Splicing: still so much to learn

ADRIAN R. KRAINER
Cold Spring Harbor Laboratory, Cold Spring Harbor, New York 11724, USA

This issue marks the 20th year since the launching of *RNA* in March of 1995. It is a time to celebrate, as well as to consider the substantial accomplishments achieved in these two decades, both in the RNA field as a whole, and in the specific areas in which each of us works—in my case, human pre-mRNA splicing mechanisms and regulation, and its relevance to human disease. Moreover, this is a time to reflect about where our field is headed, and what are the unique opportunities and obstacles in our path.

With no intention of being exhaustive in listing major achievements, I would first single out certain technical advances from which virtually everyone in the RNA field and beyond is substantially benefitting, and which satisfyingly emerged from careful characterization of novel natural phenomena. One of these is RNA interference, which has led to various powerful, widely used tools for targeted gene knockdown, as well as to major efforts in therapeutics development. A more recent example consists of the prokaryotic CRISPR/Cas systems, which are rapidly being adapted as powerful tools for targeted genome editing.

Other important technical developments include methods for massively parallel RNA-sequence analysis, and related techniques for systematically footprinting binding sites of proteins or RNP complexes on RNA (e.g., ribosome profiling, CLIP). These methods take advantage of the availability of complete or nearly complete genome sequences, and require appropriate use of computational and statistical tools.

Advances in the field of pre-mRNA splicing have continued steadily, and resulted in a comprehensive inventory of small RNAs and proteins involved in the various stages of spliceosome assembly, transesterification catalysis, and release of mature mRNA. Important insights have emerged concerning the interplay between splicing and other steps in mRNA biogenesis, including transcription and the various relevant features of chromatin, processing at the 5′ and 3′ ends, mRNA export and localization, and mRNA turnover. New complexes have been identified, notably the nuclear exosome and the exon-junction complex, with functionally important roles in RNA turnover and quality control.

Structure determination continues to provide crucial insights, and the splicing field awaits a breakthrough comparable to the high-resolution structures of ribosomes, which led to a renaissance in the field of translation. However, the dynamic nature of the spliceosome makes this especially challenging. In the meantime, many detailed structures of subassemblies and individual components of the spliceosome, or fragments thereof, have been obtained, allowing steady progress in elucidating structure-function relationships. Prominent examples include structures of the U1 snRNP and of a large fragment of PRP8.

The discovery and characterization of the minor spliceosome, which processes so-called U12-dependent introns, began after the inception of *RNA* (though its existence was predicted slightly earlier). Although only a tiny proportion of genes possess U12-dependent introns, the origin and evolution of parallel spliceosome pathways are very intriguing, comparisons with the major pathway have yielded insights for both pathways, and certain mutations in minor introns or in components of the minor spliceosome are disease-causing. Thus, this is a splicing pathway that deserves continued attention.

Characterization of the basic mechanisms and regulation of pre-mRNA splicing has enabled important advances in understanding the pathogenesis of various diseases, including genetic diseases, cancer, and infectious diseases, as well as created opportunities for therapeutics development. At the level of single-gene lesions, our understanding of which mutations cause missplicing has been augmented by insights into splicing-regulatory elements (enhancers and silencers), minor and non-canonical splice-site consensus sequences, and alternative base-pairing registers with snRNAs. Mutations in spliceosome components can also cause disease; for example, specific recurrent mutations in particular sets of components give rise to myelodysplastic syndrome and retinitis pigmentosa. In addition, sequestration of a regulatory splicing factor, MBNL, by an RNA-repeat expansion gives rise to myotonic dystrophy, and reduced levels of a snRNP-assembly factor, SMN, results in spinal muscular atrophy.

Mechanistic knowledge has enabled the development of targeted therapeutics that are now being tested in the clinic.

Corresponding author: krainer@cshl.edu
Article and publication date are at http://www.rnajournal.org/cgi/doi/10.1261/rna.050641.115. Freely available online through the *RNA* Open Access option.

For example, antisense/RNaseH-mediated cleavage is being used to destroy the *DMPK* mRNA that sequesters MBNL (myotonic dystrophy), and splicing modulation by an antisense oligonucleotide or by a small molecule is being used to restore correct *SMN2* pre-mRNA splicing and thereby increase the levels of SMN (spinal muscular atrophy). In addition, forcing exon skipping by antisense blocking of splice sites or splicing enhancers is being used to restore the correct translational reading frame in the context of *DMD* mutations in Duchenne muscular dystrophy, or to trigger nonsense-mediated mRNA decay and reduce gene expression.

Much work remains to be done to fully understand the logic of splicing, from the mechanistic details of how an individual intron is removed through the coordinated action of five snRNAs and hundreds of polypeptides, to the processing of numerous introns of widely different sizes from single pre-mRNAs, the high-fidelity recognition of >200,000 pairs of splice sites with substantial sequence variation, and the regulated choice of alternative splice sites in different contexts, as guided in part by the combinatorial effects of many splicing activators and repressors recognizing a diverse set of exonic and intronic enhancers and silencers. Such detailed knowledge will be essential to correctly interpret the information encoded in our genome, and to gain further insights into disease mechanisms. Multidisciplinary approaches, including various experimental and computational methods, are indispensable to effectively tackle these questions.

Studies involving RNA-seq yield massive datasets, whose analysis requires appropriate bioinformatic and statistical expertise. Interrogating the entire transcriptome can be extremely informative, and not surprisingly, this type of approach is becoming commonplace. However, in my opinion, rigorous standards have yet to be implemented for authors, referees, editors, and readers to be able to effectively present, evaluate, and understand this type of data. In addition, published high-throughput data generated for a specific purpose can be subsequently mined to address different questions, as long as these data meet appropriate standards for collection and reporting.

In the last 20 years, there have been substantial changes in how scientific findings are published. Increasingly, this is a source of much frustration for scientists. Many journals arbitrarily limit the length of the Methods section and the number of references allowed, even though these are crucial for others to be able to reproduce the findings and for scholarly presentation. Important details have to be relegated to an online supplement that often is not rigorously reviewed or edited, and which may not remain accessible indefinitely. Arbitrary editorial rejections, destructive reviews, deliberate delays, and unwarranted requests for additional experiments—without consideration of cost or time—have become unacceptably common. Regrettably, these trends slow down the progress of science and waste valuable resources. Yet, the strong pressure to publish in high-profile journals—associated with competition for limited research funds and with career advancement—exacerbates the problem.

Fortunately, *RNA* is one of a few quality journals that protect us from these destructive trends. The editors are active scientists with recognized expertise in their respective areas, and personal knowledge of researchers in the field, which greatly helps to keep the review process honest and efficient, and to prevent or manage potential conflicts. We should be grateful for their selfless dedication, which has made possible the first 20 years of *RNA*'s informative publications on all aspects of RNA science.

Splicing: still so much to learn

Adrian R. Krainer

RNA 2015 21: 500-501