Alternative splicing is a mechanism by which pre-mRNAs are processed to allow a single gene to encode multiple proteins, contributing substantially to the diversity of proteins in most eukaryotic cells. Alternative splicing is thought to be frequently dysregulated in cancer, and the potential of splicing to affect hundreds or thousands of genes suggests that improper splicing could be a driver of carcinogenesis (David and Manley, 2010). Common mutations in the splicing machinery have been identified in hematological malignancies including leukemias, but their actual effects on splicing have been unclear. Postdoctoral fellow Janine Ilagan and colleagues in the laboratory of Dr. Robert Bradley (Public Health Sciences and Basic Sciences Divisions), Dr. Aravind Ramakrishnan (Clinical Research Division), and Dr. Phil Bradley (Public Health Sciences) investigated the effects of specific mutations in the splicing factor U2AF1. U2AF1 mutations led to changes in the splicing of hundreds of genes, many involved in biological pathways dysregulated in cancer. Notably, they found that different U2AF1 mutations gave rise to distinct splicing outcomes, consistent with structural predictions.

U2AF1 mutations are very specific, affecting only codons S34 and Q157, which fall within the first and second zinc finger domains of the protein. The authors thus reasoned that a comprehensive study of the effects of U2AF1 was feasible. They first undertook an analysis of splicing in transcript data from acute myeloid leukemia (AML) patients with U2AF1 mutations and found that these mutations led to similar patterns of altered splicing. To experimentally test the effect of U2AF1 mutations on splicing, the authors generated erythroleukemia cell lines expressing wild-type or mutant forms of U2AF1 and performed RNA sequencing (RNA-seq). This revealed alterations in thousands of splicing events, though many of these were observed with low frequency. To identify targets of U2AF1 that might contribute to pathogenesis, the authors identified alterations in splicing in common between their cell line models and AML patient samples. This approach revealed 54 splicing alterations, many of which were in genes that have been previously implicated in myeloid disorders. These included the DNA methyltransferase DNMT3B, the Polycomb protein ASXL1, the macroH2A histone variant, the DNA damage-activated kinase ATR, and the apoptosis-related protein CASP8.

The authors next investigated how U2AF1 mutations might alter splicing. U2AF1 recognizes the nucleotide sequence yAGr (where y is C or U and r is A or G) within the 3′ splice site, and so they
hypothesized that mutations in U2AF1 might alter its sequence specificity. To test this, they identified consensus 3’ splice sites in AML patient samples carrying U2AF1 mutations. This analysis revealed different consensus nucleotides at the positions flanking the AG in the yAGr sequence, suggesting that U2AF1 mutations do indeed alter its sequence recognition properties. Further analysis showed that alterations in the left flanking position of the 3’ splice site (y in yAGr) were associated with S34 mutations, while changes in the right position (r in yAGr) were mainly found in conjunction with Q157 mutations. In vitro analysis confirmed that U2AF1 mutants in fact have an altered 3’ splice site sequence preference, and computational predictions of the structure of U2AF1 in complex with RNA indicated that S34 and Q157 are in proximity to the left and right flanking positions of the 3’ splice site, respectively, indicating a structural basis for altered sequence specificity in U2AF1 mutants.

Taken together, the data presented in this paper show that U2AF1 mutations alter splicing in patient samples, cultured cells, and in vitro, demonstrating the potential importance of altered splicing to cancer. "Surprisingly, our study suggests that even subtle changes in normal splice site recognition can drive tumorigenesis," said Dr. Bradley. "The next step is to determine how specific mis-spliced genes disrupt normal hematopoiesis."


Structural overview of a theoretical model of the U2AF1:RNA complex. The zinc finger domains are colored cyan and the RNA is colored salmon. The frequently mutated positions S34 and Q157 are shown in stick representation.

Image provided by Dr. Robert Bradley