Submitter Name: Lindsay M Payer
Submitted email: lindsaypayer@jhmi.edu
PI Name (if different): Kathleen H Burns
PI email (if different): kburns@jhmi.edu

Alu retrotransposon: A common structural variant that can regulate transcript level and structure leading to disease risk

Lindsay M Payer[1], Jared P Steranka[1], and Kathleen H Burns[1,2]

[1]Department of Pathology, Johns Hopkins University School of Medicine; [2]McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine

Alu elements are the most common mobile DNA and the most polymorphic accounting for ~17% of structural variants. However, the functional repertoire of *Alu* polymorphisms has not been fully evaluated. We systematically identified functional *Alu* variants and evaluated their contribution to disease risk. We mapped 809 previously published *Alu* variants to linkage disequilibrium blocks (r^2>0.8) with GWAS (Genome-Wide Association Studies) signals, a significant enrichment of *Alu* variants at GWAS loci (p=0.013). We propose that a subset of these variants modulate disease risk. We identified two mechanisms by which they may alter gene regulation; *Alu* variants can alter transcript level or structure. Using standard luciferase assays, we tested 110 loci for effects on transcript level and observed a continuum of effects. To validate, we focused on 4 *Alu* associated with breast cancer risk with significant effects on luciferase. Using CRISPR, we established isogenic lines varying only in *Alu* genotype. Consistent with our luciferase assays, these data reveal modulation of transcript levels of adjacent genes, including *MYC*. To test the effect of *Alu* variants on transcript structure, we used a minigene reporter assay to test 5 *Alu* variants that map near exon-intron boundaries by measuring exon incorporation. Three *Alu* altered exon incorporation rates. The *CD58 Alu* is of particular interest with its hypothesized role in multiple sclerosis risk. Collectively, these data demonstrate that, as a class, *Alu* polymorphisms commonly alter transcript levels and content, highlighting their role in epigenetic regulation and potential contribution to disease risk which could include substance use disorders.