Several examples of gene regulation via transposable element (TE) exaptation have been reported in various murine cell types. Here binding sites for transcription factors (TFs) key to the regulation of the circadian clock in mouse liver cells were examined to determine the extent to which TEs have impacted the circadian regulatory network. Published chromatin immuno-precipitation sequencing (ChIP-seq) data for the peak summits of 6 essential regulators, BMAL1, CLOCK, CRY1, CRY2, PER1, and PER2, enabled intersections with the RepeatMasker library of all mouse TEs. From these intersections, we determined that at least 15% of each factors’ binding sites directly overlap TEs. These results were shown to be of increased significance as independently generated ChIP-seq data, DNaseI hotspots and H3K27ac sites corroborated functional binding. Examination of gene ontology data and confirmation of the binding of multiple transcription factors (TF) added further weight to the functional significance of these binding sites.

Of the TE derived binding sites, a significant fraction emerged from a particular SINE subfamily, RSINE1. It appears that these RSINE1 elements independently developed the E-boxes commonly associated with circadian regulator binding. However while the presence of an E-box in open chromatin is usually necessary for binding, it is not sufficient. Further motif analysis suggests that RSINE1 elements are uniquely poised for binding as they contain a number of additional motifs, confirmed to provide binding sites for downstream regulatory proteins also implicated in the circadian clock. Additionally, these binding sites are closer to other non-TE sites than one would expect by chance. Our results expand upon the model of epistatic capture, in which a TE with a pre-existing motif gains secondary, tissue-specific TF sequences after insertion. To this model, we add details about the environmental effects, which seemingly prevent or encourage epistatic capture.