturbation. Furthermore, it is crucial to understand the timing at which regulatory decisions are occurring, since environmental stimuli may be directly acting on the speed or order at which RNA processing decisions are being made to regulate the final transcriptome. Finally, focusing on a genome-wide cellular milieu rather than individual events is crucial to delineate which variability has phenotypic consequences, rather than just being a consequence of cellular noise. Along these lines, emerging single-cell techniques for assaying gene expression and RNA processing (Hayashi et al., 2018; Shalek et al., 2013; Vu et al., 2018) will characterize these mechanisms at a higher resolution and may provide insights into the variability of cellular responses across a population of cells.

ACKNOWLEDGMENTS

We thank Eleni Jaeklein, Allison Richards, Roger Pique-Regi and members of the Luca and Pai groups for helpful discussions. This work was partially supported by funds from the University of Massachusetts Medical School (A.A.P.), NIH R01GM109215 (F.L.) and AHA 14SDG20450118 (F.L.).

CONFLICT OF INTEREST

The authors have declared no conflicts of interest for this article.

RELATED WIREs ARTICLES

[Modulation of alternative splicing by anticancer drugs](#)

[Alternative-splicing defects in cancer: Splicing regulators and their downstream targets, guiding the way to novel cancer therapeutics](#)

[Genetic variation of pre-mRNA alternative splicing in human populations](#)

ORCID

Athma A. Pai  <http://orcid.org/0000-0002-7995-9948>

Francesca Luca  <http://orcid.org/0000-0001-8252-9052>

FURTHER READING

- Li, W. V., & Li, J. J. (2018). *Modeling and analysis of RNA-seq data: A review from a statistical perspective*. arXiv. Retrieved from <https://arxiv.org/pdf/1804.06050.pdf>. This review details statistical considerations, limitations, and potential advancements for the analysis of RNA-seq data, across both gene expression, isoform, and localized RNA processing differential analyses.
- Pereria, M. A., Imada, E. L., & Guedes, R. L. M. (2017). RNA-seq: Applications and best practices. In Marchi (Ed.), *Applications of RNA-Seq and Omics Strategies: From microorganisms to human health*. IntechOpen. Retrieved from <https://www.intechopen.com/books/applications-of-rna-seq-and-omics-strategies-from-microorganisms-to-human-health/rna-seq-applications-and-best-practices>. This chapter works through applications of RNA-seq technology and experimental considerations while designing studies to measure RNA processing.

REFERENCES

- Alsafadi, S., Houy, A., Battistella, A., Popova, T., Wassef, M., Henry, E., ... Stern, M.-H. (2016). Cancer-associated SF3B1 mutations affect alternative splicing by promoting alternative branchpoint usage. *Nature Communications*, 7, 10615.
- Anvar, S. Y., Allard, G., Tseng, E., Sheynkman, G. M., de Klerk, E., Vermaat, M., ... 't Hoen, P. A. C. (2018). Full-length mRNA sequencing uncovers a widespread coupling between transcription initiation and mRNA processing. *Genome Biology*, 19(1), 46.
- Barbosa-Morais, N. L., Irimia, M., Pan, Q., Xiong, H. Y., Gueroussov, S., Lee, L. J., ... Blencowe, B. J. (2012). The evolutionary landscape of alternative splicing in vertebrate species. *Science*, 338(6114), 1587–1593.
- Barreiro, L. B., Tailleux, L., Pai, A. A., Gicquel, B., Marioni, J. C., & Gilad, Y. (2012). Deciphering the genetic architecture of variation in the immune response to mycobacterium tuberculosis infection. *Proceedings of the National Academy of Sciences of the United States of America*, 109(4), 1204–1209. <https://doi.org/10.1073/pnas.1115761109>
- Bentley, D. L. (2014). Coupling mRNA processing with transcription in time and space. *Nature Reviews Genetics*, 15(3), 163–175.
- Berg, M. G., Singh, L. N., Younis, I., Liu, Q., Pinto, A. M., Kaida, D., ... Dreyfuss, G. (2012). U1 snRNP determines mRNA length and regulates isoform expression. *Cell*, 150(1), 53–64.
- Bhatt, D. M., Pandya-Jones, A., Tong, A.-J., Barozzi, I., Lissner, M. M., Natoli, G., ... Smale, S. T. (2012). Transcript dynamics of proinflammatory genes revealed by sequence analysis of subcellular RNA fractions. *Cell*, 150(2), 279–290.
- Biamonti, G., & Caceres, J. F. (2009). Cellular stress and RNA splicing. *Trends in Biochemical Sciences*, 34(3), 146–153.
- Biamonti, G., & Vourc'h, C. (2010). Nuclear stress bodies. *Cold Spring Harbor Perspectives in Biology*, 2(6), a000695.
- Bogard, N., Linder, J., Rosenberg, A. B., & Seelig, G. (2018). *Predicting the impact of cis-regulatory variation on alternative polyadenylation*. bioRxiv. Retrieved from <https://doi.org/10.1101/300061>
- Bond, U. (1988). Heat shock but not other stress inducers leads to the disruption of a sub-set of snRNPs and inhibition of in vitro splicing in HeLa cells. *EMBO Journal*, 7(11), 3509–3518.
- Boutz, P. L., Bhutkar, A., & Sharp, P. A. (2015). Detained introns are a novel, widespread class of post-transcriptionally spliced introns. *Genes & Development*, 29(1), 63–80.
- Bray, N. L., Pimentel, H., Melsted, P., & Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology*, 34(5), 525–527.
- Brugiolo, M., Herzel, L., & Neugebauer, K. M. (2013). Counting of co-transcriptional splicing. *F1000Prime Reports*, 5, 9.
- Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y., & Greenleaf, W. J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nature Methods*, 10(12), 1213–1218.
- Coulon, A., Ferguson, M. L., de Turris, V., Palangat, M., Chow, C. C., Larson, D. R., & Black, D. L. (2014). Kinetic competition during the transcription cycle results in stochastic RNA processing. *eLife*, 3, e03939.
- Darman, R. B., Seiler, M., Agrawal, A. A., Lim, K. H., Peng, S., Aird, D., ... Buonamici, S. (2015). Cancer-associated SF3B1 hotspot mutations induce cryptic 3' splice site selection through use of a different branch point. *Cell Reports*, 13(5), 1033–1045.
- Davari, K., Lichti, J., Gallus, C., Greulich, F., Uhlenhaut, N. H., Heinig, M., ... Glasmacher, E. (2017). Rapid genome-wide recruitment of RNA polymerase II drives transcription, splicing, and translation events during T cell responses. *Cell Reports*, 19(3), 643–654.
- Di Giammartino, D. C., Shi, Y., & Manley, J. L. (2013). PARP1 represses PAP and inhibits polyadenylation during heat shock. *Molecular Cell*, 49(1), 7–17.
- Dvinge, H., & Bradley, R. K. (2015). Widespread intron retention diversifies most cancer transcriptomes. *Genome Medicine*, 7(1), 45.
- Dvinge, H., Kim, E., Abdel-Wahab, O., & Bradley, R. K. (2016). RNA splicing factors as oncoproteins and tumour suppressors. *Nature Reviews. Cancer*, 16(7), 413–430.
- Elkon, R., Ugalde, A. P., & Agami, R. (2013). Alternative cleavage and polyadenylation: Extent, regulation and function. *Nature Reviews. Genetics*, 14(7), 496–506.
- Frohman, M. A., Dush, M. K., & Martin, G. R. (1988). Rapid production of full-length cDNAs from rare transcripts: Amplification using a single gene-specific oligonucleotide primer. *Proceedings of the National Academy of Sciences of the United States of America*, 85(23), 8998–9002.

- Fu, X.-D., & Ares, M. (2014). Context-dependent control of alternative splicing by RNA-binding proteins. *Nature Reviews. Genetics*, 15(10), 689–701.
- Garber, M., Grabherr, M. G., Guttman, M., & Trapnell, C. (2011). Computational methods for transcriptome annotation and quantification using RNA-seq. *Nature Methods*, 8(6), 469–477.
- Gerstberger, S., Hafner, M., & Tuschl, T. (2014). A census of human RNA-binding proteins. *Nature Reviews. Genetics*, 15(12), 829–845.
- Gozani, O., Potashkin, J., & Reed, R. (1998). A potential role for U2AF-SAP 155 interactions in recruiting U2 snRNP to the branch site. *Molecular and Cellular Biology*, 18(8), 4752–4760.
- Gutierrez-Arcelus, M., Ongen, H., Lappalainen, T., Montgomery, S. B., Buil, A., Yurovsky, A., ... Dermitzakis, E. T. (2015). Tissue-specific effects of genetic and epigenetic variation on gene regulation and splicing. *PLoS Genetics*, 11(1), e1004958.
- Han, J., Xiong, J., Wang, D., & Fu, X.-D. (2011). Pre-mRNA splicing: Where and when in the nucleus. *Trends in Cell Biology*, 21(6), 336–343.
- Harlen, K. M., & Churchman, L. S. (2017). The code and beyond: Transcription regulation by the RNA polymerase II carboxy-terminal domain. *Nature Reviews. Molecular Cell Biology*, 18(4), 263–273.
- Hasin-Brumshtein, Y., Khan, A. H., Hormozdiari, F., Pan, C., Parks, B. W., Petyuk, V. A., ... Smith, D. J. (2016). Hypothalamic transcriptomes of 99 mouse strains reveal trans eQTL hotspots, splicing QTLs and novel non-coding genes. *eLife*, 5, e15614. <https://doi.org/10.7554/eLife.15614.002>
- Hayashi, T., Ozaki, H., Sasagawa, Y., Umeda, M., Danno, H., & Nikaido, I. (2018). Single-cell full-length total RNA sequencing uncovers dynamics of recursive splicing and enhancer RNAs. *Nature Communications*, 9(1), 619.
- Hermiston, M. L., Xu, Z., Majeti, R., & Weiss, A. (2002). Reciprocal regulation of lymphocyte activation by tyrosine kinases and phosphatases. *Journal of Clinical Investigation*, 109(1), 9–14.
- Hooper, J. E. (2014). A survey of software for genome-wide discovery of differential splicing in RNA-Seq data. *Human Genomics*, 8(1), 3.
- Hsin, J.-P., & Manley, J. L. (2012). The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes & Development*, 26(19), 2119–2137.
- Hu, R., Warri, A., Jin, L., Zwart, A., Riggins, R. B., Fang, H.-B., & Clarke, R. (2015). NF- κ B signaling is required for XBP1 (unspliced and spliced)-mediated effects on antiestrogen responsiveness and cell fate decisions in breast cancer. *Molecular and Cellular Biology*, 35(2), 379–390.
- Huang, Q., Liu, D., Majewski, P., Schulte, L. C., Korn, J. M., Young, R. A., ... Hacohen, N. (2001). The plasticity of dendritic cell responses to pathogens and their components. *Science (New York, N.Y.)*, 294(5543), 870–875.
- Ilagan, J. O., Ramakrishnan, A., Hayes, B., Murphy, M. E., Zebari, A. S., Bradley, P., & Bradley, R. K. (2015). U2AF1 mutations alter splice site recognition in hematological malignancies. *Genome Research*, 25(1), 14–26.
- Ip, J. Y., Tong, A., Pan, Q., Topp, J. D., Blencowe, B. J., & Lynch, K. W. (2007). Global analysis of alternative splicing during T-cell activation. *RNA (New York, N.Y.)*, 13(4), 563–572.
- Jacobsen, M., Hoffmann, S., Cepok, S., Stei, S., Ziegler, A., Sommer, N., & Hemmer, B. (2002). A novel mutation in PTPRC interferes with splicing and alters the structure of the human CD45 molecule. *Immunogenetics*, 54(3), 158–163.
- Jacobsen, M., Schweer, D., Ziegler, A., Gaber, R., Schock, S., Schwinzer, R., ... Hemmer, B. (2000). A point mutation in PTPRC is associated with the development of multiple sclerosis. *Nature Genetics*, 26(4), 495–499.
- Jan, C. H., Friedman, R. C., Ruby, J. G., & Bartel, D. P. (2011). Formation, regulation and evolution of *Caenorhabditis elegans* 3'UTRs. *Nature*, 469(7328), 97–101.
- Kakaradov, B., Xiong, H. Y., Lee, L. J., Jojic, N., & Frey, B. J. (2012). Challenges in estimating percent inclusion of alternatively spliced junctions from RNA-seq data. *BMC Bioinformatics*, 13(Suppl. 6), S11.
- Kapahnke, M., Banning, A., & Tikkanen, R. (2016). Random splicing of several exons caused by a single base change in the target exon of CRISPR/Cas9-mediated gene knockout. *Cell*, 5, 4.
- Katz, Y., Wang, E. T., Airolidi, E. M., & Burge, C. B. (2010). Analysis and design of RNA sequencing experiments for identifying isoform regulation. *Nature Methods*, 7(12), 1009–1015.
- Kim, E., Ilagan, J. O., Liang, Y., Daubner, G. M., Lee, S. C.-W., Ramakrishnan, A., ... Abdel-Wahab, O. (2015). SRSF2 mutations contribute to myelodysplasia by mutant-specific effects on exon recognition. *Cancer Cell*, 27(5), 617–630.
- Knowles, D. A., Burrows, C. K., Blischak, J. D., Patterson, K. M., Serie, D. J., Norton, N., ... Gilad, Y. (2018). Determining the genetic basis of anthracycline-cardiotoxicity by response QTL mapping in induced cardiomyocytes. *eLife*, 7, e33480. <https://doi.org/10.7554/eLife.33480.001>
- Knowles, D. A., Davis, J. R., Edgington, H., Raj, A., Fave, M. J., Zhu, X., ... Battle, A. (2017). Allele-specific expression reveals interactions between genetic variation and environment. *Nature Methods*, 14(7), 699–702.
- Kodzius, R., Kojima, M., Nishiyori, H., Nakamura, M., Fukuda, S., Tagami, M., ... Carninci, P. (2006). CAGE: Cap analysis of gene expression. *Nature Methods*, 3(3), 211–222.
- Koh, C. M., Bezzi, M., Low, D. H. P., Ang, W. X., Teo, S. X., Gay, F. P. H., ... Guccione, E. (2015). MYC regulates the core pre-mRNA splicing machinery as an essential step in lymphomagenesis. *Nature*, 523(7558), 96–100.
- Lai, W. K. M., & Pugh, B. F. (2017). Understanding nucleosome dynamics and their links to gene expression and DNA replication. *Nature Reviews. Molecular Cell Biology*, 18(9), 548–562.
- Lappalainen, T., Sammeth, M., Friedlander, M. R., 't Hoen, P. A., Monlong, J., Rivas, M. A., ... Hamosh, A. (2013). Transcriptome and genome sequencing uncovers functional variation in humans. *Nature*, 501(7468), 506–511.
- Lee, Y., & Rio, D. C. (2015). Mechanisms and regulation of alternative pre-mRNA splicing. *Annual Review of Biochemistry*, 84, 291–323.
- Li, Y. I., Knowles, D. A., Humphrey, J., Barbeira, A. N., Dickinson, S. P., Im, H. K., & Pritchard, J. K. (2018). Annotation-free quantification of RNA splicing using LeafCutter. *Nature Genetics*, 50(1), 151–158.
- Li, Y. I., van de Geijn, B., Raj, A., Knowles, D. A., Petti, A. A., Golan, D., ... Pritchard, J. K. (2016). RNA splicing is a primary link between genetic variation and disease. *Science*, 352(6285), 600–604.
- Lin, C. Y., Lovén, J., Rahl, P. B., Paranal, R. M., Burge, C. B., Bradner, J. E., ... Young, R. A. (2012). Transcriptional amplification in tumor cells with elevated c-Myc. *Cell*, 151(1), 56–67.
- Liu, R., Loraine, A. E., & Dickerson, J. A. (2014). Comparisons of computational methods for differential alternative splicing detection using RNA-seq in plant systems. *BMC Bioinformatics*, 15, 364.
- Lynch, K. W., & Weiss, A. (2000). A model system for activation-induced alternative splicing of CD45 pre-mRNA in T cells implicates protein kinase C and Ras. *Molecular and Cellular Biology*, 20(1), 70–80.
- Mallory, M. J., Allon, S. J., Qiu, J., Gazzara, M. R., Tapescu, I., Martinez, N. M., ... Lynch, K. W. (2015). Induced transcription and stability of CELF2 mRNA drives widespread alternative splicing during T-cell signaling. *Proceedings of the National Academy of Sciences of the United States of America*, 112(17), E2139–E2148.
- Mangravite, L. M., Engelhardt, B. E., Medina, M. W., Smith, J. D., Brown, C. D., Chasman, D. I., ... Krauss, R. M. (2013). A statin-dependent QTL for GATM expression is associated with statin-induced myopathy. *Nature*, 502(7471), 377–380.
- Maranville, J. C., Luca, F., Richards, A. L., Wen, X., Witonsky, D. B., Baxter, S., ... Di Rienzo, A. (2011). Interactions between glucocorticoid treatment and cis-regulatory polymorphisms contribute to cellular response phenotypes. *PLoS Genetics*, 7, e1002162. <https://doi.org/10.1371/journal.pgen.1002162>
- Martincorena, I., & Campbell, P. J. (2015). Somatic mutation in cancer and normal cells. *Science (New York, N.Y.)*, 349(6255), 1483–1489.

- Martinez, N. M., Agosto, L., Qiu, J., Mallory, M. J., Gazzara, M. R., Barash, Y., ... Lynch, K. W. (2015). Widespread JNK-dependent alternative splicing induces a positive feedback loop through CELF2-mediated regulation of MKK7 during T-cell activation. *Genes & Development*, 29(19), 2054–2066.
- Martinez, N. M., Pan, Q., Cole, B. S., Yarosh, C. A., Babcock, G. A., Heyd, F., ... Lynch, K. W. (2012). Alternative splicing networks regulated by signaling in human T cells. *RNA (New York, N.Y.)*, 18(5), 1029–1040.
- Martinez, N. M., & Lynch, K. W. (2013). Control of alternative splicing in immune responses: Many regulators, many predictions, much still to learn. *Immunological Reviews*, 253(1), 216–236.
- Mayr, C., & Bartel, D. P. (2009). Widespread shortening of 3UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell*, 138(4), 673–684.
- Medzhitov, R., & Horng, T. (2009). Transcriptional control of the inflammatory response. *Nature Reviews Immunology*, 9(10), 692–703.
- Medzhitov, R., & Janeway, C. A. (1998). Innate immune recognition and control of adaptive immune responses. *Seminars in Immunology*, 10(5), 351–353.
- Mele, M., Ferreira, P. G., Reverter, F., DeLuca, D. S., Monlong, J., Sammeth, M., ... Guigo, R. (2015). Human genomics. The human transcriptome across tissues and individuals. *Science*, 348(6235), 660–665.
- Merkin, J., Russell, C., Chen, P., & Burge, C. B. (2012). Evolutionary dynamics of gene and isoform regulation in mammalian tissues. *Science*, 338(6114), 1593–1599.
- Mou, H., Smith, J. L., Peng, L., Yin, H., Moore, J., Zhang, X.-O., ... Xue, W. (2017). CRISPR/Cas9-mediated genome editing induces exon skipping by alternative splicing or exon deletion. *Genome Biology*, 18(1), 108.
- Moyerbrailean, G. A., Davis, G. O., Harvey, C. T., Watz, D., Wen, X., Pique-Regi, R., & Luca, F. (2015). A high-throughput RNA-seq approach to profile transcriptional responses. *Scientific Reports*, 5, 14976.
- Moyerbrailean, G. A., Richards, A. L., Kurtz, D., Kalita, C. A., Davis, G. O., Harvey, C. T., ... Luca, F. (2016). High-throughput allele-specific expression across 250 environmental conditions. *Genome Research*, 26(12), 1627–1638.
- Munding, E. M., Shiue, L., Katzman, S., Donohue, J. P., & Ares, M., Jr. (2013). Competition between pre-mRNAs for the splicing machinery drives global regulation of splicing. *Molecular Cell*, 51(3), 338–348.
- Muñoz, M. J., Pérez Santangelo, M. S., Paronetto, M. P., de la Mata, M., Pelisch, F., Boireau, S., ... Kornblihtt, A. R. (2009). DNA damage regulates alternative splicing through inhibition of RNA polymerase II elongation. *Cell*, 137(4), 708–720.
- Naro, C., Bielli, P., Pagliarini, V., & Sette, C. (2015). The interplay between DNA damage response and RNA processing: The unexpected role of splicing factors as gatekeepers of genome stability. *Frontiers in Genetics*, 6, 142.
- Nasim, M. T., & Eperon, I. C. (2006). A double-reporter splicing assay for determining splicing efficiency in mammalian cells. *Nature Protocols*, 1(2), 1022–1028.
- Nedelec, Y., Sanz, J., Baharian, G., Szpiech, Z. A., Pacis, A., Dumaine, A., ... Barreiro, L. B. (2016). Genetic ancestry and natural selection drive population differences in immune responses to pathogens. *Cell*, 167(3), 657–669.
- Nie, Z., Hu, G., Wei, G., Cui, K., Yamane, A., Resch, W., ... Levens, D. (2012). C-Myc is a universal amplifier of expressed genes in lymphocytes and embryonic stem cells. *Cell*, 151(1), 68–79.
- Ongen, H., & Dermitzakis, E. T. (2015). Alternative splicing QTLs in European and African populations. *American Journal of Human Genetics*, 97(4), 567–575.
- Orengo, J. P., Bundman, D., & Cooper, T. A. (2006). A bichromatic fluorescent reporter for cell-based screens of alternative splicing. *Nucleic Acids Research*, 34(22), e148.
- Pai, A. A., Baharian, G., Pagé Sabourin, A., Brinkworth, J. F., Nédélec, Y., Foley, J. W., ... Barreiro, L. B. (2016). Widespread shortening of 3' untranslated regions and increased exon inclusion are evolutionarily conserved features of innate immune responses to infection. *PLoS Genetics*, 12(9), e1006338.
- Pandya-Jones, A., Bhatt, D. M., Lin, C.-H., Tong, A.-J., Smale, S. T., & Black, D. L. (2013). Splicing kinetics and transcript release from the chromatin compartment limit the rate of Lipid A-induced gene expression. *RNA (New York, N.Y.)*, 19(6), 811–827.
- Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., & Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods*, 14(4), 417–419.
- Pique-Regi, R., Degner, J. F., Pai, A. A., Gaffney, D. J., Gilad, Y., & Pritchard, J. K. (2011). Accurate inference of transcription factor binding from DNA sequence and chromatin accessibility data. *Genome Research*, 21(3), 447–455.
- Qu, W., Gurdziel, K., Pique-Regi, R., & Ruden, D. M. (2017). Identification of splicing quantitative trait loci (sQTL) in *Drosophila melanogaster* with developmental lead (Pb2+) exposure. *Frontiers in Genetics*, 8, 145.
- Quesada, V., Conde, L., Villamor, N., Ordóñez, G. R., Jares, P., Bassaganyas, L., ... López-Otín, C. (2011). Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. *Nature Genetics*, 44(1), 47–52.
- Richards, A. L., Watz, D., Findley, A., Alazizi, A., Wen, X., Pai, A. A., ... Luca, F. (2017). Environmental perturbations lead to extensive directional shifts in RNA processing. *PLoS Genetics*, 13(10), e1006995.
- Rosenberg, A. B., Patwardhan, R. P., Shendure, J., & Seelig, G. (2015). Learning the sequence determinants of alternative splicing from millions of random sequences. *Cell*, 163(3), 698–711.
- Sadee, W., Wang, D., Papp, A. C., Pinsonneault, J. K., Smith, R. M., Moyer, R. A., & Johnson, A. D. (2011). Pharmacogenomics of the RNA world: Structural RNA polymorphisms in drug therapy. *Clinical Pharmacology and Therapeutics*, 89(3), 355–365.
- Saha, A., Kim, Y., Gewirtz, A. D. H., Jo, B., Gao, C., McDowell, I. C., ... Zhu, J. (2017). Co-expression networks reveal the tissue-specific regulation of transcription and splicing. *Genome Research*, 27(11), 1843–1858.
- Sandberg, R., Neilson, J. R., Sarma, A., Sharp, P. A., & Burge, C. B. (2008). Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer micro-RNA target sites. *Science*, 320(5883), 1643–1647.
- Scotti, M. M., & Swanson, M. S. (2016). RNA mis-splicing in disease. *Nature Reviews Genetics*, 17(1), 19–32.
- Shalek, A. K., Satija, R., Adiconis, X., Gertner, R. S., Gaublot, J. T., Raychowdhury, R., ... Regev, A. (2013). Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. *Nature*, 498(7453), 236–240.
- Shalgi, R., Hurt, J. A., Krykbaeva, I., Taipale, M., Lindquist, S., & Burge, C. B. (2013). Widespread regulation of translation by elongation pausing in heat shock. *Molecular Cell*, 49(3), 439–452.
- Shalgi, R., Hurt, J. A., Lindquist, S., & Burge, C. B. (2014). Widespread inhibition of posttranscriptional splicing shapes the cellular transcriptome following heat shock. *Cell Reports*, 7(5), 1362–1370.
- Shapiro, I. M., Cheng, A. W., Flytzanis, N. C., Balsamo, M., Condeelis, J. S., Oktay, M. H., ... Gertler, F. B. (2011). An EMT-driven alternative splicing program occurs in human breast cancer and modulates cellular phenotype. *PLoS Genetics*, 7(8), e1002218.
- Shi, Y., & Manley, J. L. (2007). A complex signaling pathway regulates SRp38 phosphorylation and pre-mRNA splicing in response to heat shock. *Molecular Cell*, 28(1), 79–90.
- Shin, C., Feng, Y., & Manley, J. L. (2004). Dephosphorylated SRp38 acts as a splicing repressor in response to heat shock. *Nature*, 427(6974), 553–558.
- Shirai, C. L., Ley, J. N., White, B. S., Kim, S., Tibbitts, J., Shao, J., ... Walter, M. J. (2015). Mutant U2AF1 expression alters hematopoiesis and pre-mRNA splicing in vivo. *Cancer Cell*, 27(5), 631–643.
- Shkreta, L., Toutant, J., Durand, M., Manley, J. L., & Chabot, B. (2016). SRSF10 connects DNA damage to the alternative splicing of transcripts encoding apoptosis, cell-cycle control, and DNA repair factors. *Cell Reports*, 17(8), 1990–2003.

- Shukla, S., Kavak, E., Gregory, M., Imashimizu, M., Shutinoski, B., Kashlev, M., ... Oberdoerffer, S. (2011). CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature*, 479(7371), 74–79.
- Smale, S. T. (2010). Selective transcription in response to an inflammatory stimulus. *Cell*, 140(6), 833–844.
- Solier, S., Barb, J., Zeeberg, B. R., Varma, S., Ryan, M. C., Kohn, K. W., ... Pommier, Y. (2010). Genome-wide analysis of novel splice variants induced by topoisomerase I poisoning shows preferential occurrence in genes encoding splicing factors. *Cancer Research*, 70(20), 8055–8065.
- Somarelli, J. A., Schaeffer, D., Bosma, R., Bonano, V. I., Sohn, J. W., Kemeny, G., ... Garcia-Blanco, M. A. (2013). Fluorescence-based alternative splicing reporters for the study of epithelial plasticity in vivo. *RNA (New York, N.Y.)*, 19(1), 116–127.
- Steijger, T., Abril, J. F., Engström, P. G., Kokocinski, F., RGASP Consortium, Hubbard, T. J., ... Bertone, P. (2013). Assessment of transcript reconstruction methods for RNA-seq. *Nature Methods*, 10(12), 1177–1184.
- Stoilov, P., Lin, C. H., Damoiseaux, R., Nikolic, J., & Black, D. L. (2008). A high-throughput screening strategy identifies cardiotonic steroids as alternative splicing modulators. *Proceedings of the National Academy of Sciences of the United States of America*, 105(32), 11218–11223.
- Taylor, D. L., Knowles, D. A., Scott, L. J., Ramirez, A. H., Casale, F. P., Wolford, B. N., ... Collins, F. S. (2018). Interactions between genetic variation and cellular environment in skeletal muscle gene expression. *PLoS One*, 13(4), e0195788.
- Tian, B., & Manley, J. L. (2013). Alternative cleavage and polyadenylation: The long and short of it. *Trends in Biochemical Sciences*, 38(6), 312–320.
- Tomasetti, C., Li, L., & Vogelstein, B. (2017). Stem cell divisions, somatic mutations, cancer etiology and cancer prevention. *Science (New York, N.Y.)*, 355(6331), 1330–1334.
- Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., ... Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nature Biotechnology*, 28(5), 511–515.
- Tresini, M., Martelijn, J. A., & Vermeulen, W. (2016). Bidirectional coupling of splicing and ATM signaling in response to transcription-blocking DNA damage. *RNA Biology*, 13(3), 272–278.
- Tsalikis, J., Pan, Q., Tattoli, I., Maisonneuve, C., Blencowe, B. J., Philpott, D. J., & Girardin, S. E. (2016). The transcriptional and splicing landscape of intestinal organoids undergoing nutrient starvation or endoplasmic reticulum stress. *BMC Genomics*, 17, 680.
- Turner, M., & Díaz-Muñoz, M. D. (2018). RNA-binding proteins control gene expression and cell fate in the immune system. *Nature Immunology*, 19(2), 120–129.
- Vaquero-Garcia, J., Barrera, A., Gazzara, M. R., González-Vallinas, J., Lahens, N. F., Hogenesch, J. B., ... Barash, Y. (2016). A new view of transcriptome complexity and regulation through the lens of local splicing variations. *eLife*, 5, e11752. <https://doi.org/10.7554/eLife.11752.001>
- Vilborg, A., Passarelli, M. C., Yario, T. A., Tycowski, K. T., & Steitz, J. A. (2015). Widespread inducible transcription downstream of human genes. *Molecular Cell*, 59(3), 449–461.
- Vilborg, A., Sabath, N., Wiesel, Y., Nathans, J., Levy-Adam, F., Yario, T. A., ... Shalgi, R. (2017). Comparative analysis reveals genomic features of stress-induced transcriptional readthrough. *Proceedings of the National Academy of Sciences of the United States of America*, 114(40), E8362–E8371.
- Vu, T. N., Wills, Q. F., Kalari, K. R., Niu, N., Wang, L., Pawitan, Y., & Rantalainen, M. (2018). Isoform-level gene expression patterns in single-cell RNA-sequencing data. *Bioinformatics*, 34(14), 2392–2400.
- Wang, Z., & Burge, C. B. (2008). Splicing regulation: From a parts list of regulatory elements to an integrated splicing code. *RNA (New York, N.Y.)*, 14(5), 802–813.
- Xie, M., Li, M., Vilborg, A., Lee, N., Shu, M.-D., Yartseva, V., ... Steitz, J. A. (2013). Mammalian 5'-capped microRNA precursors that generate a single microRNA. *Cell*, 155(7), 1568–1580.
- Xu, Y., Wang, Y., Luo, J., Zhao, W., & Zhou, X. (2017). Deep learning of the splicing (epi)genetic code reveals a novel candidate mechanism linking histone modifications to ESC fate decision. *Nucleic Acids Research*, 45(21), 12100–12112.
- Ye, C. J., Chen, J., Villani, A.-C., Subramaniam, M., Gate, R. E., Bhangale, T., ... Hacohen, N. (2017). Genetic analysis of isoform usage in the human anti-viral response reveals influenza-specific regulation of erap2 transcripts under balancing selection. *bioRxiv*. Retrieved from <https://www.biorxiv.org/content/early/2017/10/31/188961> doi: <https://doi.org/10.1101/188961>
- Yost, H. J., & Lindquist, S. (1986). RNA splicing is interrupted by heat shock and is rescued by heat shock protein synthesis. *Cell*, 45(2), 185–193.
- Yost, H. J., & Lindquist, S. (1991). Heat shock proteins affect RNA processing during the heat shock response of *Saccharomyces cerevisiae*. *Molecular and Cellular Biology*, 11(2), 1062–1068.
- Yost, H. J., Petersen, R. B., & Lindquist, S. (1990). RNA metabolism: Strategies for regulation in the heat shock response. *Trends in Genetics*, 6(7), 223–227.
- Younis, I., Berg, M., Kaida, D., Dittmar, K., Wang, C., & Dreyfuss, G. (2010). Rapid-response splicing reporter screens identify differential regulators of constitutive and alternative splicing. *Molecular and Cellular Biology*, 30(7), 1718–1728.
- Yue, M., & Ogawa, Y. (2018). CRISPR/Cas9-mediated modulation of splicing efficiency reveals short splicing isoform of Xist RNA is sufficient to induce X-chromosome inactivation. *Nucleic Acids Research*, 46(5), e26.
- Zhong, X.-Y., Ding, J.-H., Adams, J. A., Ghosh, G., & Fu, X.-D. (2009). Regulation of SR protein phosphorylation and alternative splicing by modulating kinetic interactions of SRPK1 with molecular chaperones. *Genes & Development*, 23(4), 482–495.
- Zhou, R., Chun, R. F., Lisse, T. S., Garcia, A. J., Xu, J., Adams, J. S., & Hewison, M. (2015). Vitamin D and alternative splicing of RNA. *Journal of Steroid Biochemistry and Molecular Biology*, 148, 310–317.

How to cite this article: Pai AA, Luca F. Environmental influences on RNA processing: Biochemical, molecular and genetic regulators of cellular response. *WIREs RNA*. 2018;e1503. <https://doi.org/10.1002/wrna.1503>