Identifying exon boundaries in novel or non-model species is a challenging Bioinformatics problem. Being able to predict the exon boundaries in the genes of non-model species is important given the quickly growing numbers of transcriptome sequencing projects that are identifying SNPs. In order to use such *in silico* SNP predictions on genomic DNA samples from novel or non-model species, accurate exon boundary predictions are needed. As the number of sequenced genomes increases across diverse taxa, one pattern to emerge is the shared gene order along chromosomes, as well as the shared exon boundaries across orthologous genes. Given this conservation of exon boundaries, comparative genomics provides a route for extending identified exon boundaries species having whole genome sequence to non-model species having assembled transcriptomes.

CEPiNS is a novel pipeline for mining exon boundaries of predicted gene sets from species with genomic sequence, and then using this information to identify the exon boundaries in the cDNA of a novel species through codon based alignment. The pipeline uses the freeware SPIDEY, which is an exon boundary prediction tool using genomic and cDNA sequence and is part of NCBI’s toolkit, and BLAST (blastn, blastp, tblastx), which finds protein-coding nucleotide sequence regions based on their corresponding amino acid translations. To demonstrate the performance of CEPiNS, the genome and predicted gene set from *Drosophila melanogaster* were used to predict the exon boundaries in the predicted gene set from *D. willistoni*, using these as a proxy for the contigs of an assembled transcriptome. These two species are approximately 40 million years divergent. Then these exon boundaries were verified in *D. willistoni* using that species whole genome sequence data.

CEPiNS provides an important comparative genomics tool necessary for developing SNPs, identified in transcriptome sequencing of novel species, which will work on genomic DNA.