

# Simulate Exon Skipping Frequencies

I am using your FLUX Simulator to generate RNA-Seq data. Because I am investigating differential exon usage, I hope to simulate, say among the genes in a real dataset, there are 1,000 genes have differential exon usages between two conditions. I only consider the simplest case of exon skipping, and the following files are available:

1. an annotation GTF file
2. BAM file for each sample (from TopHat)
3. BED file for each sample (from TopHat)

I studied the FLUX Simulator pages and examples, but I am not sure how to “pick up” 1,000 genes (from these files available) and somehow modify them to contain differential exon usages (e.g. exon skipping in condition 1 only). I think what I need eventually are one GTF file, BAM file for each sample, and BED file for each sample. The output of FLUX Simulator includes .BED and .FASTQ which I believe are relevant to what I need.

Thank you in advance for your suggestions! I appreciate if you could provide a toy example for this situation (e.g. .par file). Thanks!