

---

# Finding the unexpected—how we identified a second class of introns and the U12-dependent spliceosome

---

**RICHARD A. PADGETT**

Department of Molecular Biology, Lerner Research Institute, Cleveland Clinic Foundation, Cleveland, Ohio 44195, USA

One of the RNA stories that was unfolding at the time the *RNA* journal was starting up was the discovery that there were two independent types of pre-mRNA introns in some metazoans including humans. As with many unexpected findings in science, this one was as much accidental as deliberate.

My first encounter with what has come to be called the U12-dependent or minor intron class was sometime in 1989 when it was brought into my office by a UT Southwestern colleague, Alan Duby, who had cloned and sequenced the human *MATN1* gene. He showed me the splice site sequences of the last intron in this gene and asked my opinion of them. They were, indeed, very odd, having unusual terminal dinucleotides of AT and AC as well as almost no other similarities to normal splice sites. I suggested initially that this was a cloning artifact but he had a partial sequence from the orthologous chicken gene showing strangely identical presumptive splice sites. Alan's 1990 *Journal of Biological Chemistry* paper on the characteristics of this gene is, to my knowledge, the first to suggest that there might be a second type of intron. A single example is hard to work with but a new graduate student, Steve Hall, and I checked some obvious guesses using Alan's intron sequence but found no sequence signatures of Group I or Group II self-splicing introns or anything else remarkable that would help us understand what we had.

Matters largely sat there until an early bioinformatic paper from Ian Jackson appeared in 1991 in which he identified a variety of non-consensus splice sites from sequence databases. To our delight, he had uncovered a second example of an intron with splice sites identical to the *MATN1* intron; this time in the human P120 gene (now known as *NOP2*). He had also found the human *MATN1* gene intron from Alan's work and, comparing the two, suggested that they might represent a new class of introns but did not pursue this further. What struck us with the human P120 intron was that it was about 100 nucleotides long—much shorter than the *MATN1* intron and suitable for PCR amplification

across it. We designed a pair of degenerate sequence primers in the two flanking P120 exons and went around to nearby labs collecting cell lines from various species from which to amplify the introns. We were able to amplify, clone and sequence DNA from several mammalian species which showed that the 5' splice site was perfectly conserved for 8 nucleotides which was also identical to the two *MATN1* 5' splice sites, while the 3' splice site not only had a conserved CAC sequence but also an identical 8 nucleotide sequence at a variable distance upstream of the CAC. These features strongly suggested that these were splice site consensus elements but were very different from those known at the time.

So now we had two phylogenetically conserved examples of unusual splice sites. Work in yeast published around this time had shown that a conventional intron could be spliced in vivo when the terminal dinucleotides were changed from GU-AG to AU-AC. All of the *MATN1* and P120 introns were AU-AC so we wondered were these like the yeast case or were they a different class entirely? The highly conserved sequences of the novel splice sites and their differences from the known consensus for regular introns suggested to us that this was a new class, but we needed more evidence.

One way we tried to find more examples to bolster our case was through a poster at an RNA meeting showing our sequences. I dragged as many splicing colleagues as I could to the poster and asked them if they had seen any similar cases. After the meeting, I received calls from two friends who said they had been asked by others in their institutions about strange introns that they thought might be similar. When I contacted these individuals and asked them to read the sequences at the ends of their introns to me, I was thrilled to see that these matched our other introns perfectly. This way, we learned that the *Drosophila prospero* gene and the mouse *Msh3* gene had introns with similar splice sites.

During this time, we were also wondering about the mechanism that spliced these very non-consensus introns. Recalling that the initial 1980 papers suggesting that the newly discovered introns of eukaryotes were somehow recognized by base pairing with snRNAs such as U1 was based only on

---

**Corresponding author:** padgetr@ccf.org

Article and publication date are at <http://www.rnajournal.org/cgi/doi/10.1261/rna.050245.115>. Freely available online through the *RNA* Open Access option.

© 2015 Padgett This article, published in *RNA*, is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

sequence, we looked for sequences in snRNAs that could base pair with our presumptive splice site sequences. We found nothing convincing in the known spliceosomal snRNAs. However, Joan Steitz's lab had sequenced additional members of the snRNA class including RNAs named U11 and U12. Our first success was with the conserved intronic 3' element that we thought might represent the branch site sequence of these unusual introns. In conventional introns, this sequence can base pair to a region of U2 snRNA such that the branch site adenosine is unpaired (bulged out) within the U2-branch site helix. When we looked at U12 snRNA, we found a region of complementary sequence that would make a similar structure with our element, at least on paper. Other work from the Steitz lab had shown that U12 snRNP associates with the U11 snRNP and noted that these two snRNAs had secondary structures similar to U2 (for U12 snRNA) and U1 (for U11 snRNA). When we looked at the sequence of U11 snRNA, we found that the 5' end could form 5 contiguous base pairs with the putative 5' splice site sequence. Amazingly, this was very similar to what U1 snRNA was known to do. These findings led us to propose that these unusual introns were spliced with the help of U11 and U12 snRNPs thus providing functions for these orphan snRNAs and an explanation for how these splice sites were identified.

We wrote up these results and sent the paper out only for it to be rejected on the basis that we had no proof for any of these ideas. We were told that these odd introns were likely just natural examples of the AU-AC splice sites shown to be functional in yeast. It was very clear that the bar was now set higher than when the U1 hypothesis was made a decade prior. On the other hand, we were well aware that we were proposing a fairly radical idea, akin to suggesting that some mRNAs were translated using an alternative genetic code by a second type of ribosome. We were, of course, trying to produce direct evidence but our attempts to get the P120 intron to splice *in vitro* were unsuccessful as were attempts to detect splicing in injected *Xenopus* oocyte nuclei. We continued shopping our paper around and eventually got it accepted in *Journal of Molecular Biology* in 1994.

Since our *in vitro* experiments were unproductive, we switched to an *in vivo* approach to test our U12 hypothesis. A powerful genetic method to probe RNA-RNA interactions is to show that mutations causing mis-pairing in one RNA partner can be reversed or suppressed by compensatory mutations in the other RNA partner that restore base pairing (but not by similar mutations that do not restore pairing). This had been previously done for the interactions of U1 and U2 snRNAs with splice sites in both yeast and mammals. So Steve set out to do this experiment in cell culture using a

minigene construct from the P120 gene and an engineered version of the U12 snRNA gene.

He was able to demonstrate that RNA from the transfected wild type P120 minigene was spliced correctly while a single mutation in the 3' conserved element blocked splicing. He further showed that co-transfecting a wild type U12 gene had no effect on the P120 mutant but a U12 gene with a compensatory mutation in the site predicted by our model restored splicing. Alleles that did not restore base pairing were inactive. These results finally provided direct genetic evidence for the function of U12 snRNA in the splicing of these non-consensus introns. During this same period of time, Woan-Yuh Tarn in the Steitz lab had been working with the P120 intron *in vitro*. She was able to solve the problem of getting it to splice by recognizing that the RNA transcript was being bound by the abundant U2 snRNPs. Blocking U2 snRNA allowed splicing by U11 and U12 snRNPs. The work from both labs was published in 1996, proving the U11/U12 hypothesis and establishing that these introns were spliced by an alternative spliceosomal mechanism.

Woan-Yuh and Joan went on to discover two completely new snRNAs required for *in vitro* splicing of the P120 intron that they named U4atac and U6atac, and showed that U5 snRNA was required to splice both classes of introns. While early literature often referred to these non-consensus introns as ATAC introns due to the AU-AC dinucleotides that originally made them stand out, we now know that both AU-AC and GU-AG classes exist which has led to calling them U12-dependent or U12-type introns to distinguish them from the more common U2-dependent or U2-type introns.

In the succeeding two decades we have learned a great deal about these introns with many of the papers appearing in *RNA*. We now recognize that there are roughly 700 U12-type introns in almost as many human genes (which contain U2-type introns as well). Phylogenetic studies suggest that both classes of introns were present in the earliest common ancestor of all eukaryotes but that the U12-type introns have been lost in many independent lineages. Careful comparisons of the two splicing systems have provided insights into the mechanism and regulation of splicing including hints that U12-type introns may play roles in cellular processes at the systems level.

In closing, I would like to acknowledge the help of the members of my own lab as well as the larger RNA community. I hope this story shows how much this community became part of the process of discovery. Finally, on a sad note, Steve Hall, who did so much of the early work on these introns, became ill prior to completing his dissertation and died of AIDS shortly after his work was published.